

Aire unleashes stalled RNA polymerase to induce ectopic gene expression in thymic epithelial cells

Matthieu Giraud^{a,1}, Hideyuki Yoshida^a, Jakub Abramson^{a,2}, Peter B. Rahl^b, Richard A. Young^{b,c}, Diane Mathis^{a,3}, and Christophe Benoist^{a,3}

^aDivision of Immunology, Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA 02115; and ^bWhitehead Institute for Biomedical Research and ^cDepartment of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Contributed by Christophe Benoist, November 28, 2011 (sent for review October 31, 2011)

Aire is a transcriptional regulator that induces expression of peripheral tissue antigens (PTA) in thymic medullary epithelial cells (MECs), driving immunological self-tolerance in differentiating T cells. To elucidate its mechanistic pathways, we examined its transcriptional impact in MECs in vivo by microarray analysis with mRNA-spanning probes. This analysis revealed initiation of Aire-activated genes to be comparable in Aire-deficient and wild-type MECs, but with a block to elongation after 50–100 bp in the absence of Aire, suggesting activation by release of stalled polymerases by Aire. In contrast, patterns of activation by transcription factors such as Klf4 were consistent with regulation of initiation. Mapping of Aire and RNA polymerase-II (Pol-II) by ChIP and high-throughput sequencing (ChIP-seq) revealed that Aire bound all Pol-II-rich transcriptional start sites (TSS), irrespective of its eventual effect. However, the genes it preferentially activated were characterized by a relative surfeit of stalled polymerases at the TSS, which resolved once Aire was introduced into cells. Thus, transcript mapping and ChIP-seq data indicate that Aire activates ectopic transcription not through specific recognition of PTA gene promoters but by releasing stalled polymerases.

autoimmunity | thymus

The capacity of T lymphocytes to react to a virtually infinite spectrum of foreign antigens carries as a corollary potential reactivity to self-constituents and, consequently, autoimmunity. Self-reactive T cells are eliminated during maturation in the thymus, requiring that gene products normally expressed only by parenchymal cells of specialized organs be made visible to differentiating thymocytes. The solution to this conundrum seems to be, at least in part, the specialized ability of thymic stromal cells, in particular medullary epithelial cells (MECs), to transcribe a large repertoire of genes encoding these peripheral tissue antigens (PTAs) (1). This ectopic gene expression is controlled in a large part by the Aire transcription factor, which is expressed almost exclusively in highly differentiated MECs. Mice with an *Aire* gene defect express only a fraction of the PTA repertoire (2) and develop immune infiltrates and autoantibodies directed at multiple peripheral tissues, as do human patients with a mutated *AIRE* gene (3).

Aire is an unusual transcription factor, and a number of observations suggest that it does not function as a classic transactivator. Aire affects the expression of thousands of genes in MECs, which are regulated by very different pathways in peripheral parenchymal tissues (4), with a strong stochastic element as to which individual cell and which of the two chromosomes actually expresses a given PTA (5, 6). This regulator does not have a clear DNA binding motif, and its transcriptional footprint is highly dependent on the cell type in which it is expressed (7, 8). Aire binds chromatin through an interaction between its PHD1 domain and the amino-terminal tail of histone H3, but only in its hypomethylated state, a mark associated with poorly transcribed genes (9–12).

The precise molecular mechanisms that Aire uses to regulate transcription remain elusive. A large-scale screen based on mass

spectrometry of coimmunoprecipitated proteins revealed Aire's connection to factors involved in a number of nuclear processes: chromatin structure/modification, transcriptional elongation, pre-mRNA processing, and nuclear transport (8). A systematic and large-scale RNAi screen of Aire's transcriptional allies identified a number of previously unrecognized elements of the Aire pathway, with a preponderance of factors involved in transcriptional elongation rather than initiation, consistent with work showing an effect of Aire on elongation in transfection experiments, and an interaction of Aire with P-TEFb (positive transcription elongation factor b) (13).

Elongation factors might be needed for processive elongation of Pol-II through Aire target genes, or because of an effect of Aire on promoter-proximal Pol-II stalling. Also referred to as poised-polymerase, paused-polymerase, abortive elongation premature termination, this mode of transcriptional regulation was initially thought to apply to special loci with very fast responses to inducers (14, 15) but was more recently recognized to be widespread (16–18): initiation occurs constitutively, Pol-II proceeds for 40–50 bp, but is then blocked by the action of dominant pause factors such as DSIF (DRB sensitivity-inducing factor) and NELF (negative elongation factor) (19, 20). Release from this block, which allows Pol-II to proceed along the gene, confers tissue specificity or inducer specificity (21–23). Gene-specific transcription factors can relieve the block by seeding the recruitment or activation of the pause-release factor P-TEFb (24), whose kinase activity phosphorylates DSIF and NELF, dislodging NELF and converting DSIF into an activator (25–27), and phosphorylates Ser2 on the C-terminal domain of Pol-II. The overall result is the unleashing of Pol-II and a number of the other proexpression activities, including H3K4 methylation and assembly of splicing factors (28, 29).

Here, we analyzed the genome-wide impact of Aire on elongation, mapping its chromosomal interaction sites and testing, in primary MECs *ex vivo*, the hypothesis that Aire affects Pol-II complexes nonproductively stalled at the promoter of its target genes.

Results

Pol-II stalled at the transcriptional start sites (TSS) results in the accumulation of short RNAs (30–33). Because MECs are far too rare for direct analysis of transcriptional elongation, we sought to identify the hallmarks of elongation control through microarray

Author contributions: M.G., H.Y., J.A., P.B.R., R.A.Y., D.M., and C.B. designed research; M.G., H.Y., J.A., and P.B.R. performed research; M.G., H.Y., J.A., P.B.R., R.A.Y., D.M., and C.B. analyzed data; and M.G., H.Y., J.A., P.B.R., R.A.Y., D.M., and C.B. wrote the paper.

The authors declare no conflict of interest.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE33878).

¹Present address: Department of Immunology, Cochin Institute, Paris, France.

²Present address: Department of Immunology, Weizmann Institute of Science, Rehovot, Israel.

³To whom correspondence should be addressed. E-mail: cbdm@hms.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1119351109/-DCSupplemental.

analysis, with expression profiles generated from sorted MECs of *Aire*-knockout (KO) and WT mice on the C57BL/6 (B6) × NOD (GEO GSE33878). These experiments used Affymetrix MoGene

ST1.0 microarrays, which include many short probes (“features”) spread across the whole length of the mRNA, with several probes for each exon. We reasoned that analysis of the array data

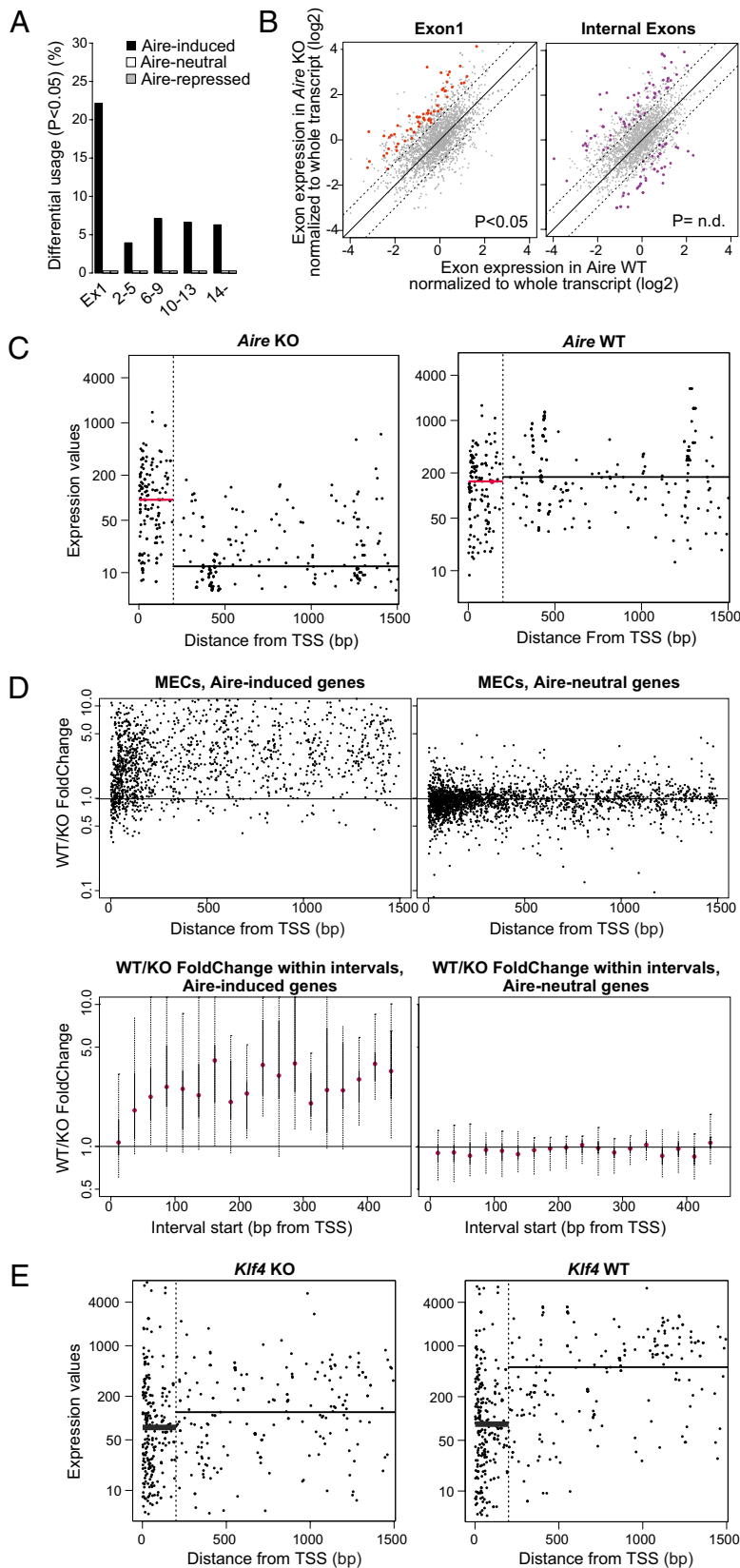


Fig. 1. “Feature-level” analysis of microarray data identifies a role for Aire in releasing paused Pol-II. (A) Expression values for individual microarray features were analyzed relative to whole-transcript averages, and “imbalanced” exons in datasets from MECs from *Aire*-KO and *Aire*-WT mice were flagged (PLATA t test $P < 0.05$). The proportion of imbalanced exons is shown for 420 transcripts most strongly induced by Aire, or an equal number of expression-matched *Aire*-neutral genes, or 84 transcripts repressed by Aire (Table S1). (B) Exon-level expression was normalized to the whole-transcript average in either WT or *Aire* KO datasets. Transcripts with imbalanced exon1 representation are shown in red (Left) or imbalanced representation of internal exons in purple (Right). (C) The expression level of each feature for transcripts showing an exon1 imbalance was plotted relative to its distance from the TSS for *Aire*-induced genes, in *Aire* KO or WT MECs. (D) Expression level of each feature for all *Aire*-induced (Left; overall WT/KO ratio > 3) or *Aire*-neutral genes (Right) plotted as a function of distance from the TSS, either individually (Upper) or grouped in 25-bp bins (Lower). (E) The expression level of each feature for transcripts showing an exon1 imbalance was plotted relative to its distance from the TSS for *Klf4*-induced genes, in *Klf4* KO or WT cells (data from ref. 35). A “TSS proximal–distal” index was calculated in the KO samples as the ratio of the medians of expression values in proximal (< 200 bp from TSS) vs. distal features.

at the feature level might reveal a footprint of Aire's effects on elongation, either at the very beginning of the gene (if Aire were to release stalled polymerases) or spread more uniformly throughout the target genes (if it were to generally facilitate elongation). We developed an R-implementation of the PLATA algorithm, a software tool that detects differential splicing through an imbalance of signals from individual exons relative to the entire gene (34). In effect, the algorithm detects in the KO data the over- or underrepresentation of intensity for single features relative to the whole gene, the relative efficiency of the feature being normalized by the feature/gene ratio observed in the control WT dataset. Analysis of ST1.0 profiles of Aire-induced genes in MECs from *Aire*-KO and WT mice showed a much greater proportion of imbalanced exons for the first than for internal exons (at greater than twofold relative variation from gene-level and t test $P < 0.05$). Such a skew was not observed for Aire-neutral or Aire-repressed genes (Fig. 1A). We then analyzed the directionality of these imbalances by plotting the expression values of each imbalanced exon relative to the whole transcript in WT vs. KO MECs. The directionality of these imbalanced exons was completely skewed, with higher signals in *Aire*-KO relative to WT MECs (Fig. 1B, *Left*), indicating that these sequences were always overrepresented in RNA from KO MECs relative to the rest of the transcript. In contrast, there was no such bias with imbalanced internal exons, which likely reflect differential splicing that Aire could potentially influence in either direction (Fig. 1B, *Right*).

To better understand this imbalanced representation of first exons in Aire-induced genes, we selected transcripts presenting an exon1 imbalance and plotted the expression value at each feature relative to its distance from the TSS. In *Aire*-KO MECs, a sharp drop was observed between features located within 200 bp of the TSSs and those further downstream (Mann-Whitney $P = 2.10^{-16}$) (Fig. 1C, *Left*). This drop was absent for the same genes in *Aire*-WT MECs (Fig. 1C, *Right*). For a broader perspective on Aire-activated genes, we analyzed all genes induced at least twofold by Aire, calculating the WT/KO ratio for each feature. For many of those features mapping in the first 50–100 bp from the TSS, the ratio remained mainly around 1, increasing only further 3' [Fig. 1D, scatter plot of all features (*Upper*), or after averaging WT/KO ratios within a sliding window (*Lower*)]. This drop in expression after 50–100 bp in KO MECs, coming after signals comparable to those of WT, suggests that initiation

of transcription of Aire target genes by Pol-II occurred normally even in the absence of Aire but that elongation was quickly extinguished, a block that could be lifted by Aire.

As a comparator, we analyzed published ST1.0 datasets from cells for *Klf4* (35), a conventional DNA-binding transcription factor whose levels of induction of target genes are comparable to Aire. In *Klf4*-KO cells, there was no drop in the expression signals distal to the TSS (Fig. 1E), actually just the opposite, which we interpret as reflecting the involvement of additional TSS downstream of the canonical TSS (36).

To validate this means to evaluate the impact of transcriptional regulators on Pol-II elongation, we generated ST1.0 datasets from 4D6 cells treated with flavopiridol, an inhibitor of CDK9 (the enzymatically active half of P-TEFb), which prevents polymerase release and inhibits Aire-driven induction of the four Aire-sensitive genes that we tested (Fig. 2A). Transcripts in flavopiridol-treated cells seemed quite similar to Aire-induced transcripts in MECs, with a drop in average expression after 100 bp (Fig. 2B).

Last, we analyzed published ST1.0 datasets from KOs or knockdowns of sequence-specific transcription initiation factors [Foxl2, LMX1B, and Stat3 (37–39)] or factors involved in the stalling or release of Pol-II [CDK8, which encodes a kinase found in some variants of Mediator, a complex known to regulate the recruitment and activation of P-TEFb (40, 41)]. We selected the genes induced by each factor that presented an exon1 imbalance and calculated from the KO datasets a "proximal–distal" index, the difference between the median of the single-probe expression values within 200 bp of the TSSs and those farther away. As Aire, the factors involved in elongation (CDK8, or flavopiridol treatment) showed a positive index, whereas initiation factors had a negative index (Fig. 2C). This difference between the signature of initiation and pause-release factors further substantiates the localization of Aire's effect to the release of stalled Pol-II.

This inference that Aire effects Pol-II release on its target genes rested primarily on analysis of the resulting transcripts; it seemed important to analyze by independent means the global effect of Aire on the positioning of Pol-II on the genome. We performed ChIP coupled to high-throughput sequencing (ChIP-seq), mapping Aire and Pol-II on the genome. To achieve the high cell numbers needed for this technique, we used 293 cells in which we previously showed Aire to be transcriptionally active

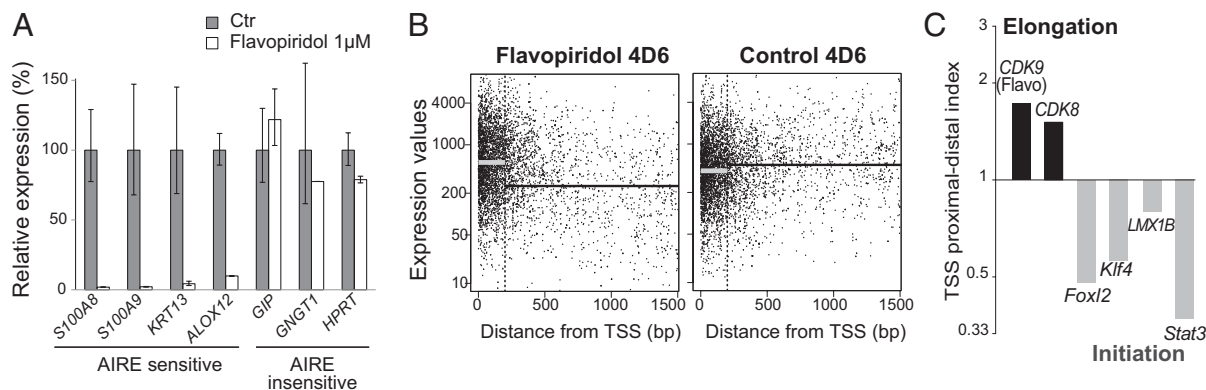


Fig. 2. Validation of the microarray-based detection of stalled polymerases. (A) RT-PCR analysis of the effect of flavopiridol on Aire-transfected 4D6 cells for a panel of Aire-sensitive or -insensitive genes. Relative expression levels are shown as the mean \pm SD from three replicates. (B) 4D6 epithelial cells were treated with flavopiridol for 24 h before expression profiling. The expression level of each microarray feature for flavopiridol sensitive transcripts showing an exon1 imbalance is plotted vs. its distance from the TSS, in flavopiridol-treated and control 4D6 cells. (C) A "TSS proximal–distal" index was calculated in KO samples for several transcription factors from plots such as those of Fig. 1C as the ratio of the medians of expression values in proximal (<200 bp from TSS) vs. distal features. TSS "proximal–distal" indices calculated for sensitive genes, in expression datasets of KO of factors involved in the release of stalled Pol-II (black) or of classic transcription factors impacting initiation (gray).

upon transfection (8). Two independent experiments were performed (Table S2), with highly concordant results. Signals for Aire were weaker than for Pol-II, such that only 1,562–11,400 enriched regions were detected at a statistical significance of 10^{-5} or better (23) vs. 38,900 for Pol-II. As illustrated for two typical genes, Aire was detected primarily, although not exclusively, at the TSS regions, with a local distribution coinciding with the peak of Pol-II (Fig. 3A). On representations at higher resolution, the peak of Aire binding largely overlapped with that of Pol-II (Fig. 3B and C). Integrated heatmaps of the distributions are shown in Fig. 3D, ranked relative to mRNA expression levels in 293 cells. The strongest Aire binding was seen on genes with the highest expression, and correspondingly the highest Pol-II density at the TSS (Fig. 3D). This general correlation of Aire binding

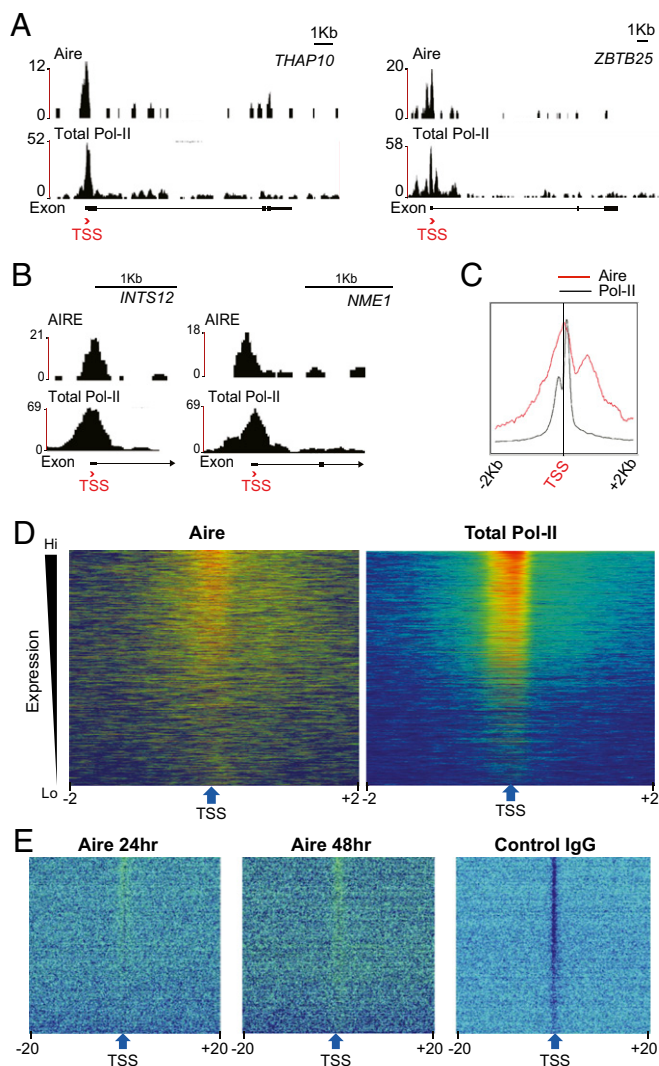


Fig. 3. Aire accumulates at the TSS of Pol-II-rich genes. (A) ChIP-seq analysis was performed in control or Aire-transfected 293 cells. Binding of Aire (Upper) and total Pol-II (Lower) over the whole length of an Aire-induced (THAP10) and an Aire-neutral gene (ZBTB25), shown as examples. (B) As in A, close-up representation of the TSS region of an Aire-induced (INTS12) and an Aire-neutral gene (NME1). (C) Cumulative distribution of Aire and Pol-II in a 4-kb window around the TSS of 2,272 genes. (D) Distribution of total Pol-II and Aire in 25-bp bins in a window of 4 kb around the TSS of all genes, ranked by expression values. (E) As in D, representation of Aire-bound DNA in an independent experiment at two different times after Aire transfection; Right: chromatin precipitated with control IgG (normal mouse IgG1).

with the density of Pol-II was independent of Aire's transcriptional effect, because most of the genes displayed in Fig. 3D are not notably influenced by Aire. This distribution of Aire around the TSS was not an immunoprecipitation artifact, as evidenced by the use of an irrelevant control antibody in an independent ChIP-seq experiment, in which Aire binding was tested at two time points after transfection (Fig. 3E). The complex around the TSS actually leads to an underrepresentation of the region in control precipitates.

These results do not agree with an earlier report from a more limited ChIP-on-chip experiment (42), which indicated a random distribution of Aire. This discrepancy likely stems from the lower sensitivity and span of the ChIP-chip, but perhaps also because stable AIRE transfectants were used in the ChIP-chip study.

The correlation of Aire and Pol-II density was surprising because it is known that it predominantly influences poorly transcribed genes. To further analyze Aire's specificity of action, we plotted the expression level for each gene vs. the total Pol-II in a 500-bp window around its TSS, in control-transfected 293 cells (Fig. 4A). As expected, the most active genes had the most Pol-II bound at the TSS. Genes sensitive to Aire induction (determined from microarray analysis of Aire-transfected 293 cells, fold change >2 ; red dots in Fig. 4A; Table S3) showed a skewed distribution, generally displaced toward higher levels of Pol-II compared with the general distribution of all genes (ANOVA $P = 5.2 \times 10^{-14}$) or of Aire-repressed genes (fold change <0.5 ; blue dots in Fig. 4A). This excess of Pol-II relative to the steady-state transcript levels was also seen in the representation of

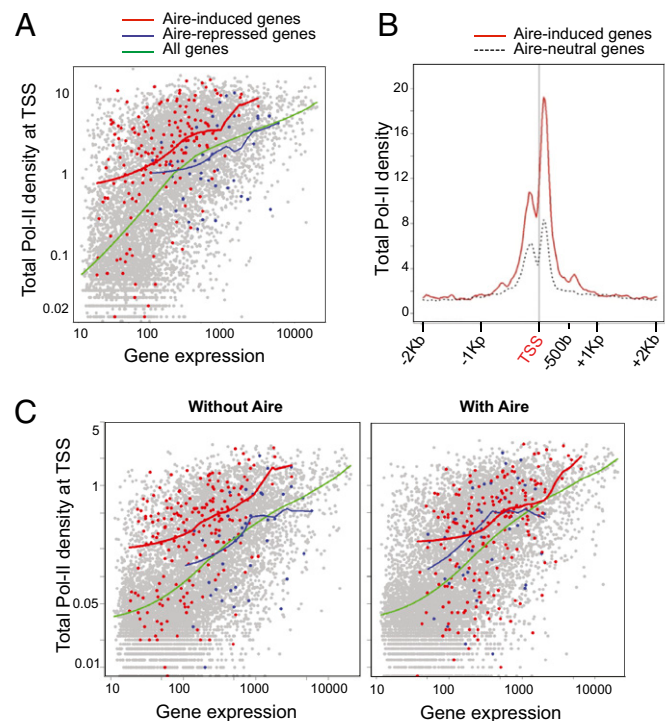


Fig. 4. Aire activates transcripts with a surfeit of stalled polymerases. (A) Mean signal density of Pol-II in windows (–300 to +200 bp) around the TSS plotted against the microarray expression levels of the corresponding gene. Gray, all Aire-neutral genes; red, Aire-induced genes; blue, Aire-repressed genes. Red, blue, and green lines are spline fits for Aire-induced, Aire-repressed, and neutral genes, respectively. (B) Mean density of total Pol-II in 25-bp bins around the TSS of Aire-induced and Aire-neutral genes. (C) Effect of Aire transfection: as in A, mean Pol-II signal density around the TSS plotted against microarray expression levels, in control (Left) or Aire-transfected (Right) 293 cells. Color coding as in A.

Fig. 4B, where the local density of Pol-II around the TSS was compared for Aire-induced genes and a set of Aire-neutral genes matched for expression levels. The disposition of Pol-II was similar, but the levels were higher for genes susceptible to activation by Aire. Thus, although Aire bound the TSS of all genes, those that responded to its presence had a surfeit of TSS-bound Pol-II relative to their activity, consistent with an impact on stalled Pol-II.

Next, we examined the effect of transfected Aire on the distribution of Pol-II on Aire-induced or Aire-neutral genes, by comparing ChIP-seq datasets from control and Aire-transfected 293 cells. With the same gene expression vs. TSS Pol-II representation as above, there was a shift after Aire transfection, with a normalization of the discrepancy between the density of Pol-II and expression levels in Aire-induced and -neutral genes (Fig. 4C). These patterns were consistent with the conclusion that Aire affects the release of stalled Pol-II on its target genes.

Discussion

The function of Aire-dependent ectopic gene expression in MECs is the display of a wide array of proteins to impose self-tolerance on the differentiating thymocyte pool, clearly an unusual challenge to gene regulation. The present results provide two important insights into its mechanism. First, that Aire's activity in MECs is associated with release of stalled polymerases. Second, that there is a disconnect between its physical targeting in the genome: although Aire localizes to all Pol-II-rich TSS, it preferentially affects only a subset of them.

A unique approach to analyze gene expression profiles from MECs of adult mice revealed, in the absence of Aire, a relative overabundance of short transcripts from Aire's target genes, similar to those typically produced by stalled polymerases (30–32). The validity of this mode of analysis was confirmed by testing in a similar manner the transcriptional footprint of several known transcriptional initiation factors and of flavopiridol, a pharmacological inhibitor of P-TEFb and, thereby, transcriptional elongation. This unique application of feature-level microarray analysis to the detection of products of stalled polymerases may have more general use, because it requires considerably fewer cells than ChIP-seq or RNA-seq and can be applied retrospectively to existing microarray datasets. Interestingly, the profile observed in Aire's absence (quasi-normal abundance of RNAs immediately downstream of the TSS, followed by a rapid drop) was different from that observed with Klf4, with which TSS-proximal signals in its target genes varied in lockstep with the rest of the gene.

The ChIP-seq analysis also brought a unique perspective: within the limits of sensitivity of our detection, Aire seemed to bind all TSS regions, and the highest amount of Aire was found on the promoters with the most Pol-II. Pol-II-rich promoters are also those with the highest density of NELF (23, 43). Of all of the promoters to which it bound, Aire activated only a subset, and Aire-induced genes were those that had, in the absence of Aire, a surfeit of polymerases relative to their final expression levels. These are characteristics of genes with stalled polymerases and, indeed, the introduction of Aire relieved the relative surfeit of bound polymerases.

If Aire and its allies partake in pause release, how is this accomplished at the biochemical level? One might propose that one or more of the factors identified in a screen for Aire-interacting proteins (8) may be involved, in particular the elongation-promoting Top2/FACT/DNA-PK complex, which may help release the elongation block. In addition, we have recently identified several cofactors of Aire transcriptional activation in a large-scale

RNAi screen; some of the factors identified in this screen are likely to influence the release of stalled polymerases (e.g., HEXIM1, which sequesters P-TEFb), and it will be important to elucidate these molecular details.

Thus, we propose that Aire's action may operate according to the following scenario. First, upon its induction in terminally differentiated MECs, Aire would generically "land" on Pol-II-rich promoter regions, its functional selectivity not being due to this original recruitment but to the state of Pol-II complexes, such as abundance of NELF or DSIF, and perhaps to local histone marks such as H3K4me(0), which Aire binds electively and functionally (9, 10). Aire would then boost the recruitment of several transcription-promoting factors, relieving inefficiencies at elongation and processing steps. This "sluggish" gene state that Aire targets may result from the absence in thymic epithelial cells of the cognate transcription factors that drive tissue-specific expression in their "normal" cell type. This model explains many previously noted features of Aire's activity: how it can activate so many loci, why the targets vary from cell type to cell type (7), and why PTAs are preferentially targeted. Is this mode of transcriptional control unique to Aire? Myc and Tif1g have also been shown to relieve stalled polymerases to achieve distinctive footprints (22, 23). One can imagine situations in which such a form of control might be advantageous, for instance in transcriptional reprogramming, and it may not be a coincidence that Aire is also present in the testis and ovary. It is also possible that terminally differentiated MECs are by nature particularly open to such effects, and it is important to recall that some degree of PTA expression occurs in MECs, even in Aire's absence. From a teleological standpoint, generically unleashing sluggish genes with blocked polymerases would seem an effective strategy to achieve widespread induction of silent genes and thus drive the induction of organismal tolerance.

Materials and Methods

Gene Expression Profiling. Total RNA, isolated from sorted MECs of 4-wk-old individual mice, was profiled as previously described (44) on Affymetrix Mouse Gene ST1.0 arrays. For feature-level analysis, normalized feature-level expression files, generated by aroma.affymetrix, were processed with an R-implementation of the PLATA algorithm (34), and the expression of a given exon was flagged as imbalanced if the ratio between exon WT/KO fold change to transcript WT/KO fold change was >2 or <0.5 , with a *t* test *P* value <0.05 for a WT vs. KO comparison of expression values of all features mapping to the exon, after normalization to whole-transcript expression. An index was then computed as the ratio of the medians of expression values in promoter-proximal (TSS to 200 bp) vs. distal (>200 bp) features of the microarray.

Chromatin Immunoprecipitation. 293FT cells were transfected for pCMV-Tag2C (control) or pCMV-Tag2C containing FLAG-tagged murine Aire (Aire) with TransIT-293 Transfection Reagent (Mirus). The cells were cross-linked by formaldehyde 48 h after the transfection for ChIP, as previously described (23). Antibodies used are as follows: total Pol-II: Santa Cruz sc-899; FLAG-tagged Aire: Sigma (M2 antibody) F1804; control IgG: Santa Cruz sc-3877. For ChIP-seq analysis, Illumina sequencing and analysis were as previously described (45), and sequence counts integrated in 25-bp bins, with plots generated in R. A detailed description of data analysis is provided in *SI Materials and Methods* and *Dataset S1*.

ACKNOWLEDGMENTS. We thank S. Davis, G. Frampton, and H. Paik for help with computational analysis. This work was supported by National Institutes of Health Grants R01-DK060027 and R01-AI088204 (to D.M. and C.B.) and R01-HG002668 and R01CA46455 (to R.A.Y.). M.G. received fellowships from the Campbell and Hall Charity Fund, The Harold Whitworth Pierce Charitable Trust, and the Philippe Foundation. H.Y. was supported by a fellowship from the Human Frontier Science Program, J.A. by Juvenile Diabetes Research Foundation Fellowships 10-2007-601 and 03-2005-1138, and P.B.R. by American Cancer Society Postdoctoral Fellowship 120272-PF-11-042-01-DMC.

1. Kyewski B, Klein L (2006) A central role for central tolerance. *Annu Rev Immunol* 24: 571–606.

2. Anderson MS, et al. (2002) Projection of an immunological self shadow within the thymus by the aire protein. *Science* 298:1395–1401.

3. Mathis D, Benoist C (2009) Aire. *Annu Rev Immunol* 27:287–312.
4. Venanzi ES, Melamed R, Mathis D, Benoist C (2008) The variable immunological self: Genetic variation and nongenetic noise in Aire-regulated transcription. *Proc Natl Acad Sci USA* 105:15860–15865.
5. Villaseñor J, Besse W, Benoist C, Mathis D (2008) Ectopic expression of peripheral-tissue antigens in the thymic epithelium: Probabilistic, monoallelic, misinitiated. *Proc Natl Acad Sci USA* 105:15854–15859.
6. Derbinski J, Pinto S, Rösch S, Hexel K, Kyewski B (2008) Promiscuous gene expression patterns in single medullary thymic epithelial cells argue for a stochastic mechanism. *Proc Natl Acad Sci USA* 105:657–662.
7. Guerau-de-Arellano M, Mathis D, Benoist C (2008) Transcriptional impact of Aire varies with cell type. *Proc Natl Acad Sci USA* 105:14011–14016.
8. Abramson J, Giraud M, Benoist C, Mathis D (2010) Aire's partners in the molecular control of immunological tolerance. *Cell* 140:123–135.
9. Org T, et al. (2008) The autoimmune regulator PHD finger binds to non-methylated histone H3K4 to activate gene expression. *EMBO Rep* 9:370–376.
10. Koh AS, et al. (2008) Aire employs a histone-binding module to mediate immunological tolerance, linking chromatin regulation with organ-specific autoimmunity. *Proc Natl Acad Sci USA* 105:15878–15883.
11. Chignola F, et al. (2009) The solution structure of the first PHD finger of autoimmune regulator in complex with non-modified histone H3 tail reveals the antagonistic role of H3R2 methylation. *Nucleic Acids Res* 37:2951–2961.
12. Koh AS, Kingston RE, Benoist C, Mathis D (2010) Global relevance of Aire binding to hypomethylated lysine-4 of histone-3. *Proc Natl Acad Sci USA* 107:13016–13021.
13. Owen I, et al. (2007) AIRE recruits P-TEFb for transcriptional elongation of target genes in medullary thymic epithelial cells. *Mol Cell Biol* 27:8815–8823.
14. Rougvie AE, Lis JT (1988) The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of *D. melanogaster* is transcriptionally engaged. *Cell* 54:795–804.
15. Bentley DL, Groudine M (1986) A block to elongation is largely responsible for decreased transcription of c-myc in differentiated HL60 cells. *Nature* 321:702–706.
16. Muse GW, et al. (2007) RNA polymerase is poised for activation across the genome. *Nat Genet* 39:1507–1511.
17. Zeitlinger J, et al. (2007) RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nat Genet* 39:1512–1516.
18. Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA (2007) A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* 130:77–88.
19. Wada T, et al. (1998) DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev* 12:343–356.
20. Yamaguchi Y, et al. (1999) NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell* 97:41–51.
21. Hargreaves DC, Horng T, Medzhitov R (2009) Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell* 138:129–145.
22. Bai X, et al. (2010) TIF1gamma controls erythroid cell fate by regulating transcription elongation. *Cell* 142:133–143.
23. Rahl PB, et al. (2010) c-Myc regulates transcriptional pause release. *Cell* 141:432–445.
24. Marshall NF, Price DH (1995) Purification of P-TEFb, a transcription factor required for the transition into productive elongation. *J Biol Chem* 270:12335–12338.
25. Ivanov D, Kwak YT, Guo J, Gaynor RB (2000) Domains in the SPT5 protein that modulate its transcriptional regulatory properties. *Mol Cell Biol* 20:2970–2983.
26. Fujinaga K, et al. (2004) Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. *Mol Cell Biol* 24:787–795.
27. Yamada T, et al. (2006) P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. *Mol Cell* 21:227–237.
28. Saunders A, Core LJ, Lis JT (2006) Breaking barriers to transcription elongation. *Nat Rev Mol Cell Biol* 7:557–567.
29. Lin S, Coutinho-Mansfield G, Wang D, Pandit S, Fu XD (2008) The splicing factor SC35 has an active role in transcriptional elongation. *Nat Struct Mol Biol* 15:819–826.
30. Kapranov P, et al. (2007) RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* 316:1484–1488.
31. Core LJ, Waterfall JJ, Lis JT (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 322:1845–1848.
32. Nechaev S, et al. (2010) Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. *Science* 327:335–338.
33. Seila AC, Core LJ, Lis JT, Sharp PA (2009) Divergent transcription: A new feature of active promoters. *Cell Cycle* 8:2557–2564.
34. Sandberg R, Neilson JR, Sarma A, Sharp PA, Burge CB (2008) Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* 320:1643–1647.
35. Tetreault MP, et al. (2010) Esophageal squamous cell dysplasia and delayed differentiation with deletion of krüppel-like factor 4 in murine esophagus. *Gastroenterology* 139:171–181, e9.
36. Carninci P, et al. (2006) Genome-wide analysis of mammalian promoter architecture and evolution. *Nat Genet* 38:626–635.
37. Uhlenhaut NH, et al. (2009) Somatic sex reprogramming of adult ovaries to testes by FOXL2 ablation. *Cell* 139:1130–1142.
38. Rasclé A, et al. (2009) The LIM-homeodomain transcription factor LMX1B regulates expression of NF-kappa B target genes. *Exp Cell Res* 315:76–96.
39. Musteanu M, et al. (2010) Stat3 is a negative regulator of intestinal tumor progression in Apc(Min) mice. *Gastroenterology* 138:1003–1011, e1–e5.
40. Donner AJ, Ebmeier CC, Taatjes DJ, Espinosa JM (2010) CDK8 is a positive regulator of transcriptional elongation within the serum response network. *Nat Struct Mol Biol* 17:194–201.
41. Takahashi H, et al. (2011) Human mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell* 146:92–104.
42. Org T, et al. (2009) AIRE activated tissue specific genes have histone modifications associated with inactive chromatin. *Hum Mol Genet* 18:4699–4710.
43. Gilchrist DA, et al. (2010) Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell* 143:540–551.
44. Zipris D, Lazarus AH, Crow AR, Hadzija M, Delovitch TL (1991) Defective thymic T cell activation by concanavalin A and anti-CD3 in autoimmune nonobese diabetic mice. Evidence for thymic T cell anergy that correlates with the onset of insulinitis. *J Immunol* 146:3763–3771.
45. Marson A, et al. (2008) Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* 134:521–533.