Transcription-Dependent Generation of a Specialized Chromatin Structure at the TCRβ Locus

Joaquin Zacarías-Cabeza,[*,†,‡,1,2] Mohamed Belhocine,[§,¶,2] Laurent Vanhille,[§,¶] Pierre Cauchy,[*,†,‡,§,¶] Frederic Koch,[*,†,‡] Aleksandra Pekowska,[*,†,‡] Romain Fenouil,[*,†,‡] Aurélie Bergon,[§,¶,‖] Marta Gut,[#] Ivo Gut,[#] Dirk Eick,[**] Jean Imbert,[§,¶,‖] Pierre Ferrier,[*,†,‡] Jean-Christophe Andrau,[*,†,‡,††] and Salvatore Spicuglia[*,†,‡,§,¶]

V(D)J recombination assembles Ag receptor genes during lymphocyte development. Enhancers at AR loci are known to control V(D)J recombination at associated alleles, in part by increasing chromatin accessibility of the locus, to allow the recombination machinery to gain access to its chromosomal substrates. However, whether there is a specific mechanism to induce chromatin accessibility at AR loci is still unclear. In this article, we highlight a specialized epigenetic marking characterized by high and extended H3K4me3 levels throughout the Dβ-Jβ-Cβ gene segments. We show that extended H3K4 trimethylation at the *Tcrb* locus depends on RNA polymerase II (Pol II)–mediated transcription. Furthermore, we found that the genomic regions encompassing the two DJCβ clusters are highly enriched for Ser[5]-phosphorylated Pol II and short-RNA transcripts, two hallmarks of transcription initiation and early transcription. Of interest, these features are shared with few other tissue-specific genes. We propose that the entire DJCβ regions behave as transcription "initiation" platforms, therefore linking a specialized mechanism of Pol II transcription with extended H3K4 trimethylation and highly accessible Dβ and Jβ gene segments. ***The Journal of Immunology*****, 2015, 194: 3432–3443.**

V(D)J recombination assembles Ag receptor genes from germline V, D, and J segments during lymphocyte development (1). In αβT cells, this leads to the subsequent expression of TCR β- and α-chains. For V(D)J recombination to occur, the presence of the lymphoid-specific proteins RAG1 and RAG2 and the ubiquitously expressed DNA repair factors from the nonhomologous end joining pathway are required (2). Control of V(D)J recombination is required to ensure cell lineage specificity, dictate the temporal order of rearrangements, and allow allelic exclusion at certain AR genes (3). This regulation mainly relies on the modulation of chromatin accessibility at the AR-associated recombination sequences (RSs)to the recombination machinery.

The accessibility model was initially based on the observation that transcription of AR germline gene segments correlated developmentally with their recombination during lymphoid cell differentiation (4). Subsequently, this model has been strengthened by findings that link V(D)J recombination to transcriptional control elements, such as AR-associated enhancers and promoters, and to several molecular parameters related to open chromatin (including association with active histone marks, DNA hypomethylation, and nuclease hypersensitivity) (3, 5). Robust germline transcription at

*Centre d'Immunologie de Marseille-Luminy, Aix-Marseille University, UM2, 13288 Marseille, France; †INSERM, U1104, 13288 Marseille, France; ‡Centre National de la Recherche Scientifique, UMR7280, F-13009 Marseille, France; §INSERM U1090, Technological Advances for Genomics and Clinics, F-13009 Marseille, France; ¶Aix-Marseille University, UMR-S 1090, Technological Advances for Genomics and Clinics, F-13009 Marseille, France; ‖Transcriptomic and Genomic Marseille-Luminy, Infrastructures en Biologie, Santé et Agronomie, 13288 Marseille, France; #Centre Nacional D'Anàlisi Genòmica, Parc Científic de Barcelona, Baldiri i Reixac, 08028 Barcelona, Spain; **Department of Molecular Epigenetics, Helmholtz Center Munich, Center for Integrated Protein Science, 80336 Munich, Germany; and ††Institut de Génétique Moléculaire de Montpellier, Centre National de la Recherche Scientifique, UMR5535, 34293 Montpellier, France

[1]Current address: Institute of Biomedical Research, University of Birmingham, Birmingham, UK.

[2]J.Z.-C. and M.B. contributed equally to this work.

Received for publication March 27, 2014. Accepted for publication February 2, 2015.

This work was supported by institutional grants from INSERM, the Centre National de la Recherche Scientifique, Aix-Marseille University, and by specific grants from the Fondation Princesse Grace de Monaco (to P.F.), the Fondation de France (to P.F.), the Association pour la Recherche sur le Cancer (to S.S., Project SFI20111203756), the Fondation pour la Recherche Médicale (to P.F.), the Agence Nationale de la Recherche (to P.F.), the Institut National du Cancer (to P.F.), the European Union's Seventh Framework Program (FP7) (to S.S., Agreement 282510-BLUEPRINT), and Initiative D'Excellence Aix-Marseille Project ANR-11-IDEX-0001-02 (to S.S.) funded by the Investissements d'Avenir French Government program. Sequencing costs for this work were supported by a European Study Group with Industry Consortium grant of the European Union (to J.-C.A., program T-DynRegSeq) from the FP7 (FP7/2007-2013) under Grant Agreement 262055. The Transcriptomic and Genomic Marseille-Luminy sequencing platform is supported by grants from Infrastructures en Biologie, Santé et Agronomie and the France Génomique National infrastructure, funded as part of the Investissements d'Avenir program managed by the Agence Nationale pour la Recherche (Contract ANR-10-INBS-09). J.Z.-C. was supported by Grant R07116AS from the Agence Nationale de la Recherche MIMe program (to P.F.). Work in J.-C.A.'s laboratory is also supported by a grant of the Fondation pour la Recherche Médicale (AJE20130728183).

The sequences presented in this article have been submitted to the National Center for Biotechnology Information's Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo) under accession numbers GSE63416, GSE64709, GSM1360722–GSM1360727, and GSM1359828.

Address correspondence and reprint requests to Dr. Pierre Ferrier, Dr. Jean-Christophe Andrau, or Dr. Salvatore Spicuglia, Centre d'Immunologie de Marseille-Luminy, Aix-Marseille University, UM2, 13288 Marseille, France (P.F.), Institut de Génétique Moléculaire de Montpellier, Centre National de la Recherche Scientifique, UMR5535, 34293 Montpellier, France (J.-C.A.), or Technological Advances for Genomics and Clinics, INSERM UMR-S 1090, Parc Scientifique de Luminy, Cedex 9, F-13009 Marseille, France (S.S.). E-mail addresses: ferrier@ciml.univ-mrs.fr (P.F.), jean-christophe.andrau@igmm.cnrs.fr (J.-C.A.), or salvatore.spicuglia@inserm.fr (S.S.)

The online version of this article contains supplemental material.

Abbreviations used in this article: ChIP, chromatin immunoprecipitation; CTD, C-terminal domain; DN, double negative; DP, double positive; FAIRE, formaldehyde-assisted isolation of regulatory elements; GEO, Gene Expression Omnibus; Pol II, polymerase II; qPCR, quantitative PCR; RS, recombination sequence; seq, sequencing; Ser[5P], phosphorylated serine 5; TIP, transcription initiation platform; TSS, transcriptional start site.

(D)J clusters is an initial activation event at all AR loci that generates a focal zone of RAG1/2 binding, termed the recombination center (6, 7). More insight into the accessibility model was provided by recent studies demonstrating that the PHD finger domain of RAG2 binds with high affinity to histone H3 trimethylated at K4 (H3K4me3) and that RAG2 is recruited to H3K4me3 domains genome-wide (6–8).

A central prediction of the accessibility model is, therefore, that transcriptional control elements and/or transcription itself are critical for allowing the recombination machinery to gain access to RSs (9). However, in most mammals' genes, highly open chromatin structure is mainly confined to the *cis*-regulatory sequences themselves (10). In particular, H3K4me3 is highly enriched at promoter regions of expressed genes but is not generally found in the body of the genes (11). Thus, the question still remains as to how chromatin accessibility is established at the recombining gene segments and associated RSs, which are often located distant from the *cis*-regulatory elements. We and others have recently shown that a subset of tissue-specific genes might display broad epigenetic marking, including extended H3K4me2 and H3K4me3, along with elevated loading of polymerase II (Pol II) (12–14). This raises the possibility that transcriptional activity throughout V(D)J rearranging loci might play a more elaborate role in the remodeling of chromatin structure and targeting of the recombinase machinery.

Genetic studies at the *Tcrb* locus have shed light on the complex cooperation between enhancer- and promoter-bound transcription factors to control V(D)J recombination during T lymphocyte development (15, 16). The mouse *Tcrb* locus spreads over ~670 kb, including a ~390-kb 5′ domain containing 21 Vβ gene segments and a 26-kb 3′ domain comprising a duplicated cluster of Dβ-Jβ-Cβ gene segments, followed by a single Vβ gene segment, Vβ31. *Tcrb* gene recombination is restricted to the T cell lineage and is activated along with locus expression. In $CD4^{-}CD8^{-}$ double negative (DN) thymocytes, V(D)J recombination proceeds in a stepwise manner (Dβ-to-Jβ joining occurring first, before Vβ-to-DJβ assembly), triggering, if productive, allelic exclusion at the *Tcrb* locus and further development into the $CD4^{+}CD8^{+}$ double positive (DP) cell stage in the αβT cell lineage, an intricate process also known as β-selection (17).

The 560-bp *Tcrb* gene enhancer (Eβ) lies at the center of the ~10-kb Cβ2-Vβ31 intervening region (18, 19). Knockout mouse models have revealed a critical function of Eβ in the efficient onset of *cis* recombination, with homozygous Eβ-deleted ($E\beta^{-/-}$) mice displaying impaired TCRβ-chain production and αβT cell development (20, 21). Further analysis implied that this element, working together with Dβ-associated promoters of germline transcripts, directs transcription, along with histone marking and chromatin opening, throughout the adjacent DJCβ clusters (22–26). Although Eβ-dependent activity is clearly required to initiate V(D)J recombination at the *Tcrb* locus, the precise mechanism or mechanisms inducing long-range histone marking and chromatin remodeling along the DJCβ regions are still poorly understood. A key feature is the Eβ-dependent transcription activity across the Dβ-Jβ recombination center, which is thought to mediate H3K4 trimethylation at this site, followed by RAG1/2 deposition (3, 16, 22, 23, 26–28).

In the current study, we used chromatin immunoprecipitation (ChIP)–sequencing (seq) technology to comprehensively map H3K4 methylation in germline *Tcrb* alleles from $Rag2^{-/-}$ thymocytes. We found that the DJCβ transcription units were highly enriched for H3K4me3 and linked to local accessibility of the Dβ and Jβ gene segments, highlighting a distinctive epigenetic marking at the *Tcrb* locus. This property was dependent on an unusual regulation of Pol II–mediated transcription in which features of transcription initiation and early elongation, including high levels of phosphorylated serine 5 (Ser^{5P}) Pol II and short-RNA transcripts, were found throughout the entire DJCβ regions. Of interest, these features are shared with a small subset of tissue-specific genes, including other *Tcr* loci. Overall, our study revealed a specialized role for Pol II transcription in the establishment of a highly accessible chromatin domain at the *Tcrb* locus.

Materials and Methods

Mice

Homozygous *Rag2*-deficient (ΔRag) (29) and Eβ-deleted (ΔEβ) (21) mice were housed under specific pathogen–free conditions and handled in accordance with European directives. Mice were bred on a C57BL/6J background and sacrificed for analysis between 4 and 6 wk of age.

KM05283 treatment

A total of 15×10^6 exponentially growing P5424 cells (30) was incubated with either 50 μM KM05283 (Maybridge, Cornwall, U.K.) or control DMSO (Sigma-Aldrich, St. Louis, MO) in RPMI 1640 medium for 16–18 h at 37°C. After incubation, cells were washed two times with 1× Dulbecco's PBS and processed for ChIPs as indicated below. Inhibition of Pol II Ser^{2} phosphorylation was confirmed by Western blot, as described previously (31).

ChIP

ChIP experiments were performed as described previously (14). For histone modification marks, we used 2×10^6 cells along with 3 μg of the following Abs: anti-H3K4me1 (ab8895; Abcam, Cambridge, U.K.), anti-H3K4me2 (ab32356; Abcam), anti-H3K4me3 (ab8580; Abcam), and anti-H3K36me3 (ab9050; Abcam). For Pol II ChIPs, the following Abs and cell numbers were used: anti–total-Pol II (Santa Cruz Biotechnology, Dallas, TX; sc-899×, 10 μg and 10×10^6 cells), anti-Ser^{2P} Pol II [rat monoclonal, clone 3E10 (32), 10 μg and 60×10^6 cells], and anti-Ser^{5P} Pol II [rat monoclonal, clone 3E8 (32); 10 μg and 30×10^6 cells]. The DNA fragments were purified and recovered using the QIAquick PCR Purification Kit (QIAGEN, Hilden, Germany). The quality of individual ChIP samples was checked at known target sites by quantitative PCR (qPCR), and DNA size was verified on a 2100 Bioanalyzer (Agilent, Santa Clara, CA). Primer sets used for qPCR are available upon request.

Formaldehyde-assisted isolation of regulatory elements

Formaldehyde-assisted isolation of regulatory elements (FAIRE) was performed as previously described (33), with slight modifications. Briefly, 20×10^6 thymocytes from ΔRag or ΔRag;ΔEβ mice were cross-linked with 1% formaldehyde for 10 min at room temperature and sonicated 14 times on an S-4000 Sonifier (Misonix, Farmingdale, NY) with 30-s pulses to give DNA fragments of length between 200 and 500 bp. The soluble chromatin of 2×10^6 thymocytes was isolated and subjected to three consecutive phenol-chloroform extractions. Samples were then incubated overnight at 65°C to reverse cross-linking. DNA was finally purified using the MinElute PCR Purification Kit (QIAGEN). DNA concentration was measured using a Nanodrop 1000 (Thermo Scientific, Illkirch, France).

ChIP-seq data generation

Sequencing of ChIP samples was performed according to the Illumina Genome Analyzer ChIP-seq protocol and aligned against the mouse mm9 genome using integrated Eland software. As prefiltering steps, only uniquely mapped tags were used for further processing, and all duplicate tags (those with identical coordinates) were filtered out to remove possible sequencing and/or alignment artifacts. Remaining tags were processed using a custom R pipeline, employing the ShortRead library3 (14). Read-count intensity profiles (wiggle files) were constructed by elongating each mapped read to the estimated fragment size, and counting the elongated read overlaps within a window of 50 nucleotides after normalization of the profile by the number of mapped reads. ChIP-seq data from total-Pol II and from micrococcal nuclease–treated H3K4me1 and H3K4me3 from ΔRag thymocytes were published previously (Ref. 34; GSE55635). Mapped reads, estimated fragment size, and Gene Expression Omnibus (GEO) accession numbers are listed in Supplemental Table I.

RNA extraction and RNA-seq experiments

Total RNA from 10×10^6 thymocytes of ΔRag mice was extracted as previously described (14). Strand-specific preparation, sequencing, and processing of short-RNA samples were carried out as explained earlier (14). RNA quantity and quality were verified using RNA Pico chips on a 2100 Bioanalyzer (Agilent). Mapped reads and GEO accession numbers are listed in Supplemental Table I. Total and polyA RNA-seq data from ΔRag thymocytes were published previously (Ref. 35; GSE44578).

ChIP-seq and RNA-seq data analyses

We first selected non-overlapping genes, harboring a single transcript annotated in the RefSeq database and longer than 8 kb (Supplemental Table II). From this set, the 300 highest expressed genes (Top-300) were selected, based on gene expression data in ΔRag thymocytes (34). To quantify the enrichment levels in H3K4me3, Ser^5 Pol II, and short-RNAs within the gene body, the ChIP-seq signal from wiggle files was quantified within the region from the transcriptional start site (TSS) to +8 kb. In the case of the DJCβ1, DJCβ2, Dδ2Jδ1, Jγ1Cγ1, and Jγ4Cγ4 clusters, the region from the Dβ1, Dβ2, Dδ2, Jγ1, and Jγ2 gene segments to +8 kb, respectively, was used to quantify H3K4me3, Ser^5 Pol II, and short-RNA levels. To directly compare expression levels between the selected RefSeq genes and the different *Tcr* gene clusters, we used polyA RNA-seq data. PolyA RNA level was estimated by counting the average number of tags at the exons of RefSeq genes and at the different *Tcr* gene clusters. All quantifications are shown in Supplemental Table II. The pausing index (also called traveling ratio) was calculated as previously described (36), using the selection of the Top-300 genes.

Analysis of ChIP-on-chip and FAIRE experiments

Enriched DNA fragments from ChIP or FAIRE experiments were hybridized together with input DNA to a previously described 15K array (Agilent) containing the whole *Tcrb* locus at 100-bp resolution (34), following the manufacturer's instructions. The results obtained with two biological replicates were averaged and converted into SGR files using CoCAS software (37). Data from ΔRag and ΔRag;ΔEβ thymocytes were normalized using the overall signal on the entire microarray (excluding the probes within the *Tcrb* locus). Normalized data were displayed in the form of log2 ratio using IGB software (http://bioviz.org/igb/).

Inflection point

We first calculated the average signal of H3K4me3 in the gene body (TSS to +8kb) for each gene. The broad H3K4me3 genes were then determined by identifying an inflection point of the average signal versus gene rank. The inflection point was computed by determining the diagonal line of the curve from endpoints, and by sliding this diagonal line to find where it is tangential. We identified 58 broad H3K4me3 genes (Supplemental Table II).

Transcription initiation platform selection

We selected promoter-associated transcription initiation platforms (TIPs), defined previously in DP thymocytes (14) and expressed in the P5424 cell line (671 TIPs). The TIPs were separated according to their size into three categories: <2 kb (557), between 2 and 2.5 kb (47), and >2.5 kb (67).

Average and boxplot profiles

Average profiles were generated by extracting the ChIP-seq signal from wiggle files around the TSS (from −2kb to +8kb), using a custom R script. Rescaled average profiles were performed by dividing the region from the TSS to the transcriptional termination site into 200 bins. To test whether the differences between different gene sets were statistically significant, we first extracted the average signal of the region of interest and plotted them in boxplot representation and performed a Student *t* test.

Gene expression analyses

Gene expression data of αβ T cells were downloaded from the Immunological Genome Project Web site (www.immgen.org) (38). A quantile normalization was then applied on gene expression of stages ETP (early thymic progenitor), DN1 (DN stage 1), DN2, DN3, DN4, ISP (immature single positive), and DPbl (DP blast). The raw expression data for 74 mouse tissues were downloaded from the National Center for Biotechnology Information GEO (accession number: GSE10246). The raw expression data were normalized by the variance stabilization and normalization method (39), and probe annotation to the NCBI37/mm9 was used for subsequent analyses. To compare the level of expression of genes between T cells and other tissues, we calculated the mean level of expression of genes in T cell samples (including T cell $CD4^+$, T cell $CD8^+$, T cell $Foxp3^+$, thymocytes DP $CD4^+/CD8^+$, thymocyte SP $CD4^+$, and thymocyte SP $CD8^+$) and in the remaining 69 samples. Statistical significance was calculated using a paired Student *t* test.

Gene ontology terms enrichment

Enrichments in Gene Ontology Terms for Biological Process were calculated using the DAVID tool (40), with default settings (count threshold: 2; EASE threshold: 0.1; multiple testing correction by the Benjamini procedure) and *Mus musculus* as background model. We selected the top 10 terms retrieved for each gene set with the lowest *p* values.

CpG density

The genomic sequences ±500 bp around the TSS of each set of genes were recovered. The total number of CpG was counted for each sequence. Statistical significance between the Top-300 and the other set of genes was calculated using a Student *t* test.

Results

A highly open chromatin structure at the DJCβ region

To assess epigenetic features associated with chromatin remodeling of the *Tcrb* locus, we analyzed the three levels of histone H3K4 methylation by ChIP of thymocytes purified from *Rag2*-deficient mice (hereafter ΔRag), followed by high-throughput sequencing (ChIP-seq). The use of the ΔRag mouse model ensures the germline configuration of *Tcrb* alleles while providing an enriched and homogeneous source of T cell precursors. We concentrated our analyses on the Eβ-proximal region, including the two DJCβ clusters (Fig. 1A). We observed that H3K4 methylation marks were not exclusively localized to the known regulatory regions (i.e., the pDβ promoters and Eβ) but, instead, extended throughout the Jβ and Cβ regions. For instance, H3K4me1 and H3K4me2 covered the entire Eβ-proximal region spanning 30 kb from ∼3 kb upstream of Dβ1 to ∼3 kb downstream of Vβ31, thus defining a domain of open chromatin that roughly corresponds to the previously described Eβ-regulated domain (22–26, 41). Intriguingly, however, H3K4me3, which has been shown to be highly enriched at promoter regions (11), but is not generally found in the body of the genes, was broadly distributed throughout the two DJCβ germline transcription units. To exclude any potential bias owing to cross-linked chromatin, we confirmed the extended profile observed for H3K4me1 and H3K4me3 at the DJCβ regions by analyzing ChIP-seq data performed with mononucleosome preparations of native chromatin from ΔRag thymocytes (34) (Fig. 1B). Moreover, Eβ-deleted alleles displayed an almost complete loss of H3K4me3 at the DJCβ regions, suggesting that this epigenetic marking depends on Eβ-mediated transcriptional activation of the locus (Fig. 1C).

We then asked whether the extended H3K4me3 profile observed at the DJCβ regions in ΔRag thymocytes could be reminiscent of a highly open chromatin. To directly determine the accessibility of the chromatin, we performed a FAIRE assay, which allowed the recovery of the soluble (i.e., nucleosome-free) fraction of the chromatin (33). As expected, FAIRE signals were highly enriched at the Eβ region in ΔRag thymocytes (Fig. 1C). In addition, we observed that regions overlapping the Dβ and Jβ gene segments also display high levels of FAIRE signal in ΔRag. We confirmed that the highly open chromatin revealed by FAIRE at the Dβ and Jβ gene segments was largely dependent on Eβ-mediated chromatin remodeling (Fig. 1C). These results confirm and extend previous observations describing extensive Eβ-dependent remodeling of the DJCβ clusters (3, 16, 22, 23, 25–27). Note, however, that residual levels of chromatin accessibility are still observed around the Dβ

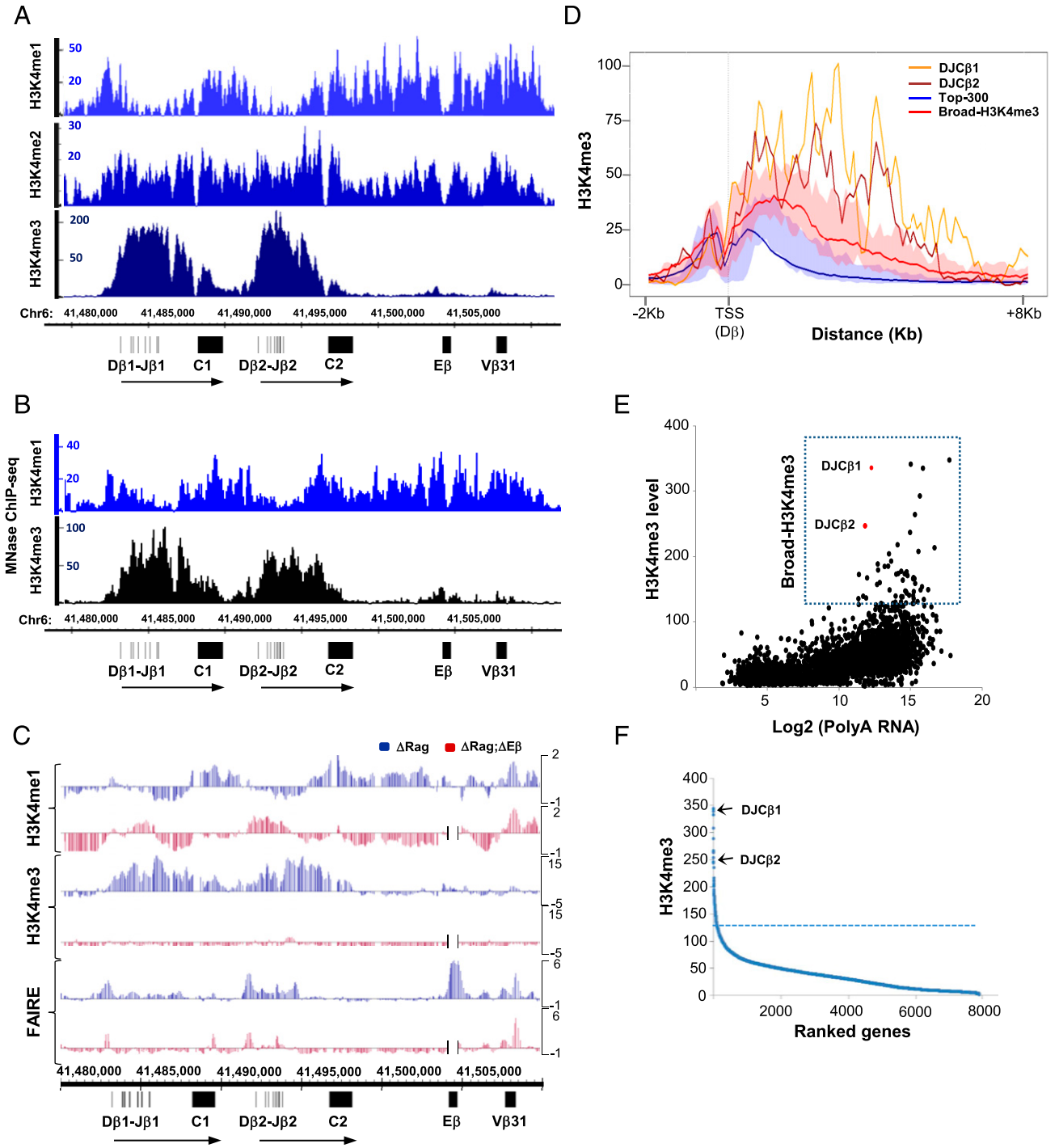

FIGURE 1. Extended H3K4me3 profiles at the DJCβ clusters. (**A**) ChIP-seq profiles of H3K4me1, H3K4me2, and H3K4me3 from cross-linked chromatin from ΔRag thymocytes at the 3′ side of the *Tcrb* locus. The genomic coordinates (mm9) and the location of the *Tcrb* gene segments and the Eβ enhancer (vertical dashes and rectangles) are indicated at the *bottom* of the panel. Horizontal arrows indicate the orientation of the two DJCβ transcription units. (**B**) ChIP-seq profiles of H3K4me1 and H3K4me3 from native chromatin from ΔRag thymocytes at the 3′ side of the *Tcrb* locus. Symbols are as in (A). (**C**) ChIP-on-chip and FAIRE profiles from ΔRag and ΔRag;ΔEβ thymocytes at the 3′ side of the *Tcrb* locus. ChIP-on-chip and FAIRE experiments were performed in duplicate and the merged ChIP(FAIRE)-over-input signals were normalized between the ΔRag and ΔRag;Eβ samples (see *Materials and Methods* for details). The log2 scale is shown. Deletion of the Eβ region in ΔRag;Eβ thymocytes is represented as a gap. (**D**) Average profiling of H3K4me3 ChIP-seq data in ΔRag thymocytes at genomic regions from −2 to +8 kb around the TSS of the set of highly expressed genes (Top-300) and the set of Broad-H3K4me3 associated genes, compared with the profiles of the two DJCβ clusters (the coordinates of the Dβ gene segments were set to 0). The shadow represents the border of the 25th and 75th percentiles. (**E**) Scatter plot showing the H3K4me3 density in the gene body of Refseq genes and in the two DJCβ clusters related to the polyA RNA-seq level. Each dot represents one gene. The two DJCβ clusters are highlighted in red. (**F**) Graph showing the ranked Refseq genes and the two DJCβ clusters related to H3K4me3 density. The two DJCβ clusters are highlighted by arrows. The threshold defining the set of Broad-H3K4me3 genes is indicated by a dotted line.

gene segments in the absence of Eβ, in agreement with an Eβ-independent role of Dβ-associated promoters (23, 25). Overall, in ΔRag thymocytes, highly accessible chromatin domains at the 3′ proximal region of the *Tcrb* locus are not restricted to the enhancer and promoter elements, but are spread over the Dβ and Jβ gene segments, thus providing a unique chomatin signature.

The extended H3K4me3 profile is a specific feature of the Tcrb *locus*

To determine whether this extended profile was a general feature of highly expressed genes, we compared the H3K4me3 profiles at the two DJCβ clusters with the average H3K4me3 profiles of a set of highly expressed genes (Top-300; Fig. 1D). As predicted, expressed genes displayed an H3K4me3 enrichment around the TSS (peaks at −0.5 and +1 kb from the TSS). In comparison, the H3K4me3 profiles at the DJCβ regions extended throughout the transcribed regions with no particular enrichment at the 5′ sides. Moreover, we found that H3K4me3 levels were 3- to-4 fold higher at the DJCβ regions than the level observed around the TSS of highly expressed genes (Fig. 1D). To directly compare the H3K4me3 enrichment within the gene body of individual genes, we calculated the density of H3K4me3 at the two DJCβ clusters and within the genomic regions from the TSS to +8 kb of mRNA genes. We next plotted the H3K4me3 values in the function of mRNA levels, obtained by polyA RNA-seq (see *Materials and Methods*). As shown in Fig. 1E, the two DJCβ clusters displayed very high levels of H3K4me3 as compared with the rest of the genes, whereas the mRNA level of the two clusters was relatively modest. We observed that a relatively small subset of genes also displayed elevated H3K4me3 enrichment (Fig. 1E). Genes ranked in function of H3K4me3 level identified 59 genes harboring substantially higher levels of H3K4me3 (Fig. 1F; see *Materials and Methods* for details). These genes displayed a broad distribution of H3K4me3 within the 5′ regions of the gene body (Fig. 1D; hereafter named Broad-H3K4me3 genes), as observed for the *Tcrb* locus, and reminiscent of previous findings of genes associated with extended H3K4 methylation (12, 13). However, the two DJCβ clusters ranked within the top 10 of the highest H3K4me3-enriched genes in ΔRag thymocytes (Fig. 1F). Thus, the active DJCβ clusters display an unusual H3K4me3-extended chromatin structure that is larger and stronger than the one observed at the vast majority of expresssed genes, without being associated with a high level of polyadenylated RNA.

Pol II–dependent chromatin remodeling

The above results raise the question of whether a specialized transcription mechanism plays a key role at the *Tcrb* locus, which ultimately leads to a highly accessible chromatin structure at the Dβ and Jβ gene segments. Chromatin accessibility at the AR loci has been generally associated with the presence of germline transcription (9). Moreover, H3K4me3 marking across the Jα segments of the *Tcra* locus has been shown to directly depend on germline transcription (42). More generally, functional links have been described between Pol II binding and H3K4 trimethylation at promoter regions (43, 44). Thus, we asked whether the atypical H3K4me3 profiles observed at the DJCβ regions may depend on Pol II–mediated transcription. To this end, elongating Pol II was blocked by inhibiting the CDK9 kinase with the KM05283 chemical compound (31, 34). We reasoned that if H3K4 trimethylation depends on local Pol II transcription, then its level was likely to decrease following the KM05283 treatment. In these experiments we used the pro–T cell line P5424, which is derived from ΔRag thymocytes and harbors a recombination-competent *Tcrb* locus (45). Efficient blocking of Pol II elongation upon KM05283 treatment was validated by global loss of phosphorylated serine 2 of the C-terminal domain (CTD) of Pol II (Ser^{2P} Pol II), as assessed by Western blot (Fig. 2A), as well as complete loss at the *Tcrb* region of H3K36me3, a mark of transcription elongation (46; Fig. 2B). Next, we performed ChIP-seq experiments for H3K4me3 and total-Pol II, using both KM05283- and DMSO-treated chromatin. Interestingly, we observed a strong decrease in H3K4me3 within the two DJCβ transcription units in KM05283-treated P5424 cells (Fig. 2C; note that the H3K4me3 profile at the *Tcrb* locus was consistent between the P5424 cell line and ΔRag thymocytes). Thus, H3K4 trimethylation at the DJCβ regions is largely dependent on Pol II transcription.

In most expressed genes, transcription initiation and elongation are regulated independently. Indeed, inhibition of transcriptional elongation normally results in the loss of Pol II within the gene body and its accumulation at promoter regions (36). This result was confirmed by increased Pol-pausing index in the KM05283-treated cells (Fig. 2D) and exemplified by the average profiles of the set of 300 highly expressed genes (Fig. 2E, *left panel*), as well as visual inspection of several expressed genes (Fig. 2F). Strikingly, however, we observed a complete loss of Pol II binding at the two DJCβ regions after KM05283 treatment (Fig. 2C, 2E, *right panel*). The specific loss of Pol II binding at Dβ promoters, but not at control genes, upon inhibition of transcription elongation was further confirmed by independent ChIP-qPCR experiments (Fig. 2G). Thus, in the absence of transcription elongation, Pol II was unable to remain stably associated to the Dβ promoters, as is the case for the vast majority of expressed genes. These results suggest that, at the *Tcrb* locus, recruitment of Pol II is directly coupled to the elongation phase of transcription.

TIPs cover the DJCβ transcription units

The above results suggest that 1) the broad distribution of H3K4me3 (and likely chromatin accessibility) is linked to Pol II–mediated transcription and 2) the regulation of the transcription process might differ between *Tcrb* locus and canonical mRNA coding genes. Pol II transcriptional activity is regulated via phosphorylation of the CTD (47). At expressed genes, phosphorylation of Ser^{5} of the CTD, which is associated with transcription initiation and early elongation, is found at the 5′ end of genes, whereas Ser^{2P}, which is required for productive elongation, is found to be enriched at the 3′ end of genes (43, 44). Moreover, Pol II phosphorylation at Ser^{5} has been shown to be required for H3K4me3 trimethylation (43, 44). To explore whether the *Tcrb* locus displays a distinctive Pol II profile, we first analyzed the distribution of phosphorylated and total forms of Pol II at the DJCβ1 cluster by ChIP-qPCR from ΔRag thymocytes. The expected patterns of Pol II phosphorylation were fully reproduced at control active genes *Actb* and *Sfrs*3: we found high levels of Ser^{5P} at the TSS, low levels of phosphorylated Pol II within the gene body, and high levels of Ser^{2P} at the 3′ end of these genes (Fig. 3A). However, in the case of the *Tcrb* locus, we found relatively high levels of Ser^{5P} Pol II throughout the DJCβ1 region, whereas the Ser^{2P} Pol II accumulated at the 3′ end of the DJCβ1 transcription unit (Fig. 3A). Indeed, although Ser^{5P} Pol II downstream of the TSS of control genes is reduced to background levels, the enrichment at equivalent regions of the DJCβ1 cluster remains elevated.

To have a more comprehensive view of Pol II profiles at the *Tcrb* locus, we performed ChIP-seq experiments for both total- and Ser^{5P} Pol II in ΔRag thymocytes. Again, we observed an accumulation of total- and Ser^{5P} Pol II across the two DJCβ regions (Fig. 3B), whereas control genes displayed the expected patterns (Fig. 3C). Note that total- and Ser^{5P} Pol II profiles were consistent between the ChIP-qPCR and ChIP-seq data (compare Figs. 3A with 3B, 3C). A more thorough analysis revealed that the Ser^{5P} Pol II profiles were quantitatively and qualitatively different between the set of highly expressed genes and the DJCβ regions (Fig. 4A). Indeed, the level of Ser^{5P} Pol II at the DJCβ regions was higher than the majority of expressed genes (Fig. 4B; the DJCβ clusters ranked in the top three of the highest Ser^{5P} Pol II–enriched genes in ΔRag thymocytes). Thus, Pol II is found in its initiating/early elongating form throughout the entire DJCβ transcription units.

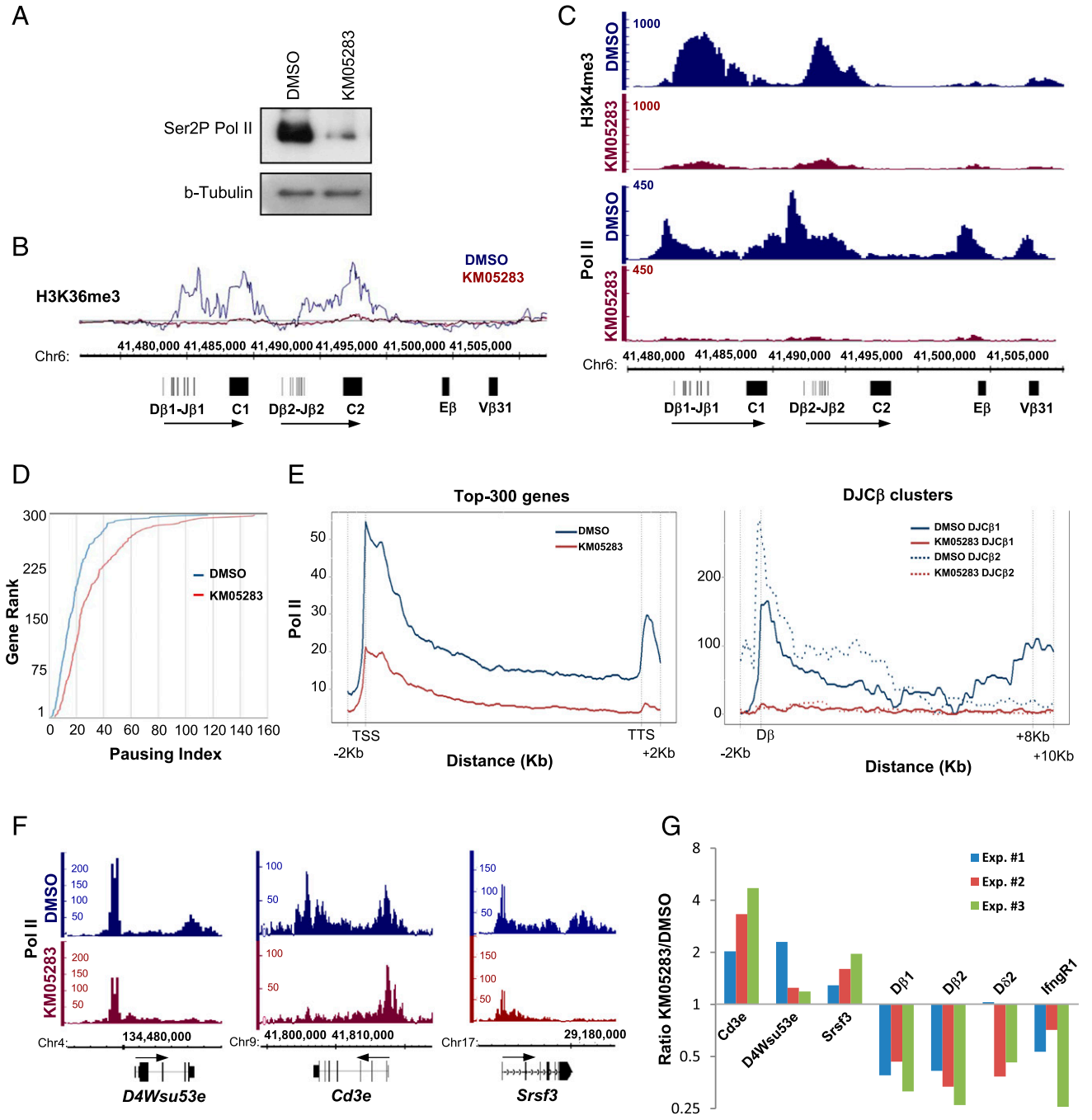

FIGURE 2. Pol II–mediated trimethylation at the DJCβ regions. (**A**) The level of Ser^{2P} Pol II was assessed by Western blot in P5424 cells treated with either DMSO or KM05283 for 18 h. (**B**) ChIP-on-chip profiles after inhibiting the elongation activity of Pol II showed a significant reduction in H3K36me3 signals at the DJCβ regions compared with the cells treated with DMSO. (**C**) Profiles of H3K4me3 and Pol II ChIP-seq from the P5424 cell line treated with KM05283 or DMSO at the 3′ region of the *Tcrb* locus. (**D**) The pausing index was calculated as previously described (32) for a selection of highly expressed genes (Top-300) in P5424 cells treated with either DMSO or KM05283. (**E**) Average profiles of Pol II ChIP-seq from the P5424 cell line treated with KM05283 or DMSO centered on the Top-300 of expressed genes (*left panel*) or the DJCβ regions (*right panel*). (**F**) Profiles of H3K4me3 and Pol II ChIP-seq from the P5424 cell line treated with KM05283 or DMSO at three control genes. Arrows indicate the sense of transcription. (**G**) ChIP-qPCR analyses of Pol II binding in the P5424 cell line treated with KM05283 or DMSO at the promoters of the indicated genes. The KM05283/DMSO ratio is shown in a log2 scale. TTS, transcriptional termination site.

A hallmark of transcription initiation in higher eukaryotes is the presence of bidirectional short-RNAs around the TSS (hereafter short-RNA), a feature related to Pol II pausing (48). Given the above results, we hypothesized that the DJCβ regions might be enriched in initiating short transcripts. To explore this possibility, we performed short-RNA–seq experiments from ΔRag thymocytes and compared them with strand-specific total (ribosomal-depleted) and polyA RNA-seq profiles previously generated (35). As expected, total and polyA RNA-seq signals overlapped with the DJCβ regions and were oriented in the sense of defined transcription units (Fig. 3C). The continuous RNA-seq signal observed at the DJCβ regions probably reflects a low splicing efficiency at this locus. Analysis of short-RNA–seq data revealed the presence of several discrete peaks of short transcripts, along with an overall enrichment of this RNA population throughout the entire DJCβ regions, suggesting that

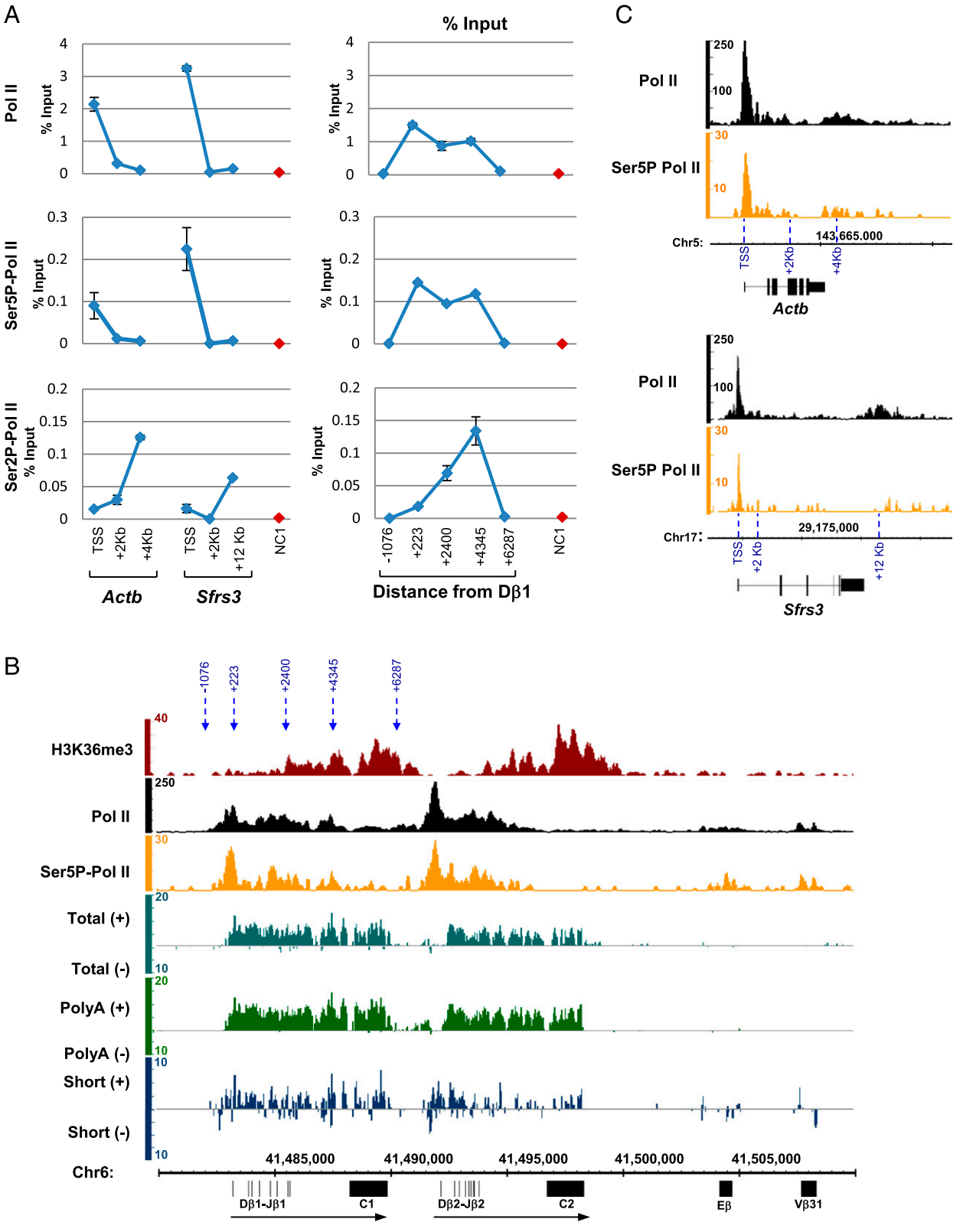

FIGURE 3. The DJCβ regions are enriched for transcription initiation features. (**A**) ChIP-qPCR assays showing the relative enrichment of total (N20), Ser^{2} and Ser^{5} phosphorylated Pol II at two active genes, at the indicated locations of the DJCβ1 region and at a negative control region (NC1). The genomic location of primer sets with respect to the Dβ1 gene segment or control genes is highlighted in (**B**) and (**C**). (B) Profiles of total and initiating (Ser^{5P}) Pol II ChIP-seq experiments, as well as total polyA and short-RNAs from directional RNA-seq experiments in ΔRag thymocytes, are shown at the 3′ region of the *Tcrb* locus. For RNA-seq the profiles are log2 scaled, and strand orientation is indicated at the *left* of each panel. Other data are as in Fig. 1. (C) The ChIP-Seq profiles of total and Ser^{5P} Pol II in ΔRag thymocytes are shown at three active genes.

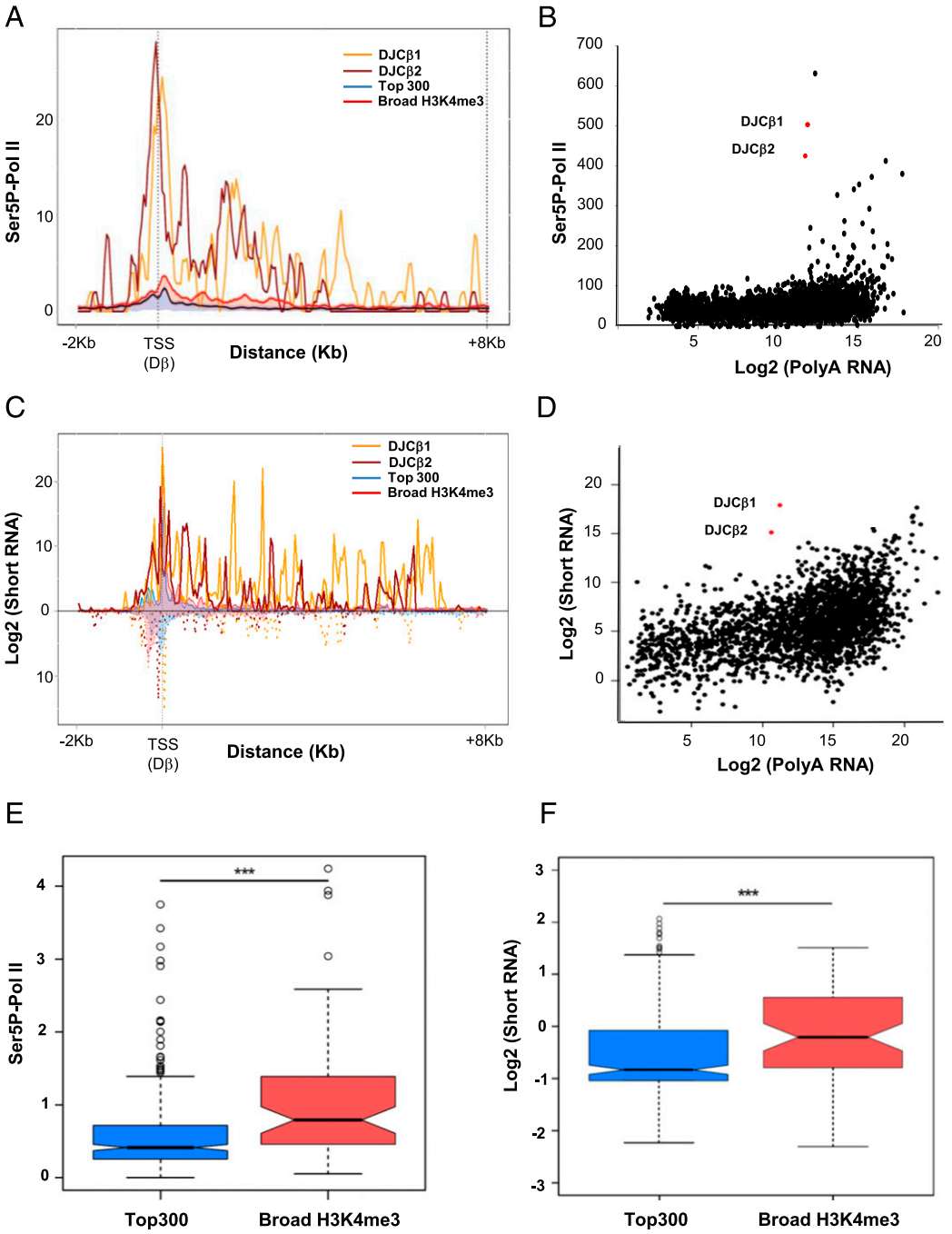

FIGURE 4. Global analyses of transcription initiation features. (**A**) Average profiling of Ser^{5P} Pol II ChIP-seq for the set of highly expressed genes (Top-300) and the set of Broad-H3K4me3 genes, compared with the profiles of the two DJCβ regions. The shadow represents the border of the 25th and 75th percentiles. (**B**) Scatter plot showing the Ser^{5P} Pol II density in the gene body of Refseq genes and in the two DJCβ clusters in the function of polyA RNA levels. (**C**) Average profiling of short-RNA–seq for the set of highly expressed genes (Top-300) and the set of Broad-H3K4me3 genes, compared with the profiles of the two DJCβ regions. (**D**) Scatter plot showing the short-RNA–seq density in the gene body of Refseq genes and in the two DJCβ clusters in the function of polyA RNA levels. (**E** and **F**) Boxplot showing the distribution of Ser^{5P} Pol II and short-RNA densities (between +1 and +4 kb from the TSS) of the Top-300 and Broad-H3K4me3 gene sets. Statistical significance was assessed by Student *t* test. $***p < 10^{-3}$.

Pol II pausing occurs at different places downstream of the Dβ promoters (Fig. 4C, 4D). This was a specific feature of the *Tcrb* locus, as the overall distribution of short-RNAs was clearly different between the DJCβ regions and the set of highly expressed genes, for which bidirectional short-RNAs accumulate around the TSS (Fig. 4C, 4D). Previously we have identified TIPs (14), which are large genomic regions associated with Ser^{5P} Pol II and TBP. TIPs were also associated with high levels of H3K4me3. The *Tcrb* might represent an extreme example of these genomic features. We concluded that the entire DJCβ regions behave as transcription initiating and early elongating platforms, thus providing a direct link between Pol II–mediated chromatin remodeling and H3K4 trimethylation at the Dβ/Jβ recombination segments.

Shared features between Tcrb *and Broad-H3K4me3 genes*

As mentioned above, a small subset of genes was found to be associated with broad H4K4me3 marking (Fig. 1D–F). These

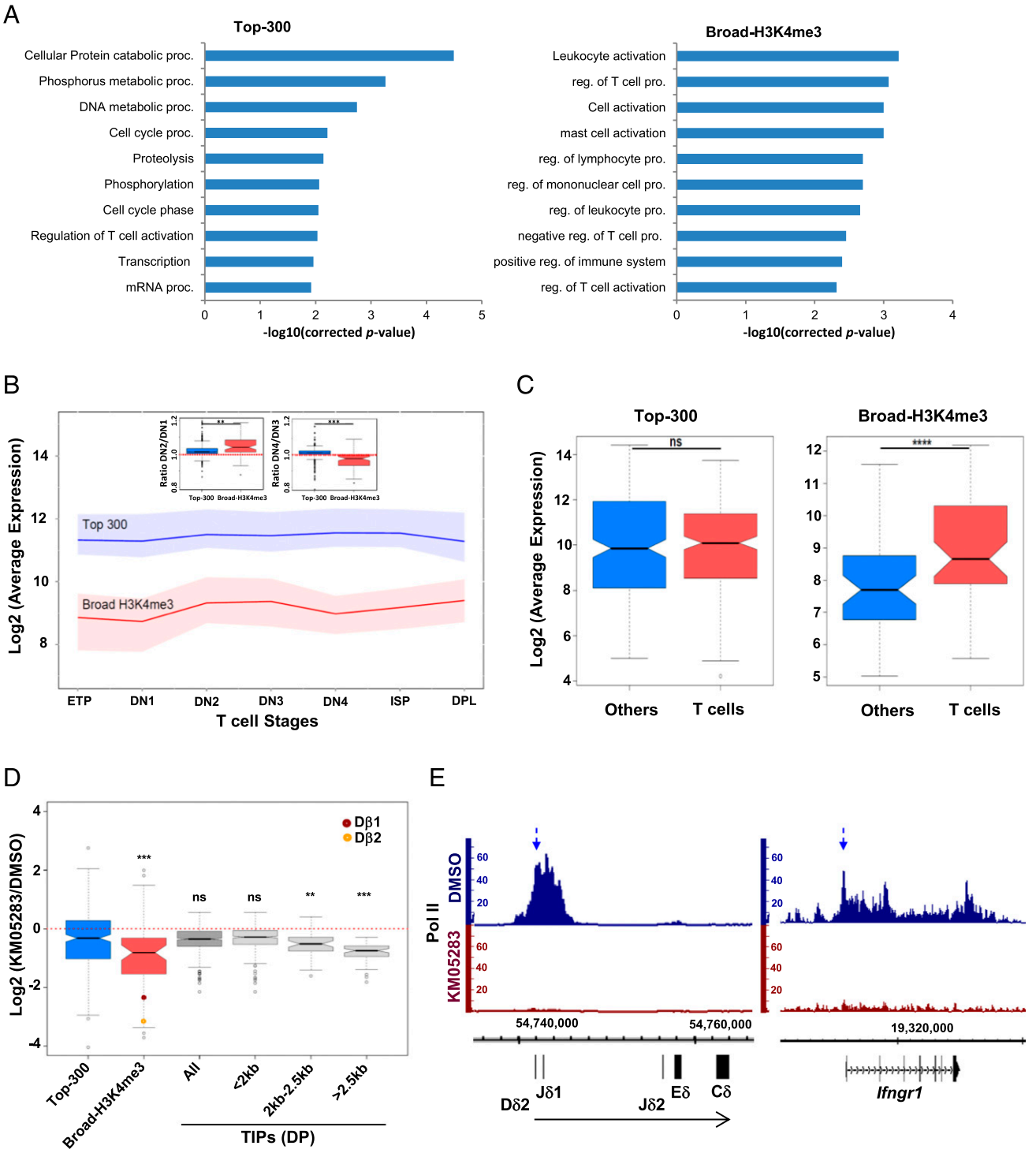

FIGURE 5. Characterization of Broad-H3K4me3 genes. (**A**) Enrichment scores of top 10 biological processes found enriched in the indicated gene sets. (**B**) Expression dynamic of Top-300 and Broad-H3K4me3 genes during T cell differentiation. Cell populations are as defined by the Immunological Genome Project and are ordered from less to most differentiated stages: ETP, early thymic progenitors; DN1–4, DN 1–4; ISP, immature single positive; and DPL, DP Blast, respectively. The *inset* shows the differential expression between the DN1-to-DN2 and DN3-to-DN4 transitions. (**C**) Boxplot showing the mean expression distribution in T cell samples and 69 other samples (see *Materials and Methods* for details). (**D**) Boxplot showing the distribution of the KM05283/DMSO ratio of Pol II binding at the promoters (±2 kb around the TSS) of the Top-300, Broad-H3K4me3, and TIPs-associated genes, as well as size-selected subsets of TIPs. The KM05283/DMSO ratios of the two DJCβ regions are also indicated. Statistical significance in (B) through (D) was calculated by the Student *t* test. (**E**) Profiles of Pol II ChIP-seq from the P5424 cell line treated with KM05283 or DMSO at the *Tcrd* locus and one Broad-H3K4me3 gene. The genomic location of primer sets used in Fig. 2G are indicated by dashed arrows. The *p* values are as follows: $^{**}p < 10^{-2}$, $^{***}p < 10^{-3}$, $^{****}p < 10^{-4}$. Proc., process; reg., regulation.

genes also displayed significant enrichment of Ser^{5P} Pol II and, to a lesser extent, short initiating transcripts (Fig. 4E, 4F, Supplemental Fig. 1A, 1B). In general, genes with high levels of H3K4me3 also displayed high levels of Ser^{5} Pol II (Supplemental Fig. 1C). Therefore, a small subset of genes with broad H3K4me3 marking also displays features of transcriptional initiation in ΔRag

thymocytes (although *Tcrb* might represent an extreme example of this phenomenon).

To gain further insight into the function of Broad-H3K4me3 genes, we analyzed the functional enrichment of the biological process and found that they were specifically enriched on T cell– and immune-related functions, whereas the set of Top-300 genes were enriched for metabolic processes (Fig. 5A). Indeed, the list of Broad-H3K4me3 genes include many genes known to be involved in T cell differentiation and signaling, such as *Lef1, Il2ra, Themis, Ifngr1, Fyb, RhoH,* and *Cd274* (Supplemental Table II). Accordingly, the set of Broad-H3K4me3 genes was highly tissue specific (Fig. 5B). Although these genes were expressed at relatively low levels in primary thymocytes, their expression was highly regulated during early T cell differentiation (Fig. 5C), namely, between DN1-to-DN2 and DN3-to-DN4 cell transitions (Fig. 4C, *insets*). Thus, the subset of Broad-H3K4me3 genes is reminiscent of the *Tcrb* locus, as they represent highly regulated genes involved in T cell function. They might represent extreme examples of genes with broad H3K4 methylation patterns described previously by us and others (12–14). To assess whether other AR genes could share the same features as the *Trcb*, we analyzed, in a similar way, gene segments of the *Tcrd* and *Tcrg* locus, which are the two other AR loci in an open chromatin configuration in ΔRag thymocytes (see *Materials and Methods* for details). We found that gene segments from *Tcrd* (spanning Dδ2-Jδ1 gene segments) and *Tcrg* (Jγ1-Cγ1 and Jγ4-Cγ4 gene segments) loci also displayed high levels of H3K4me3, Ser^{5P} Pol II, and short initiating transcripts to a similar extent as those observed for the *Tcrb* locus (Supplemental Figs. 1C, 2), thus suggesting that large initiating platforms might be a general feature of AR loci.

Finally, we asked whether Pol II binding at Broad-H3K4me3 genes was also highly sensitive to transcription elongation, as observed for the *Tcrb* locus. Quantification of Pol II levels around the TSS of the Top-300 and Broad H3K4me3 genes in P5424 cells treated with either DMSO or KM05283 demonstrated that Pol II binding is specifically lost at BroadH4K4me3 genes, although not to the same extent as observed around the Dβ gene segments. This finding was evidenced at several genes, including the *Tcrd* and *Infgr1* loci (Fig. 5E), and validated by independent ChIP-qPCR (Fig. 2G; note that the *Tcrg* locus could not be analyzed, as this gene was found to be inactive in the P5424 cell line; data not shown). To determine whether this phenomenon was a general property of TIPs, we analyzed our previously defined selection of TIPs-associated genes in DP thymocytes (14), excluding the genes that were not expressed in the P5424 cell line (see *Materials and Methods* for details). As a group, the TIPs-associated genes did not display a loss of Pol II binding at their promoters after inhibition of Pol II elongation (Fig. 5D). However, when TIPs were classified according to their size, we found that genes associated with large TIPs (>2.5 kb) significantly lost Pol II binding at their promoters (Fig. 5D). We concluded that a subset of Broad-H3K4me3 genes and large TIPs-associated genes display regulatory features similar to those of the *Tcrb* locus, including tissue-specific gene expression, the presence of a TIP, and coupled Pol II recruitment and elongation (Fig. 6).

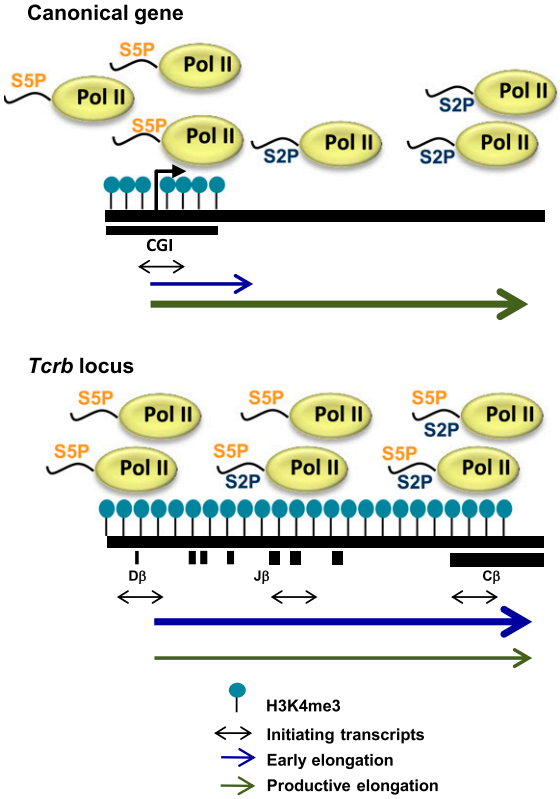

FIGURE 6. Model summarizing Pol II transcription at canonical expressed genes and DJCβ clusters. Promoters of active genes (*top panel*) are generally linked to CpG islands (CGI) and associated with high levels of initiating (Ser^{5P}) Pol II, bidirectional short transcripts, and H3K4me3. In contrast, at the *Tcrb* locus, transcription initiation hallmarks and H3K4me3 are found throughout the transcribed regions spanning the Dβ and Jβ gene segments. This feature might explain the unusual high chromatin accessibility observed at these regions.

Discussion

Previous work from our laboratory and other laboratories has shown a remarkable open chromatin structure encompassing the Dβ-Jβ recombination center, including chromatin accessibility and histone marking (3, 16, 22, 26–28). More specifically, H3K4me3 was found to be enriched at Dβ and Jβ segments using ChIP-qPCR (7). Similar extended H3K4me3 patterns have been shown across the Jα segments of the *Tcra* locus (42). In the current study, we extend these finding by showing that the distribution of H3K4 trimethylation over the DβJβ regions is both quantitatively and qualitatively different from that in the vast majority of expressed genes (Fig. 1). Although H3K4me3 generally accumulates within 2kb around the TSS of genes (11), we observed that H3K4me3 enrichment at the DJCβ clusters is much broader, spanning ≤8 kb downstream of the germline Dβ promoters, and including all Jβ gene segments. Moreover, the level of H3K4 trimethylation found at the DJCβ clusters was exceptionally high, representing one of the most enriched domains in developing thymocytes. We show that this extended profile depends on an unusual Pol II regulation process. In the case of canonical genes, Pol II accumulates around the TSS in its initiating form (high Ser^{5P} Pol II), which correlates with high enrichment of H3K4me3 and the presence of short-RNA transcripts. However, in the case of the *Tcrb* locus, the entire DJCβ regions display features of transcription initiation and Pol II pausing, including high levels of Ser^{5P} Pol II and short-RNAs. Unexpectedly, inhibition of Pol II elongation resulted in complete loss of Pol II across the DJCβ clusters (Fig. 2). To our knowledge, this is the first example in mammals whereby Pol II accumulation at the promoter is strictly dependent on transcription elongation. Remarkably, this phenomenon was also observed at the *Tcrd* locus (Figs. 2G, 5E). We propose that a high level of initiating Pol II throughout the entire DJCβ regions targets the H3K4 histone methyltransferases, resulting in an unusual extended H3K4me3 profile, and ultimately

leads to a highly accessible chromatin structure around the Dβ and Jβ gene segments (Fig. 6).

We have previously shown that tissue-specific genes expressed in T cells generally display high levels of H3K4 methylation within the 5′ region of the gene body (12). Along the same line, a recent study has shown that H3K4me3 domains that spread more broadly over genes in a given cell type preferentially mark genes that are essential for the identity and function of that cell type (13). Besides, we also described TIPs at proximal and distal sites, which were characterized by the presence of Ser^{5P} Pol II, TBP, and epigenetic marks H3K4me1 and H3K4me3 (14). In this article, we show that genes with broad H3K4me3 domains display features related to large initiation platforms (including accumulation of Ser^{5P} Pol II and short initiating transcripts) similar to the TIP genomic domains. However, TIP domains as defined previously in DP thymocytes (14) display a wide range of size, varying from 0.45 kb to 10 kb (80% of TIPs are <2 kb). Whether broad H3K4me3 and TIPs define the same type of genes remains to be precisely investigated, but our results suggest that common features are shared by both types of structures. Genes marked by the broadest H3K4me3 domains exhibit enhanced transcriptional consistency rather than increased transcriptional levels (13). Moreover, Pol II accumulation at the promoter of Broad-H3K4me3 genes tends to be dependent on transcription elongation, a phenomenon also observed at the promoters of genes associated with large TIPs (Fig. 5D). Thus, it is likely that the broad H3K4me3 domains defined in this article (in particular, those found at the *Tcr* loci) might represent a subset of larger TIPs. Indeed, larger TIPs also have a tendency to be more tissue specific (14). All in all, our results suggest the existence of a specialized transcriptional regulation mechanism restricted to a subset of tissue-specific genes. In this context, the *Tcrb* locus might represent an extreme example of this phenomenon. Our finding has implications not only for regulatory strategies used by AR loci but also for the epigenetic mechanisms that control gene expression of cell identity genes.

Are intrinsic genomic features responsible for the highly open and H3K4me3-enriched chromatin structure observed at the *Tcrb* locus? In mammals, Pol II accumulation and enrichment for active histone marks at promoters are generally linked to their high CpG content (49). We have previously shown that TIPs overlapped with CpG density, although larger TIPs displayed lower or more disperse CpG content (14). Consistently, we found that promoters of Broad-H3K4me3 and large TIPs-associated genes display significantly lower CpG density as compared with the set of Top-300 genes ($p < 0.01$ and $p < 0.0001$, respectively; Student t test; see *Materials and Methods*). The DJCβ regions do not contain any CpG island and also display relatively low G and C nucleotide content (data not shown). It is, therefore, plausible that in the absence of CpG islands, the Pol II molecules recruited at the Dβ associated promoters are immediately engaged in the elongation process while still harboring the transcription initiation mark (i.e., Ser^{5P}) and therefore remain associated with H3K4 methyltransferases (43, 44)(Fig. 6). This hypothesis would be consistent with the complete loss of Pol II at the *Tcrb* locus after inhibition of transcription elongation (Fig. 2). Another intriguing, but not mutually exclusive, possibility is that the extended H3K4me3 profile is related to the unusual structure of the *Tcrb* locus, which contains several J segments, each harboring a 5′ splicing site. A recent study has shown that H3K4 trimethylation at the 5′ border of mammalian genes is directly linked to the length of the first exon of genes (average size is 250 nt) (50). However, in the case of the DJCβ transcription units, the first splicing donors are located at the end of each Jβ segment, ranging between 641 nt and 2.5 kb from the Dβ segments, which make the first exons considerably longer than the average size. Moreover, the Jβ-associated splicing sites appear to be relatively inefficient, as judged by the high level of RNA-seq signal observed downstream of the Jβ gene segments (Fig. 3B). As described previously (50), the first exon length >500 nt results in a flat H3K4me3 profile extending to the 3′ end of the first exon, as well as increasing Pol II pausing, both features reminiscent of what is observed at the *Tcrb* locus. Thus, it is plausible that the location of Jβ gene segments, each behaving as a first exon, will result in the distinctive chromatin structure observed at the DJCβ clusters.

It has been recently demonstrated that RAG1 and RAG2 bind in vivo to focal regions, termed "recombination centers," covering mainly the J segments of AR genes and within which V(D)J recombination has been suggested to take place (7). The formation of these recombination centers depends on the AR enhancers and promoters (6), and correlates with the presence of H3K4me3 (7). Thus, given the specific requirements for chromatin accessibility and H3K4me3 enrichment at J segments to ensure efficient V(D)J recombination (3, 5), we propose that the *Tcrb* locus (and likely other AR loci) has evolved in such a way that a specialized regulation of the transcription process confers a unique long-range epigenetic marking, ultimately allowing the establishment of a highly accessible chromatin structure at the recombining Dβ/Jβ gene segments.

Data access

ChIP-seq and RNA-seq data obtained in this study have been submitted to the National Center for Biotechnology Information's GEO (http://www.ncbi.nlm.nih.gov/geo) under the following accession numbers: GSE63416 (www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE63416), GSE64709 (www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE64709), and from GSM1360722 to GSM1360727 and GSM1359828 (www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE56395). Details are available in Supplemental Table I.

Acknowledgments

We thank Dr. Eugene Oltz (Washington University, St. Louis, MO) for donating the P5424 cell line.

Disclosures

The authors have no financial conflicts of interest.

References

1. Schatz, D. G. 2004. V(D)J recombination. *Immunol. Rev.* 200: 5–11.
2. Jung, D., and F. W. Alt. 2004. Unraveling V(D)J recombination; insights into gene regulation. *Cell* 116: 299–311.
3. Osipovich, O., and E. M. Oltz. 2010. Regulation of antigen receptor gene assembly by genetic-epigenetic crosstalk. *Semin. Immunol.* 22: 313–322.
4. Yancopoulos, G. D., and F. W. Alt. 1986. Regulation of the assembly and expression of variable-region genes. *Annu. Rev. Immunol.* 4: 339–368.
5. Spicuglia, S., J. Zacarias-Cabeza, A. Pekowska, and P. Ferrier. 2010. Epigenetic regulation of antigen receptor gene rearrangement. *F1000 Biol. Rep.* 2: 2.
6. Ji, Y., A. J. Little, J. K. Banerjee, B. Hao, E. M. Oltz, M. S. Krangel, and D. G. Schatz. 2010. Promoters, enhancers, and transcription target RAG1 binding during V(D)J recombination. *J. Exp. Med.* 207: 2809–2816.
7. Ji, Y., W. Resch, E. Corbett, A. Yamane, R. Casellas, and D. G. Schatz. 2010. The in vivo pattern of binding of RAG1 and RAG2 to antigen receptor loci. *Cell* 141: 419–431.
8. Matthews, A. G., and M. A. Oettinger. 2009. RAG: a recombinase diversified. *Nat. Immunol.* 10: 817–821.
9. Abarrategui, I., and M. S. Krangel. 2009. Germline transcription: a key regulator of accessibility and recombination. *Adv. Exp. Med. Biol.* 650: 93–102.
10. Zhou, V. W., A. Goren, and B. E. Bernstein. 2011. Charting histone modifications and the functional organization of mammalian genomes. *Nat. Rev. Genet.* 12: 7–18.
11. Barski, A., S. Cuddapah, K. Cui, T. Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev, and K. Zhao. 2007. High-resolution profiling of histone methylations in the human genome. *Cell* 129: 823–837.
12. Pekowska, A., T. Benoukraf, P. Ferrier, and S. Spicuglia. 2010. A unique H3K4me2 profile marks tissue-specific gene regulation. *Genome Res.* 20: 1493–1502.

13. Benayoun, B. A., E. A. Pollina, D. Ucar, S. Mahmoudi, K. Karra, E. D. Wong, K. Devarajan, A. C. Daugherty, A. B. Kundaje, E. Mancini, et al. 2014. H3K4me3 breadth is linked to cell identity and transcriptional consistency. *Cell* 158: 673–688.
14. Koch, F., R. Fenouil, M. Gut, P. Cauchy, T. K. Albert, J. Zacarias-Cabeza, S. Spicuglia, A. L. de la Chapelle, M. Heidemann, C. Hintermair, et al. 2011. Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nat. Struct. Mol. Biol.* 18: 956–963.
15. Jackson, A. M., and M. S. Krangel. 2006. Turning T-cell receptor β recombination on and off: more questions than answers. *Immunol. Rev.* 209: 129–141.
16. Spicuglia, S., A. Pekowska, J. Zacarias-Cabeza, and P. Ferrier. 2010. Epigenetic control of Tcrb gene rearrangement. *Semin. Immunol.* 22: 330–336.
17. Michie, A. M., and J. C. Zúñiga-Pflücker. 2002. Regulation of thymocyte differentiation: pre-TCR signals and beta-selection. *Semin. Immunol.* 14: 311–323.
18. Krimpenfort, P., R. de Jong, Y. Uematsu, Z. Dembic, S. Ryser, H. von Boehmer, M. Steinmetz, and A. Berns. 1988. Transcription of T cell receptor beta-chain genes is controlled by a downstream regulatory element. *EMBO J.* 7: 745–750.
19. McDougall, S., C. L. Peterson, and K. Calame. 1988. A transcriptional enhancer 3′ of C beta 2 in the T cell receptor beta locus. *Science* 241: 205–208.
20. Bories, J. C., J. Demengeot, L. Davidson, and F. W. Alt. 1996. Gene-targeted deletion and replacement mutations of the T-cell receptor beta-chain enhancer: the role of enhancer elements in controlling V(D)J recombination accessibility. *Proc. Natl. Acad. Sci. USA* 93: 7871–7876.
21. Bouvier, G., F. Watrin, M. Naspetti, C. Verthuy, P. Naquet, and P. Ferrier. 1996. Deletion of the mouse T-cell receptor beta gene enhancer blocks alphabeta T-cell development. *Proc. Natl. Acad. Sci. USA* 93: 7877–7881.
22. Mathieu, N., W. M. Hempel, S. Spicuglia, C. Verthuy, and P. Ferrier. 2000. Chromatin remodeling by the T cell receptor (TCR)-β gene enhancer during early T cell development: Implications for the control of TCR-β locus recombination. *J. Exp. Med.* 192: 625–636.
23. Spicuglia, S., S. Kumar, J. H. Yeh, E. Vachez, L. Chasson, S. Gorbatch, J. Cautres, and P. Ferrier. 2002. Promoter activation by enhancer-dependent and -independent loading of activator and coactivator complexes. *Mol. Cell* 10: 1479–1487.
24. Ryu, C. J., B. B. Haines, D. D. Draganov, Y. H. Kang, C. E. Whitehurst, T. Schmidt, H. J. Hong, and J. Chen. 2003. The T cell receptor β enhancer promotes access and pairing of Dbeta and Jbeta gene segments during V(D)J recombination. *Proc. Natl. Acad. Sci. USA* 100: 13465–13470.
25. Oestreich, K. J., R. M. Cobb, S. Pierce, J. Chen, P. Ferrier, and E. M. Oltz. 2006. Regulation of TCRbeta gene assembly by a promoter/enhancer holocomplex. *Immunity* 24: 381–391.
26. Bonnet, M., F. Huang, T. Benoukraf, O. Cabaud, C. Verthuy, A. Boucher, S. Jaeger, P. Ferrier, and S. Spicuglia. 2009. Duality of enhancer functioning mode revealed in a reduced TCR β gene enhancer knockin mouse model. *J. Immunol.* 183: 7939–7948.
27. Whitehurst, C. E., S. Chattopadhyay, and J. Chen. 1999. Control of V(D)J recombinational accessibility of the D beta 1 gene segment at the TCR beta locus by a germline promoter. *Immunity* 10: 313–322.
28. Morshead, K. B., D. N. Ciccone, S. D. Taverna, C. D. Allis, and M. A. Oettinger. 2003. Antigen receptor loci poised for V(D)J rearrangement are broadly associated with BRG1 and flanked by peaks of histone H3 dimethylated at lysine 4. *Proc. Natl. Acad. Sci. USA* 100: 11577–11582.
29. Shinkai, Y., G. Rathbun, K. P. Lam, E. M. Oltz, V. Stewart, M. Mendelsohn, J. Charron, M. Datta, F. Young, A. M. Stall, and F. W. Alt. 1992. RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell* 68: 855–867.
30. Mombaerts, P., C. Terhorst, T. Jacks, S. Tonegawa, and J. Sancho. 1995. Characterization of immature thymocyte lines derived from T-cell receptor or recombination activating gene 1 and p53 double mutant mice. *Proc. Natl. Acad. Sci. USA* 92: 7420–7424.
31. Medlin, J., A. Scurry, A. Taylor, F. Zhang, B. M. Peterlin, and S. Murphy. 2005. P-TEFb is not an essential elongation factor for the intronless human U2 snRNA and histone H2b genes. *EMBO J.* 24: 4154–4165.
32. Chapman, R. D., M. Heidemann, T. K. Albert, R. Mailhammer, A. Flatley, M. Meisterernst, E. Kremmer, and D. Eick. 2007. Transcribing RNA polymerase II is phosphorylated at CTD residue serine-7. *Science* 318: 1780–1782.
33. Giresi, P. G., J. Kim, R. M. McDaniell, V. R. Iyer, and J. D. Lieb. 2007. FAIRE (formaldehyde-assisted isolation of regulatory elements) isolates active regulatory elements from human chromatin. *Genome Res.* 17: 877–885.
34. Pekowska, A., T. Benoukraf, J. Zacarias-Cabeza, M. Belhocine, F. Koch, H. Holota, J. Imbert, J. C. Andrau, P. Ferrier, and S. Spicuglia. 2011. H3K4 trimethylation provides an epigenetic signature of active enhancers. *EMBO J.* 30: 4198–4210.
35. Lepoivre, C., M. Belhocine, A. Bergon, A. Griffon, M. Yammine, L. Vanhille, J. Zacarias-Cabeza, M. A. Garibal, F. Koch, M. A. Maqbool, et al. 2013. Divergent transcription is associated with promoters of transcriptional regulators. *BMC Genomics* 14: 914.
36. Rahl, P. B., C. Y. Lin, A. C. Seila, R. A. Flynn, S. McCuine, C. B. Burge, P. A. Sharp, and R. A. Young. 2010. c-Myc regulates transcriptional pause release. *Cell* 141: 432–445.
37. Benoukraf, T., P. Cauchy, R. Fenouil, A. Jeanniard, F. Koch, S. Jaeger, D. Thieffry, J. Imbert, J. C. Andrau, S. Spicuglia, and P. Ferrier. 2009. CoCAS: a ChIP-on-chip analysis suite. *Bioinformatics* 25: 954–955.
38. Heng, T. S., M. W. Painter, Immunological Genome Project Consortium. 2008. The Immunological Genome Project: networks of gene expression in immune cells. *Nat. Immunol.* 9: 1091–1094.
39. Huber, W., A. von Heydebreck, H. Sültmann, A. Poustka, and M. Vingron. 2002. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* 18(Suppl 1): S96–S104.
40. Dennis, G., Jr., B. T. Sherman, D. A. Hosack, J. Yang, W. Gao, H. C. Lane, and R. A. Lempicki. 2003. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* 4: 3.
41. Carabana, J., A. Watanabe, B. Hao, and M. S. Krangel. 2011. A barrier-type insulator forms a boundary between active and inactive chromatin at the murine TCRβ locus. *J. Immunol.* 186: 3556–3562.
42. Abarrategui, I., and M. S. Krangel. 2006. Regulation of T cell receptor-α gene recombination by transcription. *Nat. Immunol.* 7: 1109–1115.
43. Selth, L. A., S. Sigurdsson, and J. Q. Svejstrup. 2010. Transcript elongation by RNA polymerase II. *Annu. Rev. Biochem.* 79: 271–293.
44. Buratowski, S. 2009. Progression through the RNA polymerase II CTD cycle. *Mol. Cell* 36: 541–546.
45. Osipovich, O., R. M. Cobb, K. J. Oestreich, S. Pierce, P. Ferrier, and E. M. Oltz. 2007. Essential function for SWI-SNF chromatin-remodeling complexes in the promoter-directed assembly of Tcrb genes. *Nat. Immunol.* 8: 809–816.
46. Guenther, M. G., S. S. Levine, L. A. Boyer, R. Jaenisch, and R. A. Young. 2007. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* 130: 77–88.
47. Heidemann, M., C. Hintermair, K. Voß, and D. Eick. 2013. Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. *Biochim. Biophys. Acta* 1829: 55–62.
48. Seila, A. C., J. M. Calabrese, S. S. Levine, G. W. Yeo, P. B. Rahl, R. A. Flynn, R. A. Young, and P. A. Sharp. 2008. Divergent transcription from active promoters. *Science* 322: 1849–1851.
49. Suzuki, M. M., and A. Bird. 2008. DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* 9: 465–476.
50. Bieberstein, N. I., F. Carrillo Oesterreich, K. Straube, and K. M. Neugebauer. 2012. First exon length controls active chromatin signatures and transcription. *Cell Reports* 2: 62–68.