Published online 17 March 2023

Nucleic Acids Research, 2023, Vol. 51, No. 10 ***4845–4866***
https://doi.org/10.1093/nar/gkad187

Short tandem repeats are important contributors to silencer elements in T cells

Saadat Hussain[1,2,†], Nori Sadouni[1,2,†], Dominic van Essen[3], Lan T.M. Dao[1,2,4], Quentin Ferré[1,2], Guillaume Charbonnier[1,2], Magali Torres[1,2], Frederic Gallardo[1,2], Charles-Henri Lecellier[5,6], Tom Sexton[7], Simona Saccani[3] and Salvatore Spicuglia[1,2,*]

[1]Aix-Marseille University, Inserm, TAGC, UMR1090, Marseille, France, [2]Equipe Labélisée Ligue Contre le Cancer, Marseille, France, [3]Institute for Research on Cancer and Ageing, IRCAN, 06107 Nice, France, [4]Present address: Vinmec Research Institute of Stem cell and Gene technology (VRISG), Hanoi, Vietnam, [5]Institut de Génétique Moléculaire de Montpellier, University of Montpellier, CNRS, Montpellier, France, [6]LIRMM, University of Montpellier, CNRS, Montpellier, France and [7]Institut de Génétique et de Biologie Moléculaire et Cellulaire – IGBMC (CNRS UMR 7104, INSERM U1258, Université de Strasbourg), 67404 Illkirch, France

Received May 09, 2022; Revised February 26, 2023; Editorial Decision February 28, 2023; Accepted March 15, 2023

ABSTRACT

The action of *cis*-regulatory elements with either activation or repression functions underpins the precise regulation of gene expression during normal development and cell differentiation. Gene activation by the combined activities of promoters and distal enhancers has been extensively studied in normal and pathological contexts. In sharp contrast, gene repression by *cis*-acting silencers, defined as genetic elements that negatively regulate gene transcription in a position-independent fashion, is less well understood. Here, we repurpose the STARR-seq approach as a novel high-throughput reporter strategy to quantitatively assess silencer activity in mammals. We assessed silencer activity from DNase hypersensitive I sites in a mouse T cell line. Identified silencers were associated with either repressive or active chromatin marks and enriched for binding motifs of known transcriptional repressors. CRISPR-mediated genomic deletions validated the repressive function of distinct silencers involved in the repression of non-T cell genes and genes regulated during T cell differentiation. Finally, we unravel an association of silencer activity with short tandem repeats, highlighting the role of repetitive elements in silencer activity. Our results provide a general strategy for genome-wide identification and characterization of silencer elements.

INTRODUCTION

The precise regulation of gene expression during normal development and cell differentiation requires the action of *cis*-regulatory elements with either activation or repression functions (1–3). Gene activation by the combined activities of promoters and distal enhancers has been extensively studied in normal and pathological contexts. In sharp contrast, gene repression by *cis*-acting silencers, defined as genetic elements that negatively regulate gene transcription in a position-independent fashion, is less well understood. Silencers were first described more than three decades ago in yeast and vertebrates (4–6). Since then, several silencers have been discovered to control the expression of key developmental and immunological model genes, and some progress has been made to characterize various features of a few of these individual silencers (3,7–13). Nevertheless, despite the widely-held belief that silencers likely represent critical general regulators of gene expression, this view is still largely conjectural, and their genome-wide distribution, mechanisms of action and involvement in disease are largely unknown. Noticeably, among the silencers that have been described in the literature, many are associated with the regulation of T cell specific genes. These included silencers associated with the expression of *TCA3/CCL1* (14), *Il2* (15), *CD4* (16), *Tcrb* (17), *ThPOK* (18), *Rag1-Rag2* (19), *CD8* (20) and *Spi1* loci (21). Several of these silencers have been shown to play an important role in cell lineage restriction: for instance, the CD4 and CD8 silencers repress the expression of the associated genes in $CD8^+$ and $CD4^+$ T cells, respectively. T cell differentiation thus provides an excellent model for the implementation of a high-throughput strategy to identify silencers.

*To whom correspondence should be addressed. Tel: +33 491828717; Email: salvatore.spicuglia@inserm.fr
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

Compared to other types of *cis*-regulatory elements, such as enhancers and insulators, silencers have been challenging to map genome-wide (22). Recent efforts included the development of a negative selection strategy (23) or prediction strategies based on chromatin signatures and 3D interactions with repressed genes (9,24–27). Episomal reporter assays have been widely used to characterize putative regulatory regions (1). The development of high-throughput reporter assays for enhancer function has enabled the testing of thousands of distinct DNA sequences simultaneously, by cloning variable DNA fragments into common reporter constructs and using high-throughput sequencing to quantify fragment activity (28). These functional approaches have led to an explosion of discoveries, including their roles in the regulation of many disease-related genes and their involvement in the development of diverse pathologies and cancer (29). In particular, the Self Transcribing Active Regulatory Region Sequencing (STARR-seq) method allows direct genome-wide investigation of enhancer activity using DNA fragments directly collected from genomic DNA (30). Based on this technique, we previously developed the CapSTARR-seq approach (31), coupling capture of regions of interest to STARR-seq reporter assay, providing a cost-effective method to assess *cis*-regulatory function in mammals.

In line with their operational definitions, assays for silencers could measure their ability to silence gene expression in *cis*, when driven by an independent 'strong' promoter (1,10,32). Therefore, to identify silencers genome-wide we sought to repurpose the STARR-seq approach by systematically testing DNase hypersensitive sites (DHS) from developing T cells, using three distinct promoter-based reporter vectors. We compared the set of silencers identified with the different vectors and evaluated their association with genomic and epigenomic features. The robustness of the approaches was extensively assessed by independent episomal reporter assays, and CRISPR/Cas9 genomic manipulation demonstrated the involvement of several endogenous silencers in the repression of neighboring genes. Repetitive elements (REs) have been suggested to contribute to *cis*-regulation, including gene repression (33,34). Here, we found that short tandem repeats (STRs) play an important role in silencing activity. Overall, we provide a general, scalable, and high-throughput approach for the high-resolution experimental dissection of silencer elements in the context of human biology and disease.

MATERIALS AND METHODS

Cell culture

Mouse P5424 T cells (35) and mouse embryonic fibroblast NIH-3T3 (3T3, ATCC: CRL-1658) cells were cultured in RPMI medium (Thermo Fisher Scientific) supplemented with heat-inactivated 10% FBS (PAA) at 37°C, 5% CO_2. Cells were passaged every 2–3 days and frequently tested for mycoplasma contamination.

Stimulation of P5424 cells

P5424 cells were grown at a density of 3×10^5 cells/ml. Cells were treated with DMSO or PMA at 10 ng/ml (P1585, Sigma) and ionomycin at 0.5 μg/ml (I3909, Sigma) for 4 h in triplicates as previously described (36).

Cloning of the STARR-seq vectors

The STARR-seq mammalian screening vector (30) has been kindly provided by Alexander Stark (Vienna, Austria). The synthetic SCP1 promoter present in the enhancer-STARR-seq vector was replaced by the promoter of the ubiquitous PGK gene as defined in (37) or the lymphoid-specific promoter of the *Rag2* gene as defined in (38) using In-Fusion homologous recombination. In the R-Ea construct the TCRα enhancer (38) was cloned downstream of the GFP cassette.

CapSTARR-seq library generation

Construction and capture of the genomic library have been described previously (31,39). The genomic library was generated from mouse C57BL/6 genomic DNA. For target enrichment, a custom-designed 3 nt resolution oligonucleotide microarray covering 28055 DHSs identified in mouse $CD4^+$ $CD8^+$ double-positive (DP) thymocytes (31) was constructed using the SureSelect technology (Agilent, 1M format) and the eArray tool default settings (https://earray.chem.agilent.com/earray/). In addition, 437 randomly-selected non-DHS regions were included. The three screening vectors were linearized with AgeI-HF and SalI-HF (New England Biolabs) by 6 h digestion, followed by agarose gel electrophoresis, extracted by QIAquick gel extraction (Qiagen), and cleaned up with Qiagen Minelute PCR purification Kit (Qiagen). After, 500 ng of amplified captured DNA was recombined with 2000 ng of linearized screening vectors in a total of 10 μl per reaction (each having 50 ng of captured DNA and 200 ng of screening vector) (Clontech In-Fusion HD). All the recombination reactions were pooled together and purified with Agencourt AMPureXP DNA beads (Thermo Fisher Scientific) and then eluted in 29 μl. Thirteen aliquots (20 μl each) of MegaX DH10B Electrocompetent Bacteria of (Thermo Fisher Scientific) were transformed with 2 μl of DNA each, according to the manufacturer's recommendation. After 1 h recovery at 37°C, the transformations were pooled together and transferred into 2 l of LB overnight for each specific promoter vector. An aliquot of each transformation was plated on LB AMP medium to estimate the number of cloned fragments. A total of 6–8 million clones were achieved by each library. Finally, plasmid libraries were purified using Qiagen Plasmid Plus Maxi Kit (Qiagen).

CapSTARR-seq library transfection

For each library, a total of 50×10^6 cells were transfected in triplicate (5 μg/1×10^6 cells) using the Neon Transfection System (Thermo Fisher Scientific). The P5424 cells were transfected with 1600V-20ms-1 pulse conditions. After transfection, cells were transferred to a complete growth medium and incubated for 24 h before isolation of the RNA.

CapSTARR-seq RNA isolation from transfected cells

RNA extraction was performed using the RNeasy miniprep kit (Qiagen) with the on-column DNaseI treatment. The

PolyA RNA fraction was isolated by μMACS mRNA isolation kit (Miltenyi Biotec) following the manufacturer's recommendations. PolyA RNA was treated with Ambion turbo DNase (Thermo Fisher Scientific) and then purified with RNeasy Minelute kit (Qiagen). Finally, mRNA was quantified by using Qubit RNA HS Kit (Thermo Fisher Scientific).

CapSTARR-seq reverse transcription and sequencing library preparation

cDNA first-strand synthesis was performed for non-stimulated (NS) and stimulated (PMI/I) cells with superscript III (Thermo Fisher Scientific) using a reporter-specific primer (5′-CAAACTCATCAATGTATCTTATCATG-3′) and 0.2 to 0.3 μg of polyA RNA per reaction for a total of 10 reactions. After the reverse transcription, 1 μl of RNaseH was added and incubated at 37°C for 1 h. The cDNA was then purified with QIAquick PCR purification kit and determined concentration using Qubit ssDNA Kit (Thermo Fisher Scientific). The cDNA was amplified using the KAPA Hifi Hot Start Ready Mix in a 2-step nested PCR. In the first PCR (98°C, 2 min; followed by 15 cycles of 98°C for 20 s, 65°C for 20 s, 72°C for 30 s), cDNA of 5 ng per reaction was amplified using two reporter-specific primers (fw: 5′-GGGCCAGCTGTTGGGGTG*T*C*C*A*C-3′) and rw: 5′-CTTATCATGTCTGCTCGA*A*G*C-3′), one of which spans the splice junction of the synthetic intron, in a total of 10 reactions. Purification of the PCR products was performed on gel using QIAquick gel extraction kit (Qiagen) followed by a clean-up with QIAquick mini elute PCR purification kit (Qiagen), to remove any residual contamination of plasmid or cDNA. Generating Ion Torrent libraries, purified PCR product was used as a template for the second PCR (5 ng/PCR, for a total of 10 PCR reactions; 98°C, 2 min; followed by 10 cycles of 98°C for 20 s, 65°C for 20 s, 72°C for 30 s) with KAPA Hifi Hot Start Ready Mix and Ion Torrent library amplification primer mix (Thermo Fisher Scientific, T_PCR_A: 5′-CCATCT CATCCCTGCGTGTC-3′ and P1amp: 5′-CCACTACGCCTCCGCTTTCCTCTCTAT G-3′). Generating the INPUT control, 10 reactions with 5 ng of reporter constructs (library) per reaction were amplified using the same conditions as mentioned above, except for forward primer in the first PCR (fw: 5′-GGGCCAGCTGTTGGGGTG*A*G*T*A*C-3′). The '*' indicates the phosphorothioates bond. To assess potential biases in library composition caused by electroporation, 10 reactions with 5 ng per reaction of the reporter constructs isolated from transfected cells were amplified as explained above. For sequencing the libraries, the sequencing indexes are added to the libraries by one simple PCR reaction using Multiplex oligos for Illumina (NEBNext Ultra RNA library prep). So, the non-transfected (plasmid input) and transfected libraries were sequenced on the Illumina NextSeq 500 platform (Supplementary Table S1).

CapSTARR-seq data processing

FastQ files of transfected (cDNA) and non-transfected (input) libraries were trimmed using sickle with -q 20 option and mapped to the mm9 reference genome using bowtie2 with default parameters. Sam files were converted using samtools to BAM files. BAM files of all replicates were pooled and converted into bed files using BedTools (v2.28.0) (40). In this approach, each read is considered a captured region and independent clone. First using basic bash commands, we create for each condition a unique list of clones present in cDNA and input libraries. To count the frequency of each unique clone, the length was set to 1 bp to avoid multiple counting clones, then it was count using BedTools intersect -c using the original bed files with the unique list. This led us to get the number of clones present in cDNA and input. The length of all clones was restored to 314 nt, corresponding to the average size of the captured fragments. Each DHS region was annotated with a unique ID, then each clone was annotated by the ID of the overlapping region. We extended the original DHS coordinates by considering the start of the first clone and the end of the last clone overlapping the DHS region using BedTools and python homemade scripts. These regions are considered as the extended DHS. The extended DHS regions were split using bins of 50 bp. Using a R homemade script, we computed the activity of the extended DHS regions computing the fold change of the sum of the clones overlapping the same region in the transfected condition over the non-transfected condition. The count of clones was normalized using FPKM and activities were centered. We excluded the region which has an FPKM <1 in the input. The subregion activity was computed in the same way, using the clone that overlaps the subregion created using the bins of 50 nt. We defined silencers as the extended DHS regions with $\log_2$(activity) lower than -1. In these regions, we defined the core silencer as the consecutive subregions with $\log_2$(activity) lower than -1. Then, we identified the edge silencer as the consecutive subregions with the lowest activity of the core silencer. A silencer can have only one edge of N subregion with strict equal minimal activity. All the regions were annotated with the two nearby genes using GREAT web-service (41). All results are summarized in Supplementary Table S2 and silencers are summarized in Supplementary Table S3. To visualize the CapSTARR-seq signal per individual cloned fragments or by regions, we generated a bed file with a color code proportional to the activity ranging from green for the positive activity to red for the negative activity (silencer). To assess reproducibility between replicates and conditions, we generated a correlogram using ggplot2 (https://onlinelibrary.wiley.com/doi/abs/10.1002/wics.147) and corrplot package (https://cran.r-project.org/web/packages/corrplot) with a Spearman non-parametric test.

CUT&tag and CUT&RUN

CUT&Tag of H3K27me3 (pAb C39055, Active motif) was performed by using CUT&Tag-IT Assay Kit (Active motif) following the manufacturer's protocol. CUT&RUN of H3K9me3 (pAb, C39062, Active motif) was performed by using CUTANA CUT&RUN kit (EpiCypher) following the manufacturer's protocol. In both experiments, 500 000 P5424 cells were used. For sequencing the libraries, the sequencing indexes were added to the libraries by using NEBNext Ultra II DNA library prep Kit (Illumina). The li-

braries were sequenced on the Illumina NextSeq 500 platform.

Definition of silencers

Putative silencers were defined as DHS regions displaying a CapSTARR-seq signal ($\log_2$(FC)) lower than or equal to −1 in any of the CapSTARR-seq experiments. For each library, a control group of the same number of DHS as the silencers were created by randomly selecting DHS with $\log_2$(FC) ranging between −0.1 and 0.1. A list of active enhancers was created using the SCP1 condition taking the DHS with a $\log_2$(FC) >1. The relative proportion of proximal (≤1 kb) and distal (>1 kb) silencers was computed using the distance to TSS of the genes provided by RefSeq (42).

Epigenomic analyses

ChIP-seq datasets from P5424 cells and DNaseI-seq and ChIP-seq datasets from mouse DP thymocytes were downloaded from Gene Expression Omnibus (GEO) database as detailed in Supplementary Table S1. ChIP-seq data were processed as described in (31). For average profiles and heatmaps, Wiggle files were converted to BigWiggle files using wigTo BigWig (43). Average profiles were generated with deepTools (44) using MNase-seq, DNaseI-seq and ChIP-seq signals from bigwig files around the DHS center (±5 kb for histone modifications, ±1 kb for TFs and chromatin features). For matrix clustering, H3K27ac, H3K4me3, H3K9me3, H3K27me3 histone marks were used as well as CTCF and DNaseI. Then clusters were generated using plotHeatmap with 8 k-means option in order to isolate as possible the different regions. Silencers showing strong enrichment for active marks (H3K27ac, H3K4me3), CTCF and repressive marks (H3K9me3, K3K27me3) were merged.

RNA-seq data

Public RNA-seq data from P5424 cells (36) and mouse T cell differentiation (45) were downloaded from GEO database (accession numbers GSE120655 and GSE48138, respectively). The raw RNA-seq data was processed as described in (36).

Association of silencers with variable genes

RefSeq genes in a window of ±100 kb around all silencer candidates were retrieved using BedTools window (v2.28.0). The different conditions were annotated using the mm9 file from UCSC. To retrieve the top 5% most variable genes, the variance through T cell differentiation was computed for each gene using merged replicates of the RNA-seq dataset (45). Those 5% most variable genes were considered as T cell regulated genes.

Tissue-specific genes

Gene tissue-specificity score was computed by adapting the calculation method of entropy of Shannon to data expression from GNF Gene Atlas (46). This led us to compare the distribution of expression of genes across different tissues. High entropy score indicates uniform distribution meaning that the gene is not tissue-specific. Low entropy (entropy ≤ 3) score indicates high tissue-specific genes.

Motif research analysis and clustering

The HOMER (47) software was used to perform research motif analysis with the option 'findMotifsGenome.pl input.file.bed mm9 output.file.bed –len 6,8,10,12,15 –size given –bg dhs'. We choose as background all the DHS sites in order to isolate the core silencer-specific binding sites. This method was also used within STR contained in the core silencers subregions. From the known motif output files, a list of transcription factors (TFs) was obtained. The relative matrix files were downloaded from JASPAR 2018 database (48). In order to reduce the redundancy, the RSAT matrix clustering tool (49) was used with default parameters to cluster motifs based on their sequences. The best transcription factor *P*-values provided by HOMER were kept by cluster and the expression of the individual genes in P5424 was retrieved (Supplementary Table S4). Then, the activity of silencers containing enriched transcription factor binding site (TFBS) based on the genomic track from JASPAR 2018 was compared.

Genomic and RE enrichment of candidate regions

Genomic distribution was analyzed using OLOGRAM (50), included in pygtftk (51). Regions outside DHS were excluded to compare our candidates with the DHS distribution. mm9 GTF files from Ensembl were provided to assess the genomic distribution of our candidates. We plotted with ggplot2 the $\log_{10}$(*P*-value) with a factor ±1 depending on whether it is enriched or depleted to obtain the enrichment score. Enrichment of REs was performed as described above using RepeatMasker (v4.0.8) file for mm9 (52).

RE analysis

The genomic position of STR elements within the core silencer was obtained using BedTools intersect to get overlapping regions between STR repetitive elements based on RepBase (53). The impact of the number of repeats on silencer activity was investigated using HipSTR reference file (54) in mice giving the details of simple repeat elements in the genome. Reference file was converted from mm10 to mm9 using LiftOver. Using BedTools, silencers containing simple repeats were identified and then separated by the number of repeats.

Functional analysis

Functional enrichment analyses of putative silencers identified in each condition and putative enhancers coming from the SCP1 strategy were produced using a custom pipeline to automate multi-sample queries to the GREAT web-service. In summary, putative silencers regions and putative enhancers regions were queried with rGREAT R package (https://github.com/jokergoo/rGREAT,http://great.stanford.edu/public/html/) against GO Biological

Process. Genomic association rules with genes were done using default GREAT 'Basal plus extension'. Significant terms were filtered as those having Binomial Fold Enrichment higher than 2 and both Binomial and Hypergeometric tests Benjamini–Hochberg adjusted *P*-values lower than 0.05. GOSemSim R package (55) was used to compute the Wang similarity distance between all terms. Terms with a similarity distance higher than 0.8 were grouped. Terms were further filtered for heatmap display to keep only the best five terms for each sample according to the binomial *P*-value. The color scale on heatmaps displays Binomial Benjamini–Hochberg adjusted (BH) *P*-values.

Luciferase reporter assays

Silencer candidates or control regions were amplified from mouse genomic DNA or synthesized *in vitro* and cloned downstream of the luciferase gene at the BamHI and Sal1 restriction site in the pGL3-Promoter vector (Promega) (Supplementary Table S5) and verified by Sanger sequencing. A total of 1×10^6 P5424 cells were co-transfected with 1 μg of the tested construct and 200 ng of *Renilla* vector using the Neon Transfection System (Thermo Fisher Scientific). Electroporation conditions for P5424 cells were maintained at 1600 V–20 ms^{-1} pulse and for NIH-3T3 were maintained at 1350 V–20 ms^{-2} pulse. After 24 h of transfection, luciferase activity was measured using the Dual-Luciferase Reporter Assay kit (Promega) on a TriStar LB-941 Reader. For all measurements, firefly luciferase values were first normalized to *Renilla* luciferase values (controlling for transfection efficiency and cell number). Data are represented as the fold decrease in relative luciferase signal over the pGL3-Promoter vector (pSV40). All experiments were performed in triplicates.

FACS analysis

A total of 1×10^6 P5424 cells were transfected, as described above, with the indicated vectors. Twenty-four hours post electroporation, GFP expression was assessed on a FACS Calibur (BD Biosciences).

Site-specific mutagenesis

Mutagenesis was performed using Q5 Site-Directed Mutagenesis Kit (New England Biolabs) following the manufacturer's instructions. The mutagenesis primers were designed using the NEBase Changer tool and are shown in (Supplementary Table S6). All the mutations were verified by sanger sequencing.

Hi-C processing

Raw Hi-C data from primary DP thymocytes were taken from Hu *et al.* (56) and processed with FAN-C (57), entailing iterative mapping to the mm9 genome assembly with bowtie2, filtering self-ligation events and PCR duplicates, binning the data to 10 kb bins and balancing the chromosome-wide matrices with the Knight–Ruiz method. Topologically associating domain (TAD) boundaries were identified by computing insulation scores (58) with windows of 100 kb (10 bins), normalizing to chromosome-wide averages of insulation scores, then filtering the local minima with the delta vector calculated for the three bins flanking the computed one, and with the difference of the minima and maxima of the delta vector being at least 0.7.

CRISPR/cas9 genome editing

For the CRISPR/Cas9 experiments of targeted silencer regions, two gRNAs were designed for each end of the targeted region using the CRISPRdirect tool (59). The designed gRNAs were cloned into a gRNA cloning vector (Addgene, 41824) as described previously (60). One million cells were transfected with 1 μg of each gRNA and 1 μg of hCas9 vector (Addgene, 41815) using the Neon Transfection System (Thermo Fisher Scientific) and cultured in 5 ml. After two days of transfection, the transfected cells were plated in 96-well plates at limiting dilution (0.5 cells per 100 μl per well) for clonal expansion. Individual cell clones were screened for homologous allele deletion after 10–14 days, by direct PCR using Phire Tissue Direct PCR Master Mix (Thermo Fisher Scientific) according to the manufacturer's protocol. For the detection of knockout or wild-type alleles, forward and reverse primers were designed to bracket the targeted regions. The clones were considered to have undergone homologous allele deletion if they had no wild-type band and at least one deletion band of the expected size (Supplementary Figure S5A–K). The homozygous deletions were verified by Sanger sequencing (Supplementary Figure S5L). All the gRNAs and primers are listed in Supplementary Table S6.

Gene expression analyses

Total RNA from P5424 from WT and deleted clones in non-stimulated or PMA/Ionomycin treated conditions was extracted using the RNAeasy mini kit (Qiagen). Three micrograms of RNA were treated with DNase I (Thermo Fisher Scientific) and were quickly reverse transcribed into cDNA using Superscript VILO Master Mix (Thermo Fisher Scientific). The Real-time PCR was performed using Power SYBR Master Mix (Thermo Fisher Scientific) on a Quant Studio 6 Flex instrument (Thermo Fisher Scientific). Primer sequences are listed in Supplementary Table S6. The expression of the gene was normalized to that of *Rlp32*. Relative expression was calculated by the ΔC_T method, and all the data shown are reported as fold change over the control. For each of the cell clones, from the three independent RNA/cDNA preparations, the Student's *t*-test was performed (unpaired, two-tailed, 95% confidence interval). Data are represented with s.d. For the conventional RT-PCR, one-twentieth of the synthesized cDNA was used as the template for the reaction.

RESULTS

Experimental strategies to identify silencer elements

To set up an experimental strategy to quantify silencer activity, we repurposed the CapSTARR-seq technique (31), an approach coupling capture of defined regions to the previously developed STARR-seq technique (30) (Figure 1A;

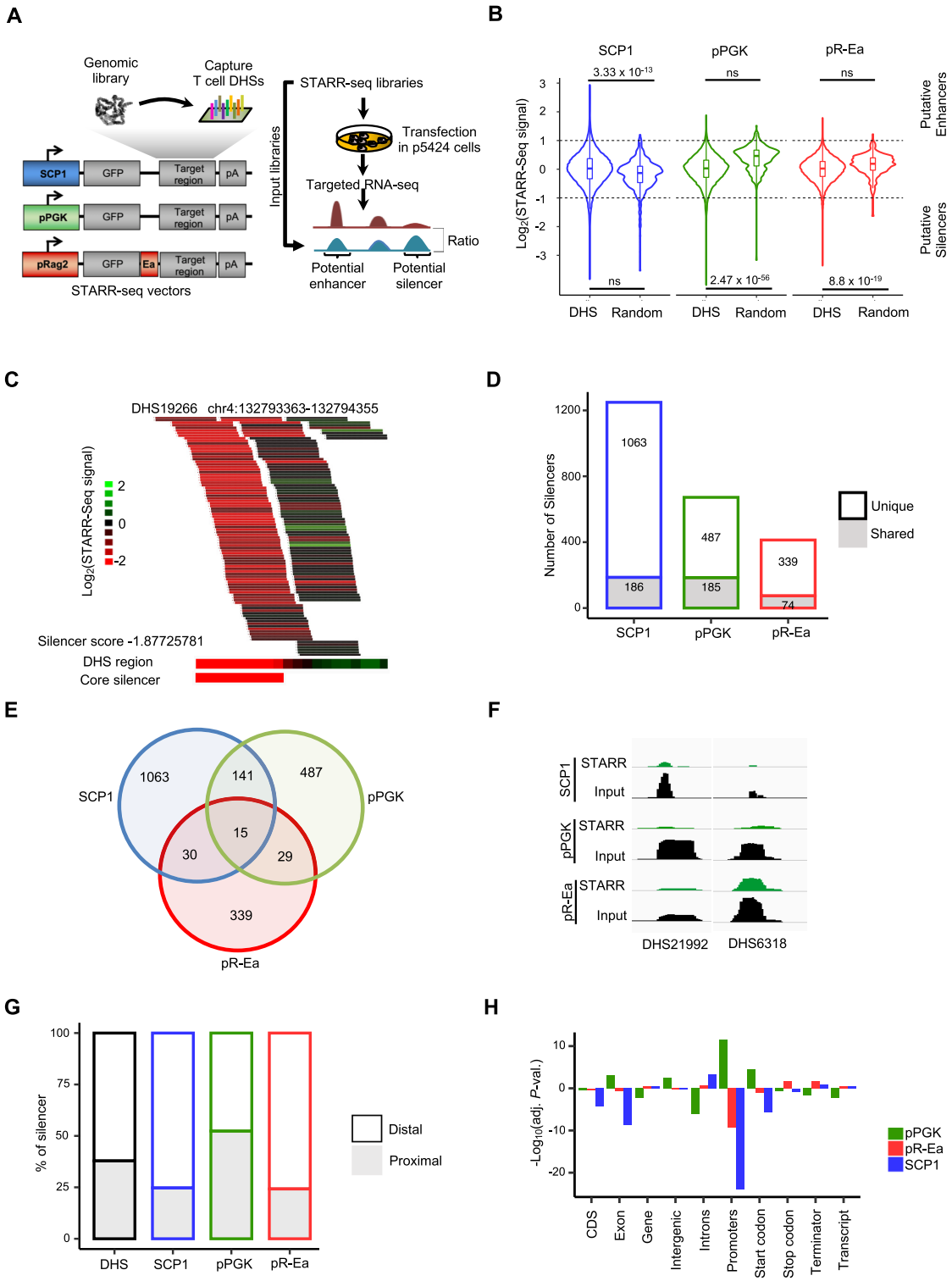

Figure 1. CapSTARR-seq for silencer assessment. (**A**) Schematic of the CapSTARR-seq strategy to assess the silencer activity in the P5424 mouse T cell line. (**B**) Distribution of CapSTARR-seq signal ($\log_2$) in the different conditions for DHS and random captured regions. The threshold for putative silencers ($\log_2(FC) \leq -1$) and enhancers ($\log_2(FC) \geq 1$) is indicated. Statistical analysis was performed using Wilcoxon test, *P*-values are displayed (ns: not significant). (**C**) The UCSC genomic track of Mouse NCBI37/mm9 around DHS19266 showing the $\log_2$ STARR-seq signal of individual clones, the captured DHS region, the core silencer and the silencer score of the region. (**D**) Bar plot showing the number of unique and shared silencer candidates identified with the three different CapSTARR-seq libraries. (**E**) Venn diagram displaying the overlap between the silencer candidates identified with the three different CapSTARR-seq libraries. (**F**) Example of a putative silencer identified with the three promoter-based CapSTARR-seq strategies. The signal for each CapSTARR-seq experiment and the corresponding Input are displayed. (**G**) Bar plot displaying the proportion of silencer candidates that are proximal (<1 kb) or distal (>1 kb) to the closer TSS. (**H**) Genomic distribution of silencer candidates compared to the whole set of DHSs. Bar plots represent the $-\log_{10}$(*P*-value) of the negative binomial test computed by OLOGRAM. Depletion is represented by negative values.

see also Materials and Methods section). We replaced the basal synthetic super core promoter 1 (SCP1) promoter present in the original STARR-seq vector with a ubiquitous strong promoter from the human *PGK* gene (pPGK) (37) and a T cell-specific promoter enhancer pair pRag2-Eα (pR-Ea) (38) (Figure 1A). Analyses of GFP expression by FACS demonstrated increased promoter activity with the pPGK and pR-Ea derived STARR-seq vectors, as compared to the original SCP1 vector (Supplementary Figure S1A, B). Although little is known about the general biochemical properties of silencers, it seems reasonable to assume that a subset of silencers may be occupied by sequence-specific transcription factors and/or lie within nucleosome-depleted genomic regions, and consequently may overlap with DNase I Hypersensitive Sites (DHS). Indeed, several known silencers have been identified as laying within DHS sites (e.g. (18,32,61)). Therefore, to isolate silencer elements genome-wide, we designed a captured library containing 28055 DHSs from mouse double-positive (DP) thymocytes, plus 437 randomly-selected non-DHS regions as negative controls.

In brief, DNA fragments of ∼400 bp were captured on a custom-designed microarray covering all the DHS and cloned by homologous recombination into the three different STARR-seq vectors (hereafter named, SCP1, pPGK and pR-Ea libraries). The STARR-seq libraries were transfected in triplicate into the mouse T-cell line P5424 and sequenced by targeted RNA-seq (Supplementary Table S2). We have used the P5424 cell line as a model of differentiating T cells in previous STARR-seq experiments (31,62). The P5424 cell line originated from early developing T cells and have a transcription signature similar to $CD4^{-}$ $CD8^{-}$ double negative (DN) thymocytes, including high expression of *Ptcra* and *Hes1* markers (35,36). However, like other DN-derived cell lines, P5424 cells also express the CD4 and CD8 surface markers, suggesting that they are blocked between the DN-to-DP transition during the β-selection process. As controls, we sequenced the libraries before transfection (hereafter named, input). The *cis*-regulatory activity was assessed by computing the $\log_2$ ratio between the targeted RNA-seq and the corresponding input signals (hereafter referred to as STARR-seq signal) after normalization and filtering (See Methods section; Figure 1B; Supplementary Table S2). A good correlation was obtained between the replicates of the same library (Spearman's correlation coefficient (ρ) ranging between 0.38 and 0.87; Supplementary Figure S1C). For subsequent analyses, the signals from the replicates of the same library were merged.

Comparison of STARR-seq signal between DHS and random regions for the different libraries showed that enhancers are significantly overrepresented in DHSs as compared to random regions in the SCP1 library (Figure 1B). In contrast, DHSs with silencer activity were significantly detected with the pPGK and pR-Ea libraries. These results suggested that using a strong promoter-based library is indeed an effective strategy to detect silencer elements. We defined the putative silencers as DHS regions with a $\log_2$ STARR-seq signal lower than or equal to −1 (Figure 1B) and found a consistent silencer activity between the two relative orientations of the fragments in the STARR-seq vector (Supplementary Figure S1D). The silencer definition resulted in a set of 1249, 672 and 413 putative silencers for the SCP1, pPGK, and pR-Ea libraries, respectively (Supplementary Table S3). Given the comparison of the STARR-seq signal between DHSs and random regions (Figure 1B), we expect that many putative silencers identified with the SCP1 library might be false positives. Figure 1C provided an example of a silencer region identified with the pPGK library showing the STARR-seq signal for the individual cloned fragments. Core silencers were identified as a region containing successive bins of 50 bp with a $\log_2$ STARR-seq signal lower than or equal to −1 (Figure 1C; Supplementary Table S2). The correlation of silencer activity between the different libraries was relatively low (Supplementary Figure S1C), indicating that the use of distinct promoters enables the identification of distinct, yet overlapping, sets of silencers, and suggesting that silencers may exhibit promoter-specific activities, as has been shown for enhancers (63). Consistently, 15%, 27.5% and 18% of silencers identified with the SCP1, pPGK and pR-Ea libraries, respectively, were also found within another library (Figure 1D, E), while only 15 putative silencers were shared between all the libraries (two examples are displayed in Figure 1F).

To assess whether there was a bias in the genomic location of silencers, we computed the specific enrichment of putative silencers obtained with each library with respect to all DHS (Figures 1G and H). While SCP1 and pR-Ea based silencers were depleted from promoter regions, the pPGK based silencers were enriched. Importantly, none of the silencer sets were enriched for terminator sequences (Figure 1H), which could represent a potential bias of the approach by artificially interfering with the quantification of the STARR-seq vector-derived transcripts.

Validation of STARR-seq identified silencers

To independently evaluate the accuracy of STARR-seq to identify silencers, we selected 24 DHS candidate silencers (13 commons in at least two libraries, 8 specifics to the pPGK and three specifics to the pR-Ea libraries), as well as 12 DHS control regions ($\log_2$ STARR-seq signal close to 0). The selected DHS were tested in a classical luciferase reporter assay in the P5424 cell line (Figure 2A). We found that 87.5% (21 out of 24) putative silencers and 33.3% (4 out of 12) control DHS regions displayed significant silencer activity in the luciferase assay. Overall, the STARR-seq identified silencers displayed a higher silencer activity in the luciferase assay as compared with the control DHS set (P-value = 0.0001; Figure 2B). Importantly, several STARR-seq-defined silencers, including DHS26112, DHS2610, DHS10824, DHS5667, DHS12366 and DHS23650, displayed strong silencer effects resulting in luciferase expression close to background levels, while this was not observed for any of the control regions. Moreover, silencer activity was independent of the orientation of the tested region with respect to the luciferase gene (Figure 2C). To further assess the silencer activity, we cloned one of the validated silencers (DHS12366) into a GFP-containing reporter vector. After transfection in P5424 cells, the vector with the DHS12366 silencer displayed reduced GFP expression as compared to the control vector (Figure 2D). Thus, consistency between the independent reporter assays indi-

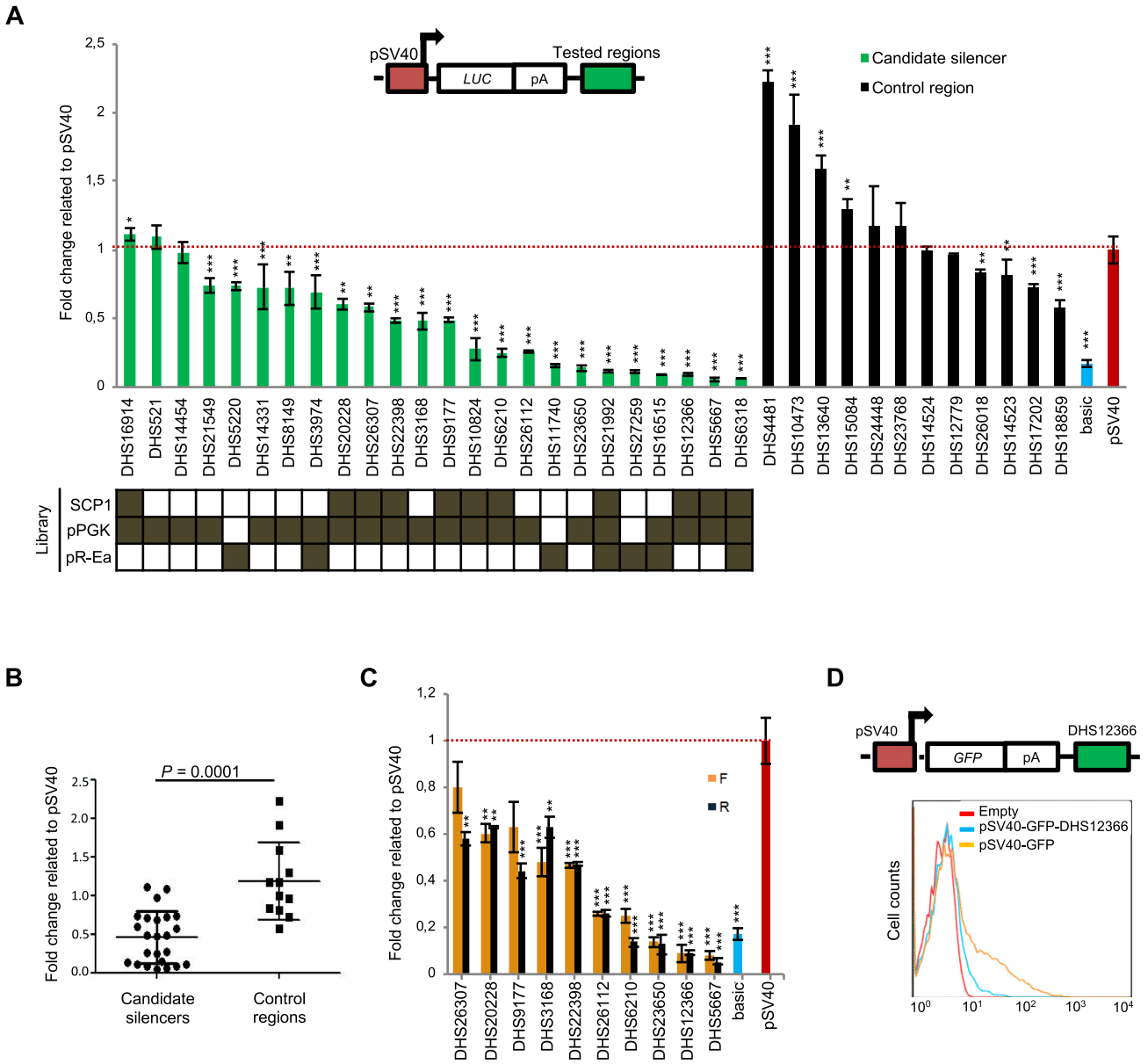

Figure 2. Validation of the CapSTARR-seq approach. (**A**) Luciferase reporter assays in P5424 cells of DHSs defined as putative silencers by CapSTARR-seq (green) or with $\log_2$ STARR-seq signal close to zero (black). The promoter-based CapSTARR-seq where the silencer was identified is indicated in the bottom panel. Data represent the normalized fold change over the pSV40 vector control. Error bars show s.d. from three independent transfections (***P-values < 0.001, **P-values < 0.01, *P-values < 0.05; two-sided Student's t-test). (**B**) Comparison of luciferase activity between silencer candidates and control regions. The two-sided Student's t-test is shown. (**C**) Assessment of orientation-dependent silencer activity for a subset of identified silencers (F: forward; R: reverse). (**D**) FACS analysis for GFP expression assessing DHS12366 silencer activity.

cates that the high-throughput assessment of silencer activity by STARR-seq is highly accurate.

Chromatin features and gene functions associated with silencers

To assess the meaningfulness of the STARR-seq identified silencers, we assessed their association with chromatin features and expression of neighbor genes as compared with corresponding control DHSs (see Methods section) and enhancers identified by the SCP1 library (Figure 1A). To explore whether silencer activity reflects the epigenetic status of the endogenous DHSs, we analyzed several chromatin features available from the P5424 cell line (Figure 3A) and primary DP thymocytes (Supplementary Figure S2), and performed Cut&TAG H3K27me3 and CUT&RUN H3K9me3 in P5424 cells (Figure 3A). Silencers were associated with a lower level of active histone marks (H3K4me3, H3K27ac and H3K4me1) (Figure 3A and Supplementary Figure S2A), as compared with control DHSs and SCP1-enhancers, but displayed a similar level of DNase I accessibility and CTCF binding (Supplementary Figure S2B). Both the repressive marks H3K27me3 and H3K9me3 were present at elevated levels

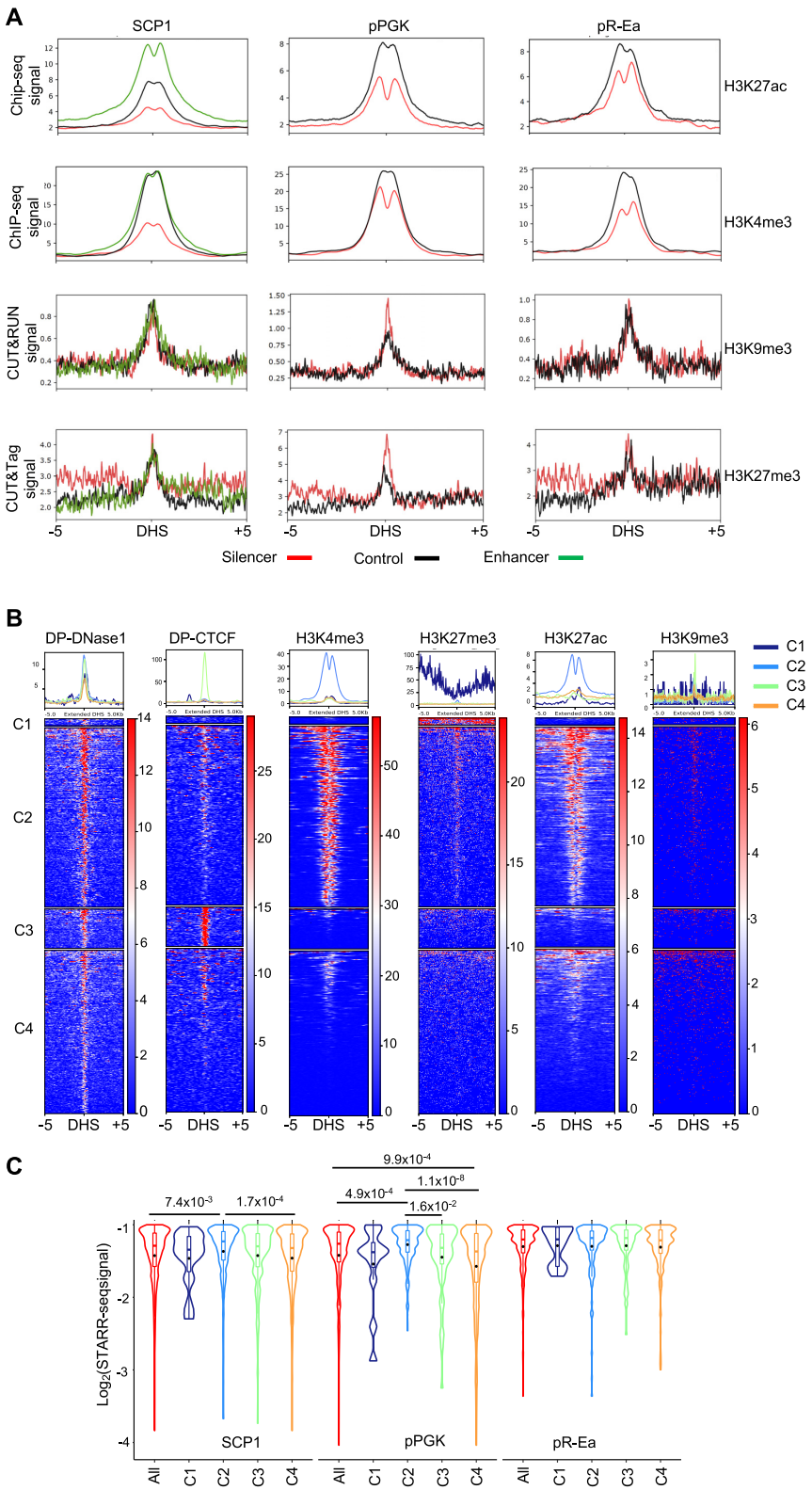

Figure 3. Chromatin features associated with silencers. (**A**) Average profiles of H3K27ac and H3K4me3 (ChIP-seq), H3K9me3 (CUT&RUN), H3K27me3 (CUT&Tag) signal from P5424 cells centered on putative silencers (red), control regions (black) and putative enhancers (green) with a window of ±5kb. (**B**) Average profiles and heatmaps of pPGK silencers clustered in the function of the signal of different histone marks and CTCF in four distinct groups. (**C**) Comparison of pPGK silencer activity within each cluster. Statistical analysis was performed using the Wilcoxon test, significant *P*-values are displayed.

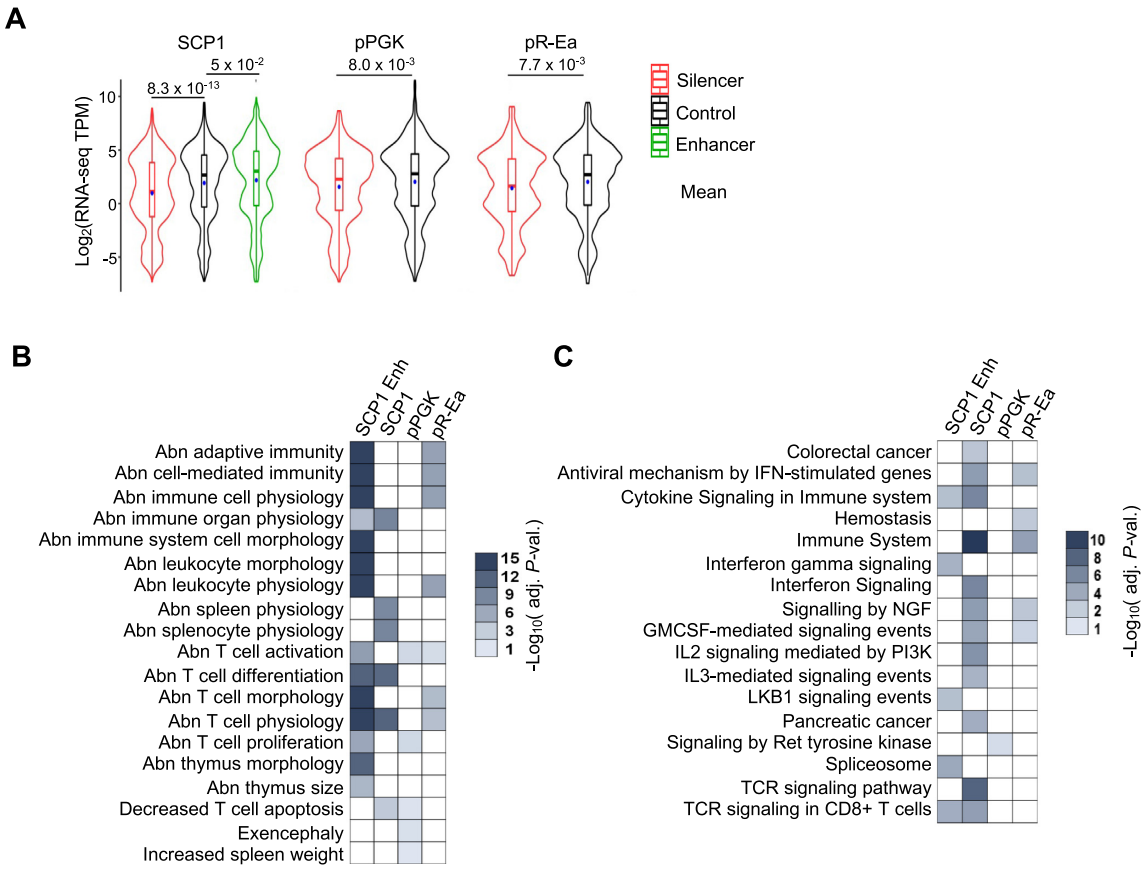

Figure 4. Gene expression and GO term analysis. (**A**) Violin plot comparing the expression in P5424 cells of genes associated with putative silencers (red), control regions (black) and putative enhancers (green). The mean values are indicated by a dot. Statistical analysis was performed using the Wilcoxon test; *P*-values are displayed. (B, C) Heatmap of top five GO terms analysis for mouse phenotype (**B**) and MSigDB pathway (**C**) enriched in genes associated with silencers or SCP1 enhancers.

at pPGK-silencers compared to control DHSs, while SCP1- and pR-Ea-silencers were not enriched in H3K27me3 nor H3K9me3 (Figure 3A). Silencers have been previously associated with H4K20me3 (23). Using H4K20me3 ChIP-seq data from mouse hematopoietic precursors (64), we observed that pPGK-silencers, and to a less extent pR-Ea-silencers, were also enriched in this histone modification (Supplementary Figure S2C).

To gain insight into the potential combination of chromatin features associated with silencers, we clustered the signal of histone modifications and CTCF around the silencer regions (Figure 3B, Supplementary Figure S3A, B). For each set of silencers, we obtained four clustering with different chromatin features, including repressed chromatin enriched in H3K27me3 (cluster C1), relative open chromatin enriched in H3K27ac and H3K4me3 (cluster C2), association with CTCF (cluster C3), and silent chromatin with sparse enrichment in CTCF and H3K9me3 (cluster C4). Characteristic examples of each cluster are presented in Supplementary Figure S4. Note that in the case of pPGK-silencers the cluster associated with open chromatin features also displayed slight enrichment in H3K27me3 and H3K9me3 at the DHSs (Figure 3B). By comparing the silencer activity within each cluster, we observed that pPGK- and SCP1-silencers present in DHS with active marks (H3K27ac and/or H3K4me3, cluster C2) have significantly less silencer activity than those found in the other three clusters (Figure 3C). Overall, these results suggested that silencer elements are found in a relatively open chromatin configuration and might be associated with different types of epigenetic signatures, including either active or repressive histone modifications.

We next used the GREAT tool (41) to associate each DHS to their neighbor genes and assess gene expression using available RNA-seq data from P5424 cells (36) (Figure 4A). Silencers identified by the three different libraries were significantly associated with genes expressed at lower levels than those associated with control DHS. As expected, enhancer regions identified by the SCP1 library were associated with genes expressed at a higher level. Functional enrichment analysis showed that the identified silencers were associated with genes involved in immune and T cell phenotypes and functions, but generally different from those associated with SCP1-enhancers (Figure 4B, C), suggesting that at least a subset of silencers might be involved in the repression of T cell associated genes.

In vivo assessment of silencer activity

To more generally assess whether silencers identified by the different STARR-seq vectors readily display silencer

function at their endogenous loci, we performed CRISPR-mediated deletions of 11 silencer candidates, including four identified by all three STARR-seq libraries (Figure 5A). For each silencer candidate, we generated at least two clones and assessed the expression of neighboring genes contained within the same TAD (Figure 5B and Supplementary Figure S6), or the two adjacent TADs when the silencer was located near a TAD border (e.g. DHS12366). Deletion of seven out of ten silencer candidates resulted in strong upregulation (>2-fold) of one or more genes within the same TAD, these included the four silencer candidates identified with the three STARR-seq libraries (DHS19456, DHS9272, DHS13190, DHS12366). For instance, the shared DHS19456 silencer is embedded within a large H3K27me3 domain and its deletion resulted in consistent and significant up-regulation of three out of four genes contained within the TAD. However, while the silencers shared between the three libraries displayed exclusive silencer activity, the deletion of the library-specific silencers in a more complex output. In fact, from the six library-specific silencers tested, the deletion of five resulted in both up- and down- regulation of neighboring genes (DHS11740, DHS8661, DHS9575, DHS10350 and DHS15125), while one resulted only in down-regulation (DHS17348) (Supplementary Figure S6). Note that deleted silencers belonged to the C3 and C4 clusters only as we were not able to generate homozygous deletions of silencers belonging to the C1 or C2 clusters. These results show that STARR-seq predicted silencers generally work as *bona fide* silencers at the endogenous locus, although in some cases both activating and inhibitory effects can be observed.

Transcription factors associated with silencers

To assess whether putative silencers were enriched for TFBS we performed motif enrichment analyses using the HOMER tool (47) on the three sets of identified core silencers, as well as, the set of active enhancers (Figure 6A; Supplementary Figure S7). Enriched TFBS were clustered in TF families and the expression in P5424 of associated genes was used to prioritize potential repressor factors (Supplementary Table S4A). Active SCP1-defined enhancers were enriched in binding sites of TFs involved in T cell differentiation, including MYB, TCF, RUNX, ROR and ETS/NFATC, consistent with previous results (31). In contrast, all three sets of silencers were enriched in TFBS bound by known transcriptional repressors, including CTCF, HOX TF family, SMAD and ZNF263. Strikingly, the pPGK- and SCP1-based silencers, but not pREa silencers, were strongly enriched in binding sites for the RE1-Silencing Transcription factor (REST), a major transcriptional repressor involved in the repression of neural genes in non-neuronal cells (65–67). To better evaluate the impact of TFBS on silencer activity, we plotted the significance of the enrichment of each TFBS against the mean activity of the core silencers (Figure 6B). We found that several TFBS were associated with strong silencer activity, including REST, SMAD3/4, MAFK and the HOX TF family. The strongest silencer activity was associated with the presence of REST or MAFK motifs in silencers found with either SCP1 or pPGK vectors, or the presence of SMAD3 in SCP1-detected silencers. Finally, we assessed the relevance of REST, SMAD and MAFK binding sites by mutating these sites in silencer candidates and assessing the silencer activity by the luciferase assay (Figures 6C, E). Mutation of either of the three tested TFBS significantly de-repressed the luciferase expression, indicating that these two TFBSs are indeed important contributors to the silencer activity.

Dynamic silencer activity mediated by TCR signaling

To explore whether silencer activity can be regulated by T cell stimulation, we performed STARR-seq experiments using the pPGK library in P5424 cells treated with PMA and Ionomycin (PMA/I), previously shown to partially mimic TCR signaling and consecutive T-cell differentiation (36) (Supplementary Figure S8A; Supplementary Tables S1–S3). We observed that a majority of silencers active in PMA/I-treated cells were stimulation specific (59%) (Supplementary Figure S8B) and were generally associated with genes downregulated after T cell stimulation (Supplementary Figure S8C). Luciferase reporter assays for three induced silencers validated their stimulation-dependent activity (Supplementary Figure S8D). Analysis of motif enrichment did not reveal any TFBS specifically enriched in PMA/I-dependent silencers (Supplementary Figure S8E). However, we observed that REST-containing motifs were exclusively enriched in constitutive silencers, suggesting that binding of REST transcriptional repressor provides general silencer activity.

REST binding site-containing silencers contribute to the repression of non-T cell genes

To gain insight into the contribution of REST to silencer activity, we determined the set of identified silencers harboring a REST binding site (Figure 6F). Only ∼5% of identified silencers with SCP1 and pPGK libraries contained REST binding sites, while the proportion significantly increased to 15.4% when considering the silencers shared between SCP1 and pPGK libraries (24 out of 156). Strikingly, silencers containing REST-binding sites were associated with significantly lower levels of expression of neighboring genes (Figure 6G). Moreover, REST binding site-containing silencers were associated with a higher proportion of tissue-specific genes not expressed in T cell precursors (Figure 6H; chi-square test, P-value = 0.001 and 0.04 for the pPGK and SCP1 libraries, respectively), and also with a reduced proportion of T-cell regulated genes compared to non-REST binding site-containing silencers (Figure 6I; i.e. genes whose expression is highly variable across T cell differentiation; see Methods section).

Among the experimentally deleted silencers (Figure 5), the DHS12366 contains a REST-binding site and displayed REST-dependent silencer activity by luciferase assay (Figure 6C). As shown in Figure 5B, deletion of the DHS12366 silencer resulted in significant up-regulation of seven genes, including two genes consistently displaying more than 2-fold increased expression (*Plin4* and *Arrdc5*). Interestingly, several of the deregulated genes were expressed in a tissue-specific manner in non-T-cell tissues (Figure 5C). Consistent with a ubiquitous activity, the DHS12366 silencer dis-

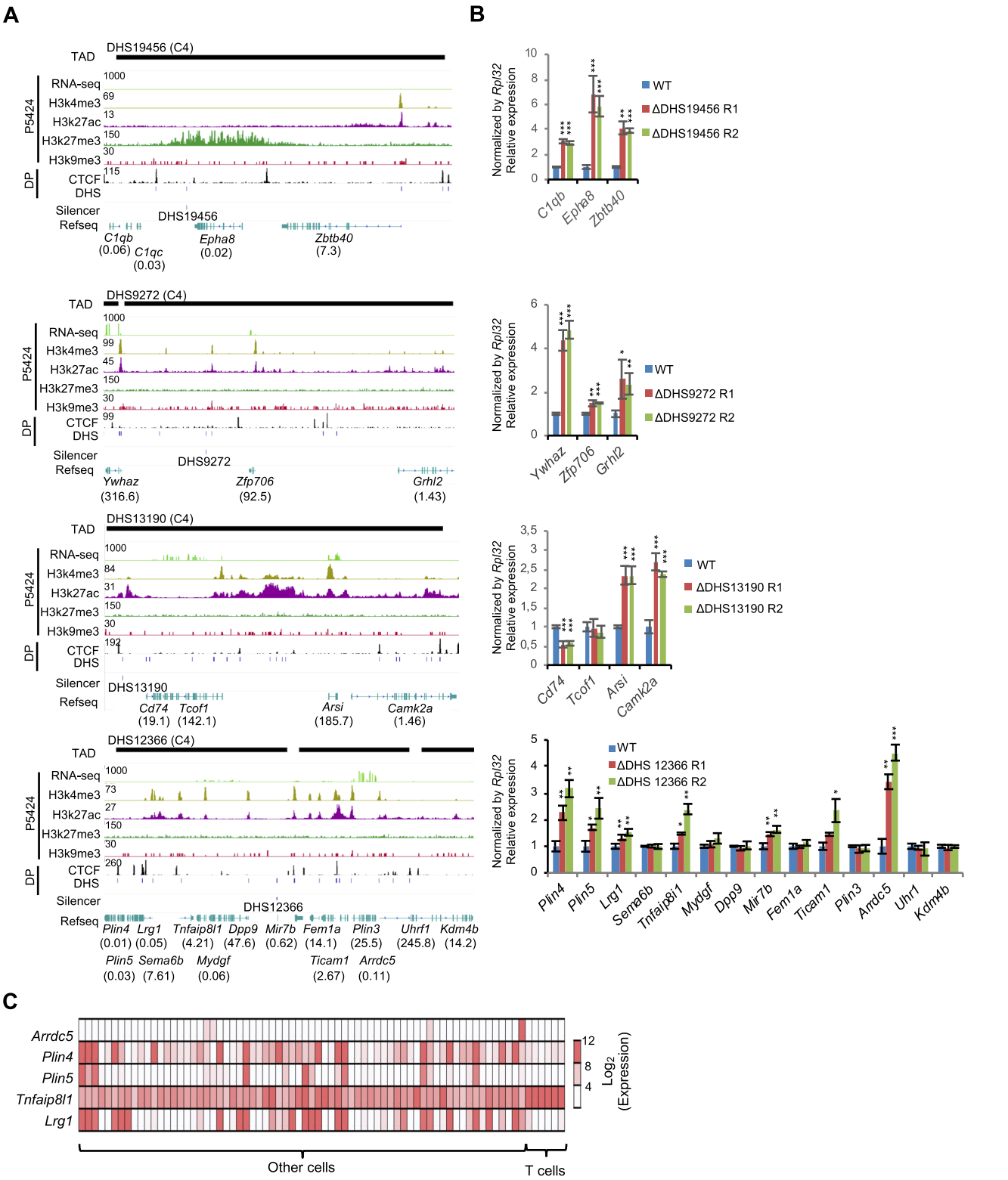

Figure 5. Functional validation of silencers by CRISPR–Cas9 genomic editing. (**A**) Genomic tracks displaying the indicated RNA-seq and chromatin signals as well as DHS and TADs surrounding the silencer regions. The relative expression values based on P5424 RNA-seq (TPM) are indicated in bracket. (**B**) Gene expression analysis of genes around the silencer region in wild-type (WT) and deleted P5424 clones. The expression of the genes was normalized to *Rlp32* and with respect to the WT value. Error bars, s.d.: ***P-values < 0.001, **P-values < 0.01, *P-values < 0.05, two-sided Student's t-test. (**C**) Heatmap displaying the relative gene expression of genes around DHS12366 locus in T cell populations and other cell types.

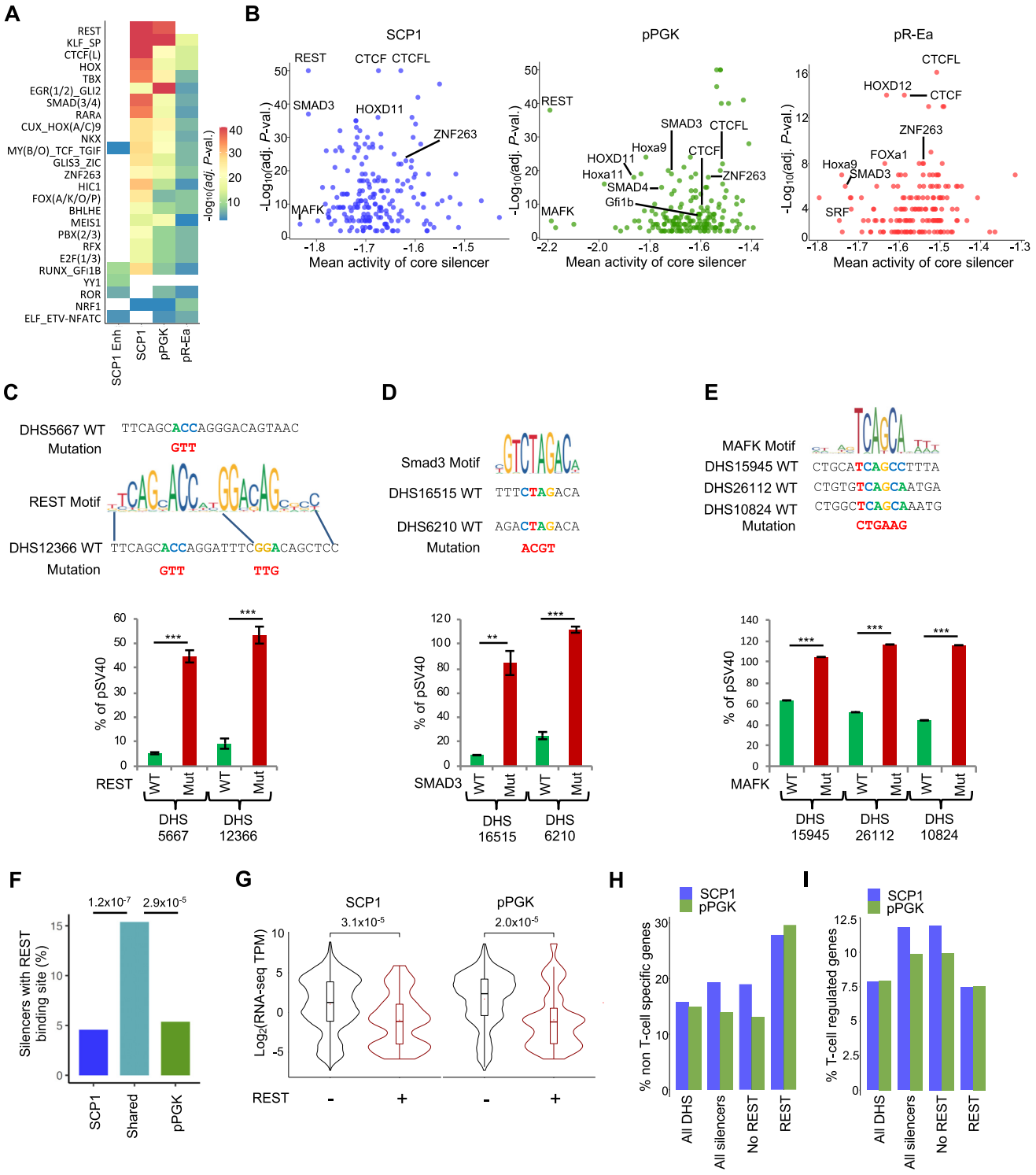

Figure 6. TFBS associated with silencers and site-directed mutagenesis. (**A**) Heatmap displaying the enrichment score of the top 10 clustered TF motifs enriched in each of the silencer sets, as well as, in the SCP1 enhancers. (**B**) Dot plots displaying the mean activity of silencers carrying a given TFBS against the enrichment score for the same TFBS. Only significantly enriched TFBS are displayed. TFs of interest are highlighted. (C–E) Validation of REST (**C**), SMAD3 (**D**) and MAFK (**E**) binding sites impact on silencer activity. The left panels display the mutated nucleotides. The right panels display the luciferase reporter assay in wild-type and mutated silencers (***P-values < 0.001, **P-values < 0.01, *P-values < 0.05; two-sided Student's t-test). (**F**) Bar plots showing the proportions of silencers harboring REST binding sites among the SCP1, pPGK or shared between the two libraries. (**G**) Expression level of genes associated with silencers containing or not REST binding sites. Significance was assessed by the Wilcoxon test. (**H, I**) The proportion of genes associated with all DHS, all silencers or silencers containing or not REST binding sites that are tissue-specifics excluding T-cells (C) or T-cell regulated genes (I).

played silencer activity also in the fibroblast cell line NIH-3T3 (Supplementary Figure S9C). Overall, our results support the idea that silencers containing REST-binding sites repress tissue-specific genes in other unrelated tissues.

The DHS23650 silencer regulates two genes involved in T cell function

Next, we sought to identify silencers involved in normal T cell function. To this aim, we searched for putative silencers that might control the expression of genes regulated across T cell development and differentiation. We reasoned that lymphoid genes regulated by silencers might have a high expression variance across T cell populations. Thus, we isolated the top 5% of highly variable genes based on available RNA-seq from different stages of thymic and peripheral T cell differentiation (45) (Figure 7A). We then retrieved the top variable genes located in a window of 100 kb around any STARR-seq-defined silencer (Figure 7B). We obtained 516 genes associated with 615 silencers. Of these, we identified the DHS23650 silencer belonging to the C4 cluster and associated with *Hcst* (Hematopoietic cell signal transducer, also known as DAP10) and *Nfkbid* (Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, delta), two genes involved in T cell differentiation and activation. The *Hcst* gene encodes for a transmembrane signaling adaptor which forms part of the immune recognition receptor complex (68,69). This receptor complex has a role in cell survival and proliferation by activation of NK and T cell responses. The *Nfkbid* gene encodes for a member of the atypical inhibitors of NF-κB TF and is particularly involved in the regulation of T cell activation and development of regulatory T cells (70,71). Strikingly, the expression of both genes is anticorrelated through T cell differentiation. In particular, *Hcst* is induced during T helper (Th) maturation, while *Nfkbid* is repressed (Figure 7C).

Hcst and *Nfkbid* were both located within the TAD containing the DHS23650 silencer (Figure 7D), suggesting that they could be a direct target of this silencer. Luciferase reporter assay demonstrated a strong silencer activity for the DHS23650 silencer in both orientations (Figure 2C). To explore the *in vivo* function of the DHS23650 silencer, we deleted this element in the P5424 cell line using CRISPR/Cas9 genome editing (Figure 7E; Supplementary Figure S5E). We analyzed the expression of all genes contained in the same TAD as the DHS23650 silencer (Figure 7E; only genes with detectable expression are shown). Only the *Hcst* gene appeared to be significantly up-regulated in the two DHS23650-deleted clones. Interestingly, we observed that *Nfkbid* was specifically upregulated by the PMA/ionomycin stimulation (Figure 7D), in agreement with its induction between the $CD4^{-}CD8^{-}$ double negative (DN) and DP stages (Figure 7C). Analysis of expression in wild-type and mutated P5424 cells stimulated by PMA/ionomycin revealed a role of DHS23650 silencer in limiting *Nfkbid* induction (Figure 7F), consistent with the conserved silencer activity of DHS23650 after PMA/ionomycin stimulation (Supplementary Figure S9A). Therefore, the DHS23650 silencer regulates both *Hcst* and *Nfkbid* genes in different stimulatory contexts. However, the effect of DHS23650 on *Hcst* and *Nfkbid* expression appeared to be mild (<2-fold) in comparison with the expression changes that these genes display during T cell differentiation (Figure 7C). This might be due to a limitation of our cell line model or indicates that DHS23650 modulates the transcription level of the target genes rather than being responsible for their complete repression.

As both *Hcst* and *Nfkbid* are important regulators of T cell differentiation and activation we analyzed the expression of several T cell markers previously validated in the P5424 cell line (Supplementary Figure S9B). Interestingly, the regulation of *Lef1*, *Ptcra* and *Bcl2* after PMA/ionomycin stimulation of P5424 cells is significantly altered in the two DHS23650 deleted clones, suggesting that the DHS23650 silencer might be required for normal T cell differentiation. Consistent with the regulation of T-cell specific genes, the silencer activity of DHS23650 was not conserved in the fibroblast cell line NIH-3T3, in contrast to the REST-containing DHS12366 silencer (Supplementary Figure S9C). Overall, these results show that our STARR-seq derived approach is able to identify a silencer element involved in the control of genes whose expression is modulated during T cell differentiation.

Repetitive elements (REs) associated with silencers

With the aim of identifying regulators associated with DHS23650 silencer activity, we looked for potential repressor candidates based on Jaspar TFBS (Supplementary Figure S7A). We identified GFI1 (Growth Factor Independent 1 Transcriptional Repressor) binding site as a potentially relevant repressor element, as this motif was enriched in silencer elements (Figure 6A, B). GFI1 is a nuclear zinc finger protein that functions as a transcriptional repressor and is essential for hematopoiesis and involved in Th2 differentiation pathway and T-cell receptor signaling (72–75). However, mutation of the GFI1 binding site did not affect the silencer activity of DHS23650 (Supplementary Figure S9D). We also observed the presence of a 16 stretches of short tandem repeats (STRs) of the 'AGGGC' unit in a head-to-tail configuration on the 5′ side of DHS23650 (Figure 8A). STRs are short sequences of DNA, normally of length 2–5 base pairs, that are generally repeated 5–50 times. It has been shown that STRs could regulate gene expression by diverse mechanisms, including the recruitment of transcriptional repressors (34,76). To experimentally validate the impact of the 'AGGGC' repeat on silencer activity, we generated a series of DHS23650 mutations and tested the silencer activity by luciferase reporter assay (Figure 8B). Deletion of the 'AGGGC' tandem repeat resulted in a significant reduction of DHS23650 silencer activity while the 'AGGGC' tandem repeat alone displayed similar silencer activity as the full-length DHS23650. To assess whether the number of 'AGGGC' repeats is important for the silencer activity of DHS23650 we mutated four, eight or all of the 'AGGGC' repeats (Figure 8B). Strikingly, increased mutations of 'AGGGC' repeats resulted in the progressive decrease of the silencer activity. Thus, 'AGGGC' tandem repeats appear to play an essential role in the silencer activity of DHS23650.

The above results raised the question of whether REs could be involved in silencer activity. Large portions of

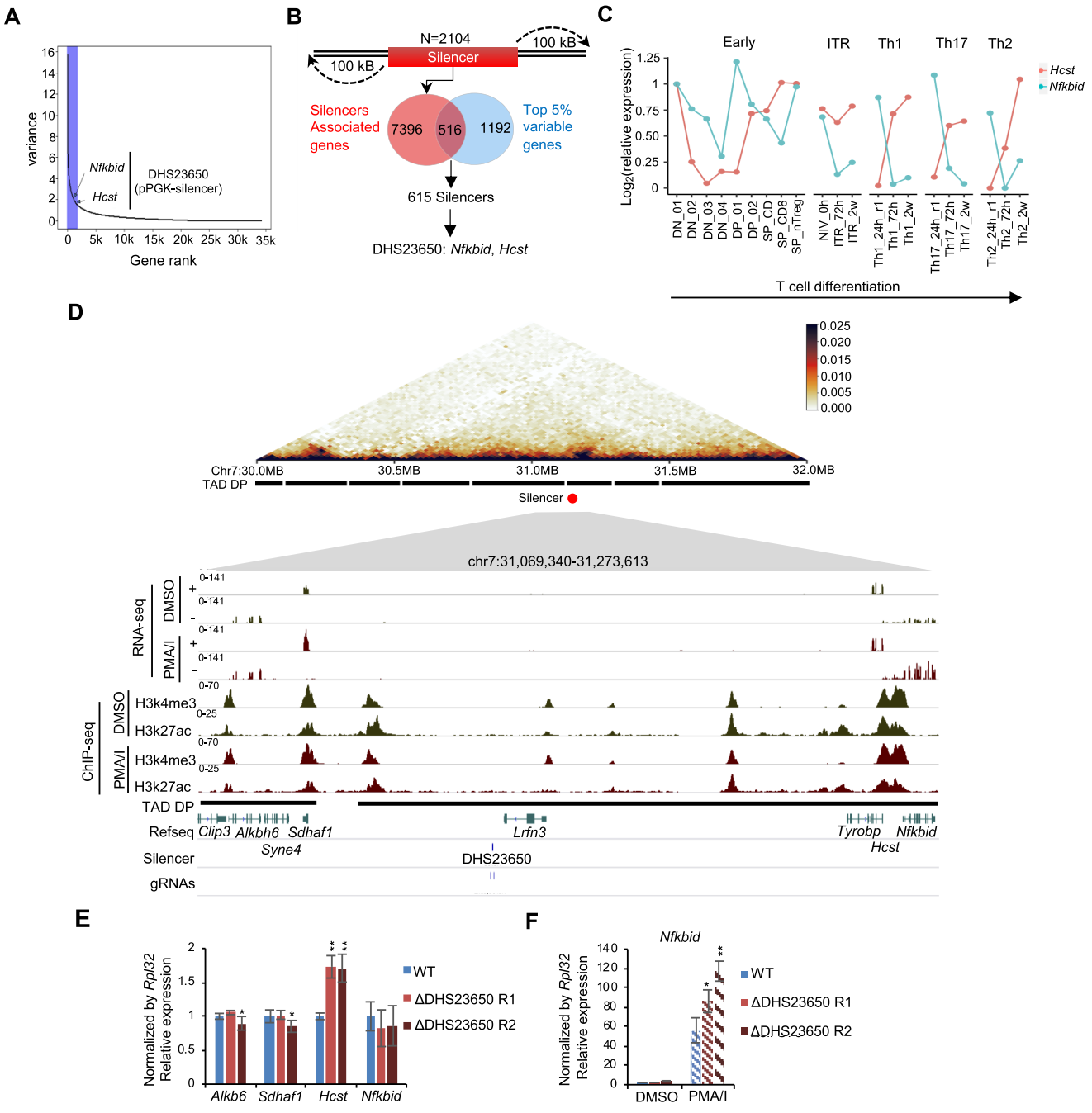

Figure 7. Functional validation of DHS23650 silencer by CRISPR/Cas9 system. (**A**) Genes ranked in the function of their expression variance across T-cell differentiation. The top 5% of variable genes are highlighted in blue. The *Nfkbid* and *Hcst* genes associated with DHS23650 silencer are also shown. (**B**) All putative silencers were associated with genes located in a window of 100 kb upstream and downstream. Of these, 516 genes (corresponding to 615 silencers) were part of the top 5% variable genes during T-cell differentiation. (**C**) Relative expression of the *Nfkbid* and *Hcst* genes during T-cell differentiation. (**D**) Top panel: Hi-C data and TADs in DP thymocytes surrounding the DHS12366 silencer. Bottom panel: genomic tracks displaying the indicated RNA-seq and ChIP-seq signals in P5424 cells stimulated or not with PMA and Ionomycin (36). (**E**) Gene expression analysis of genes around the DHS23650 locus in wild-type, ΔDHS23650 R1 and ΔDHS23650 R2 P5424 clones. (**F**) Gene expression analysis of *Nfkbid* in wild-type, ΔDHS23650 R1 and ΔDHS23650 R2 P5424 clones stimulated or not with PMA and Ionomycin. The expression of the genes was normalized to Rlp32 and with respect to the WT value as 1. Error bars, s.d.: ***P-values < 0.001, **P-values < 0.01, *P-values < 0.05 two-sided Student's t-test.

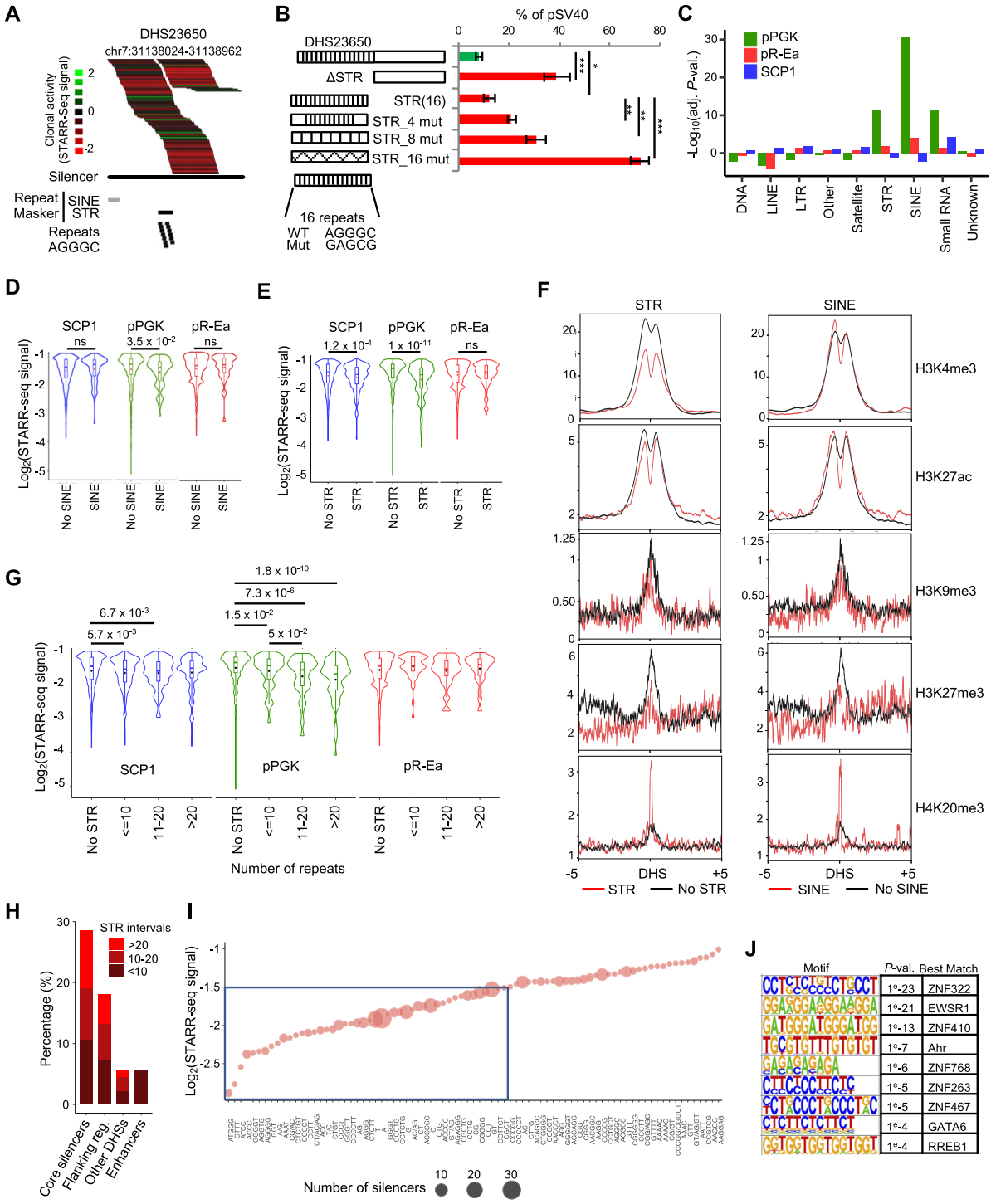

Figure 8. Analysis and validation of simple tandem repeats. (**A**) UCSC genomic track of mouse NCBI37/mm9 around the DHS23650 silencer displaying the individual clonal activity, the DHS region (silencer), the core silencer, the RepeatMasker track and the repeat units. (**B**) Luciferase reporter assay of wild-type and mutated DHS23650 silencer. The impact of STR repeats on silencer activity was assessed by either deleting the STR region or by mutating the indicated number of STR binding sites. Data represent the normalized fold change over the pSV40 vector. Error bars show s.d. from three independent transfections (***P-values < 0.001, **P-values < 0.01, *P-values < 0.05; two-sided Student's t-test). (**C**) Enrichment of repetitive elements at putative silencers. Bar plots represent the $-\log_{10}(P\text{-value})$ of the negative binomial test computed by OLOGRAM. (**D**, **E**) Activity of silencers found with the pPGK library associated with the presence of SINE (D) or STR (E) repetitive elements. Significance was assessed by a Wilcoxon test. (**F**) Average profiles of silencers with or without STR (left panels) or SINE (right panels) with a window of ±5 kb showing the enrichment of different histone marks in P5424 (H3K4me3, H3K27ac, H3K9me3, H3K27me3) or hematopoietic lineages (H4k20me3). (**G**) STARR-seq signal of silencers in the function of the number of repeat units within the STRs. Significance was assessed by a Wilcoxon test. (**H**) Comparison of the proportion of regions from pPGK with different lengths of STRs. Control enhancers are from the SCP1 library. (**I**) Activity of pPGK silencers ranked by the pattern of the STR. (**J**) Motif enrichment and associated TFs found in the subset of pPGK silencers containing STR with a $\log_2(\text{signal}) \leq -1.5$ (square in the panel I).

mammalian genomes are derived from REs, which are linked to TF binding (77–79). RE elements have been associated with both enhancer (77,80) and repressive activities (81). To more generally assess STARR-seq silencers for the occurrence of REs, we used the RepeatMasker annotation (82). The number of RE-derived sequences in silencer regions was compared to the number detected in all DHS regions (Figure 8C). pPGK-based silencers were highly enriched in SINE, STR and small RNAs. Moreover, the presence of SINEs or STRs within the silencer was associated with strong silencer activity only in those identified with the pPGK library (Figure 8D, E). We next assessed whether silencers containing SINE or STR REs were associated with a specific epigenetic signature (Figure 8F). We found that RE-containing silencers were depleted of H3K27me3, and to less extent of H3K9me3, repressive marks but were enriched in H4K20me3, a repressive histone modification previously found to silence repetitive DNA and transposons (83). These results indicate that certain families of REs might contribute to silencer function.

We suit to better assess the contribution of STRs to the activity of pPGK silencers. Core silencers were significantly associated with a higher percentage of STRs as compared with other DHS regions or the regions flanking the core silencers. We next assessed whether the length of the STRs could impact silencer activity. Reminiscent of the impact of 'AGGGC' repeat length on silencer activity of the DHS23650 silencer, we observed that the silencer activity of pPGK-based silencers significantly increased with the length of tandem repeats (Figure 8G) and was present in higher proportion in the core silencer regions (Figure 8H). Noticeable, enhancers identified by the SCP1 library were not associated with long STRs (Figure 8H) and when considering all DHS, the length of the tandem repeats was associated with significantly lower STARR-seq signal in all three libraries (Supplementary Figure S10), supporting the finding that STRs are specifically associated with silencer activity. The STRs are a family of highly heterogenous repeats (34,84), we, therefore, analyzed the contribution of commonly found repeats on the silencer activity (Figure 8I). Several STR types were associated with a strong silencer activity, with the highest activity observed with the 'ATGGG' repeat. Considering the STRs associated with a strong silencer activity ($\log_2$ FC < -1.5), we found a preferential enrichment for several members of the Kruppel-associated box (KRAB)-associated zinc finger proteins (KZFP) family of transcriptional repressors (Figure 8J), thus suggesting a potential role of KZFP on the activity of STR-containing silencers. Therefore, the STRs, by serving as a platform for the recruitment of transcriptional repressors, might play an important role in the silencer activity of *cis*-regulatory elements in T cells.

DISCUSSION

Despite the widely-held belief that silencers represent critical general regulators of gene expression, their genome-wide distribution, mechanisms of action and involvement in disease are largely unknown. A breakthrough in the analysis of distal *cis*-regulatory elements was provided by the development of high-throughput reporter assays to assess promoters (85,86), enhancers (28) or insulator activities (87). Such a strategy for the identification of silencer elements has been missing. Here, we repurposed the widely used STARR-seq approach to identify silencers systematically and efficiently. Although STARR-seq using an SCP1 basic promoter have been previously used to identify silencers (26), in the present study we evaluated and compared the use of different type of promoters and demonstrated that the use of a strong ubiquitous promoter (i.e. pPGK) allow more efficiently to assess silencer activity.

An operational assay for the silencer function of a genetic element is to isolate it and measure its ability to repress promoter activity in a given cell type (10,32). To screen for silencers in a manner optimizing throughput and functional information, we adapted a high throughput reporter assay based on the CapSTARR-seq strategy developed previously (31). We identified DNA fragments capable of negatively regulating their transcription in a minimal, episomal context in transfected cells, allowing for genome-wide screening of putative silencers from several tens of thousands of genomic regions. As an initial assessment of our approach, we analyzed a set of 28055 DHS from primary developing mouse thymocytes in a T cell line previously used in STARR-seq (31). We tested STARR-seq vectors containing either a basic (SCP1, the original STARR-seq promoter), strong (pPGK) or T-cell specific (pR-Ea) promoter. We extensively assessed the accuracy of STARR-seq to quantify silencer activity using an alternative luciferase reporter assay and CRISPR/Cas9 deletion of endogenous loci, and demonstrated its robustness to identify *bona fide* silencers. In fact, out of 11 predicted silencers assessed by CRISPR/Cas9 deletion, five resulted in consistent silencer activity, while six displayed both up- and down-regulation of neighbor genes, thus suggesting that some DHS might have dual effect as silencers and enhancers, being able to activate or repress specific neighboring genes. These results are consistent with the fact that many of the tested silencers are found within an open and active chromatin. Whether the observed dual effect depends on intrinsic features of the DHS elements or results from alteration of the 3D chromatin structures or disruption of TAD borders will require further investigation.

Overall, the basic SCP1 library allowed identifying both enhancers and silencers, consistent with previous results (26), although no significant enrichment for silencer elements was observed. In contrast, the libraries harboring the ubiquitous pPGK promoter or the T-cell specific promoter-enhancer pair (pR-Ea) were significantly enriched in DHS harboring silencer activity, thus suggesting that STARR-seq vectors with strong promoters perform well in the identification of silencer elements. We found a relatively low overlap between the silencers identified with the three libraries, suggesting that silencers may exhibit promoter-specific activities, as has been shown for enhancers (63). However, we found that silencers identify with the three libraries displayed consistent silencer activity at the endogenous locus as assessed by CRISPR/Cas9 deletion, while deletion of library-specific silencers resulted in a more complex output, including opposite activating and inhibitory effects. We also observed that silencers identified by the SCP1 and pPGK libraries shared similar characteristics,

such as enrichment for the ubiquitous REST repressor, in contrast to the pR-Ea library. However, silencers identified by the pPGK library were clearly enriched in repressive histone marks, including H3K27me3, H3KK9me3 and H4K20me3, and associated with the length of STRs, while the SCP1 and p-R-Ea-silencers were not. This suggests that the pPGK library might be enriched in typical silencers and that the use of the pPGK reporter might be a suitable vector for future silencer studies. The less efficient performance of the pR-Ea vector to identify silencers could be explained by the close proximity of the strong Ea enhancer to the tested regions, which might interfere with the silencer function. Another striking feature observed with the pPGK library was the specific enrichment of silencers overlapping gene promoters. This is reminiscent of a recent study showing that 'poised' gene promoters exhibit a silencer-like function to repress the expression of distal genes via promoter-promoter interactions (88). Together with the recent findings that many promoters can act as distal enhancers of other genes, also named Epromoters (89–91), these observations support a unifying model whereby single DNA sequences can encode different types of regulatory functions, including being a promoter for immediate genes, or an enhancer or silencer for neighboring or distal genes via linear chromatin proximity or long-range chromatin interactions (88,92). Overall, we propose that modified STARR-seq vectors, replacing the basic SCP1 promoter with a constitutive strong promoter (such as pPGK), provide an effective strategy to discover and characterize silencers genome-wide.

Several previous works have predicted silencers using indirect approaches based on 3D interactions and epigenetic signatures. A study using Capture Hi-C (CHi-C) technology suggested that transcriptionally inactive genes interacting with previously uncharacterized elements marked by repressive features may act as long-range silencers (25). They further showed that a genomic region located 1.2 Mb from the *BCL6* promoter and associated with Polycomb repressive complex 2 (PRC2) displayed silencer activity when tested in a reporter assay. Similar approaches identified PRC2-bound silencers playing an important role in mouse development (24) or tumor growth (93). Another strategy to identify candidate silencer elements has been to correlate cross-tissue epigenetic profiles (26,94). However, all the above approaches did not directly assess the silencer activity but assumed an association with repressive epigenetic signatures and long-range interaction with the target genes that prevented an unbiased identification of silencer elements. In an attempt to directly identify silencer elements, Pang and Snyder developed a lentiviral screening approach named repressive ability of silencer elements (ReSE) to identify elements capable of repressing the pro-apoptotic protein, Caspase-9 (23). In this study, the authors identified *bona fide* silencers from open chromatin (FAIRE), but the lentivirus strategy is difficult to set up and has limited complexity, while a potential bias might be introduced by the random genomic integration of the lentivirus.

In line with the previous studies, we found that silencer activity was associated with repressive marks H3K27me3, H3K9me3 and H4K20me3. However, more detailed analyses suggested that distinct classes of silencers associated with different epigenetic signatures, including, relative enrichment of active histone marks (H3K4me3 and H3K27ac), CTCF binding, and repressive marks (either H3K27me3 or H3K9me3). Moreover, a subset of silencers containing STR or SINE REs is specifically enriched in H3K20me3, consistent with the role of this histone mark in the regulation of REs (83). The CTCF binding site was enriched in our set of identified silencers and defined a subset of silencers specifically bound by this factor. The association of CTCF with silencer function is intriguing given the general role of CTCF as an insulator factor. This is, however, consistent with a previous study that showed that the T39 region bound by CTCF functions as a strong silencer, but is devoid of insulator activity, thus suggesting that in some cases CTCF might be specifically required for silencer activity (32,87).

We identified several TFs associated with silencer activities. Of these, REST appeared to be associated with the strongest silencing activity in the pPGK and SCP1 libraries. Moreover, REST binding motif-containing silencers were associated with tissue-specific genes that are not normally expressed in T cells. This observation was consistent with the results obtained after the deletion of the REST-containing silencer DHS12366. Our results agree with the widespread role of REST repressor (95) and its general association with silencer elements (26,96). REST was initially described to be involved in the repression of neural genes in non-neuronal cells (65–67). However, the function of the REST TF is not restricted to the repression of neuronal genes and might be involved in the regulation of distinct developmental pathways (97–99). Our results suggested the existence of two main types of silencer elements. One type of silencer contains the REST binding site and appears to have ubiquitous silencer activity. The other type of silencers might have more tissue-restricted silencer activity (i.e. involved in the regulation of genes expressed during cell differentiation).

Some known silencers have been shown to recruit tissue-specific transcription factors with repressive activities, but the overall set of proteins that collaborate to impart silencer function is essentially unknown. In our study, we identified several TFs involved in T cell differentiation and function for which binding sites were associated with strong silencer activity. These included the hematopoietic-specific RUNX1 transcription factor (100), the HOXA family of developmental repressors (101), as well as the TGFb-signaling dependent SMAD3/4 repressors (102). Members of the Runx family have been shown to repress T cell-specific genes by binding to well-characterized silencers such as those found in the CD4 and ThPOK loci (7,8,103–106). We have previously shown that HOXA TFs play an important role during early T cell differentiation and their maintained expression induced T-acute lymphoblastic leukemia (107). SMAD-dependent TGFb signaling also plays an essential role in T cell differentiation and function (108,109). In addition, the nuclear receptor RARA has been shown to display ligand-dependent repressor and activator function in T-cells (110). We also found an association between MAFK and strong silencer activity, while mutation of MAFK binding sites resulted in the loss of silencer function. MAFK is a member of the small Maf proteins (sMafs) which are basic region leucine zipper (bZIP)-type transcription factors. sMafs lack

the transcriptional activation domain and hence, their homodimers act as transcriptional repressors and have been shown to play an important role in hematopoietic lineages (111). The direct involvement of these factors in the function of T cell regulated silencers will need to be investigated in the future.

A striking finding of our study was the association of SINE and STR REs with silencer activity of the pPGK library. STRs have been suggested to regulate gene expression in mammalian cells through various molecular mechanisms and to contribute to gene expression variation in humans (34,76,84). We found that STRs contained in strong silencers are potential binding sites for KZFPs, including ZNF322, ZNF410 and ZNF263 which are expressed in P5424 cells (Supplementary Table S4B). Several other KZFP motifs appeared to be enriched in the set of identified silencers (Supplementary Figure S3), including Zfp281, which has been shown to sustain $CD4^+$ T lymphocyte activation by directly repressing the Ctla-4 gene (112). The presence of tandem binding sites for repressor factors is reminiscent of a previous study showing that the high density of identical motifs of ZEB1 repressor in tandem repeats can make them suitable platforms for the recruitment of transcriptional repressors (113). KZFPs play a major role in the recognition and transcriptional silencing of REs (114,115). The majority of KZFPs bind to transposable elements (TEs), including LTR, L1, SINE, and SVA families, as well as simple repeats and other variable number tandem repeats, including zinc finger repeats. KZFP tethers KRAB associated protein 1 (KAP1, also known as TRIM28) to the DNA. In turn, KAP1 can recruit different epigenetic effectors, such as histone methyltransferases, nucleosome remodeling and deacetylation (NuRD) complex or DNA methyltransferases (116,117). Overall, these results suggest that KZFP binding might be a general feature of silencer elements, while it is tempting to speculate that STRs bound by the KZFP family of transcriptional repressors have been co-opted in the mammalian genome as repressive *cis*-regulatory modules.

Besides STRs, TEs and in particular SINEs, might also play a role in silencer activity. TEs are silenced by the targeted deposition of repressive histone modifications (33). For example, in mESCs, TEs are silenced through H3K9 tri-methylation by SET domain bifurcated histone lysine methyltransferase 1 (SETDB1), which is recruited to TEs by KZFPs through interaction with KAP1 (118). SETDB1 knockout leads to widespread de-repression of class I and class II ERV elements and transcription of chimeric RNAs, suggesting that repression of these elements not only prevents mutagenic transposition but also deleterious *cis*-regulatory effects (119). The current model suggests that, rather than a defense system against transposition, the KZFP system may enable the genomic accumulation of TEs with strong *cis*-regulatory elements (such as LTR elements), which increases the likelihood of these elements being subsequently co-opted for host functions (33,118). Recent examples of TEs co-opted as silencer elements included a SINE element involved in the silencing of a T-cell specific gene (21) and endogenous retrovirus (ERV) involved in the repression of immune genes (81). Similarly, SETDB1-mediated repression of SINE B2 repeats restricts the usage of functional CTCF sites (120). These examples reveal that TE silencing not only affects TE activity but can also have collateral effects on the regulation of host-gene transcription.

Overall, our study represents an initial step toward the understanding of the molecular basis driving silencer activity, including epigenetic features and binding of transcription factors. We provided experimental evidence that the STARR-seq approach is able to identify silencers functioning in the endogenous context and likely playing a key physiological role in the regulation of T cell differentiation and function. Thus, further implementation of STARR-seq in different cellular systems will help in the functional assessment of mammalian silencers that are active in specific pathways or induced by specific stimuli.

DATA AVAILABILITY

Sequencing data generated in this study have been submitted to the Gene Expression Omnibus (GEO) under the accession codes GSE202547 (CapSTARR-seq) and GSE219157 (CUT&Tag and CUT&RUN). All other datasets used in this study are listed in Supplemental Table S1.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the Transcriptomics and Genomics Marseille-Luminy (TGML) platform for sequencing the CapSTARR-seq samples and the Marseille-Luminy cell biology platform for the management of cell culture. TGML is a member of the France Genomique consortium (ANR-10-INBS-0009). *Author contributions:* S.Sp. and S.Sa. conceived the project and secured funding. S.Sp. supervised the project. S.H., N.S. and S.Sp. conceptualized and designed the experiments. S.H. performed the majority of the experimental work, with the help of LTMD and MT. N.S. performed the majority of the bioinformatic analyses, with the help of QF and GC. C.L. contributed to the analyses of repetitive sequences. T.S. performed analyses of Hi-C data. D.vE. and S.Sa. contributed to the design of the CapSTARR-seq approach as well as bioinformatics identification of silencers. S.H., N.S. and S.Sp. wrote the manuscript. All authors contributed to reading, discussion and commenting on the manuscript.

FUNDING

Work in the laboratory of S. Spicuglia was supported by recurrent funding from Institut National de la Santé et de la Recherche Médicale and Aix-Marseille University; Cancerop\u00f4le PACA, AMIDEX [ANR-11-IDEX-0001-02]; INCA [PLBIO018-031 INCA_12619]; Ligue contre le Cancer (Equipe Labellisée 2018); ANR [ANR-18-CE12-0019, ANR-17-CE12-0035, ANR-22-CE45-0031-02]; Bettencourt Schueller Foundation (Prix coup d'élan pour la recherche française); S.H. and N.S. were supported by PhD fellowships from the Pakistani's government and Marseille Institute of Rare Diseases (MARMARA), respectively. Funding for open access charge: ANR.

Conflict of interest statement. None declared.

REFERENCES

1. Chatterjee,S. and Ahituv,N. (2017) Gene regulatory elements, major drivers of human disease. *Annu. Rev. Genomics Hum. Genet.***18**, 4.1–4.19.
2. Maston,G.A., Evans,S.K. and Green,M.R. (2006) Transcriptional regulatory elements in the human genome. *Annu. Rev. Genomics Hum. Genet.*, **7**, 29–59.
3. Ogbourne,S. and Antalis,T.M. (1998) Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *Biochem. J.*, **331(Pt 1)**, 1–14.
4. Baniahmad,A., Muller,M., Steiner,C. and Renkawitz,R. (1987) Activity of two different silencer elements of the chicken lysozyme gene can be compensated by enhancer elements. *EMBO J.*, **6**, 2297–2303.
5. Brand,A.H., Breeden,L., Abraham,J., Sternglanz,R. and Nasmyth,K. (1985) Characterization of a "silencer" in yeast: a DNA sequence with properties opposite to those of a transcriptional enhancer. *Cell*, **41**, 41–48.
6. Kadesch,T., Zervos,P. and Ruezinsky,D. (1986) Functional analysis of the murine IgH enhancer: evidence for negative control of cell-type specificity. *Nucleic Acids Res.*, **14**, 8209–8221.
7. Taniuchi,I., Sunshine,M.J., Festenstein,R. and Littman,D.R. (2002) Evidence for distinct CD4 silencer functions at different stages of thymocyte differentiation. *Mol. Cell*, **10**, 1083–1096.
8. Setoguchi,R., Tachibana,M., Naoe,Y., Muroi,S., Akiyama,K., Tezuka,C., Okuda,T. and Taniuchi,I. (2008) Repression of the transcription factor Th-POK by Runx complexes in cytotoxic T cell development. *Science*, **319**, 822–825.
9. Huang,Z., Liang,N., Goni,S., Damdimopoulos,A., Wang,C., Ballaire,R., Jager,J., Niskanen,H., Han,H., Jakobsson,T. *et al.* (2021) The corepressors GPS2 and SMRT control enhancer and silencer remodeling via eRNA transcription during inflammatory activation of macrophages. *Mol. Cell*, **81**, 953–968.
10. Petrykowska,H.M., Vockley,C.M. and Elnitski,L. (2008) Detection and characterization of silencers and enhancer-blockers in the greater CFTR locus. *Genome Res.*, **18**, 1238–1246.
11. Liu,Z., Widlak,P., Zou,Y., Xiao,F., Oh,M., Li,S., Chang,M.Y., Shay,J.W. and Garrard,W.T. (2006) A recombination silencer that specifies heterochromatin positioning and ikaros association in the immunoglobulin kappa locus. *Immunity*, **24**, 405–415.
12. Yadav,D.K., Shrestha,S., Dadhwal,G. and Chandak,G.R. (2018) Identification and characterization of cis-regulatory elements 'insulator and repressor' in PPARD gene. Epigenomics, **10**, 613–627.
13. Tran,T.H., Nakata,M., Suzuki,K., Begum,N.A., Shinkura,R., Fagarasan,S., Honjo,T. and Nagaoka,H. (2010) B cell-specific and stimulation-responsive enhancers derepress Aicda by overcoming the effects of silencers. *Nat. Immunol.*, **11**, 148–154.
14. Oh,C.K., Neurath,M., Cho,J.J., Semere,T. and Metcalfe,D.D. (1997) Two different negative regulatory elements control the transcription of T-cell activation gene 3 in activated mast cells. Biochem. J., **323 (Pt 2)**, 511–519.
15. Williams,T.M., Moolten,D., Burlein,J., Romano,J., Bhaerman,R., Godillot,A., Mellon,M., Rauscher,F.J. 3rd and Kant,J.A. (1991) Identification of a zinc finger protein that inhibits IL-2 gene expression. *Science*, **254**, 1791–1794.
16. Sawada,S., Scarborough,J.D., Killeen,N. and Littman,D.R. (1994) A lineage-specific transcriptional silencer regulates CD4 gene expression during T lymphocyte development. *Cell*, **77**, 917–929.
17. Dombret,H., Font,M.P. and Sigaux,F. (1996) A dominant transcriptional silencer located 5′ to the human T-cell receptor V·2.2 gene segment which is activated in cell lines of thymic phenotype. *Nucleic Acids Res.*, **24**, 2782–2789.
18. He,X., Park,K., Wang,H., Zhang,Y., Hua,X., Li,Y. and Kappes,D.J. (2008) CD4-CD8 lineage commitment is regulated by a silencer element at the ThPOK transcription-factor locus. *Immunity*, **28**, 346–358.
19. Yannoutsos,N., Barreto,V., Misulovin,Z., Gazumyan,A., Yu,W., Rajewsky,N., Peixoto,B.R., Eisenreich,T. and Nussenzweig,M.C. (2004) A cis element in the recombination activating gene locus regulates gene expression by counteracting a distant silencer. *Nat. Immunol.*, **5**, 443–450.
20. Yao,X., Nie,H., Rojas,I.C., Harriss,J.V., Maika,S.D., Gottlieb,P.D., Rathbun,G. and Tucker,P.W. (2010) The L2a element is a mouse CD8 silencer that interacts with MAR-binding proteins SATB1 and CDP. *Mol. Immunol.*, **48**, 153–163.
21. Hosokawa,H., Koizumi,M., Masuhara,K., Romero-Wolf,M., Tanaka,T., Nakayama,T. and Rothenberg,E.V. (2021) Stage-specific action of Runx1 and GATA3 controls silencing of PU.1 expression in mouse pro-T cells. *J. Exp. Med.*, **218**, e20202648.
22. Della Rosa,M. and Spivakov,M. (2020) Silencers in the spotlight. *Nat. Genet.*, **52**, 244–245.
23. Pang,B. and Snyder,M.P. (2020) Systematic identification of silencers in human cells. *Nat. Genet.*, **52**, 254–263.
24. Ngan,C.Y., Wong,C.H., Tjong,H., Wang,W., Goldfeder,R.L., Choi,C., He,H., Gong,L., Lin,J., Urban,B. *et al.* (2020) Chromatin interaction analyses elucidate the roles of PRC2-bound silencers in mouse development. *Nat. Genet.*, **52**, 264–272.
25. Mifsud,B., Tavares-Cadete,F., Young,A.N., Sugar,R., Schoenfelder,S., Ferreira,L., Wingett,S.W., Andrews,S., Grey,W., Ewels,P.A. *et al.* (2015) Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat. Genet.*, **47**, 598–606.
26. Doni Jayavelu,N., Jajodia,A., Mishra,A. and Hawkins,R.D. (2020) Candidate silencer elements for the human and mouse genomes. *Nat. Commun.*, **11**, 1061.
27. Zhang,P., Xia,J.H., Zhu,J., Gao,P., Tian,Y.J., Du,M., Guo,Y.C., Suleman,S., Zhang,Q., Kohli,M. *et al.* (2018) High-throughput screening of prostate cancer risk loci by single nucleotide polymorphisms sequencing. *Nat. Commun.*, **9**, 2022.
28. Santiago-Algarra,D., Dao,L.T.M., Pradel,L., Espana,A. and Spicuglia,S. (2017) Recent advances in high-throughput approaches to dissect enhancer function. F1000Res., **6**, 939.
29. Gasperini,M., Tome,J.M. and Shendure,J. (2020) Towards a comprehensive catalogue of validated and target-linked human enhancers. *Nat. Rev. Genet.*, **21**, 292–310.
30. Arnold,C.D., Gerlach,D., Stelzer,C., Boryn,L.M., Rath,M. and Stark,A. (2013) Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science*, **339**, 1074–1077.
31. Vanhille,L., Griffon,A., Maqbool,M.A., Zacarias-Cabeza,J., Dao,L.T., Fernandez,N., Ballester,B., Andrau,J.C. and Spicuglia,S. (2015) High-throughput and quantitative assessment of enhancer activity in mammals by CapStarr-seq. *Nat. Commun.*, **6**, 6905.
32. Qi,H., Liu,M., Emery,D.W. and Stamatoyannopoulos,G. (2015) Functional validation of a constitutive autonomous silencer element. *PLoS One*, **10**, e0124588.
33. Fueyo,R., Judd,J., Feschotte,C. and Wysocka,J. (2022) Roles of transposable elements in the regulation of mammalian transcription. *Nat. Rev. Mol. Cell Biol.***23**, 481–497.
34. Gymrek,M., Willems,T., Guilmatre,A., Zeng,H., Markus,B., Georgiev,S., Daly,M.J., Price,A.L., Pritchard,J.K., Sharp,A.J. *et al.* (2016) Abundant contribution of short tandem repeats to gene expression variation in humans. *Nat. Genet.*, **48**, 22–29.
35. Mombaerts,P., Terhorst,C., Jacks,T., Tonegawa,S. and Sancho,J. (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. U.S.A.*, **92**, 7420–7424.
36. Saadi,W., Kermezli,Y., Dao,L.T.M., Mathieu,E., Santiago-Algarra,D., Manosalva,I., Torres,M., Belhocine,M., Pradel,L., Loriod,B. *et al.* (2019) A critical regulator of Bcl2 revealed by systematic transcript discovery of lncRNAs associated with T-cell differentiation. *Scientific Rep.*, **9**, 4707.
37. Qin,J.Y., Zhang,L., Clift,K.L., Hulur,I., Xiang,A.P., Ren,B.Z. and Lahn,B.T. (2010) Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. *PLoS One*, **5**, e10611.
38. Wei,X.C., Dohkan,J., Kishi,H., Wu,C.X., Kondo,S. and Muraguchi,A. (2005) Characterization of the proximal enhancer element and transcriptional regulatory factors for murine recombination activating gene-2. *Eur. J. Immunol.*, **35**, 612–621.
39. Dao,L.T.M., Vanhille,L., Griffon,A., Fernandez,N. and Spicuglia,S. (2015) CapStarr-seq protocol. Protocol Exchange, https://doi.org/10.1038/protex.2015.096.
40. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.

41. McLean,C.Y., Bristor,D., Hiller,M., Clarke,S.L., Schaar,B.T., Lowe,C.B., Wenger,A.M. and Bejerano,G. (2010) GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.*, **28**, 495–501.
42. O'Leary,N.A., Wright,M.W., Brister,J.R., Ciufo,S., Haddad,D., McVeigh,R., Rajput,B., Robbertse,B., Smith-White,B., Ako-Adjei,D. *et al.* (2016) Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.*, **44**, D733–D745.
43. Kent,W.J., Zweig,A.S., Barber,G., Hinrichs,A.S. and Karolchik,D. (2010) BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics*, **26**, 2204–2207.
44. Ramirez,F., Dundar,F., Diehl,S., Gruning,B.A. and Manke,T. (2014) deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.*, **42**, W187–W191.
45. Hu,G., Tang,Q., Sharma,S., Yu,F., Escobar,T.M., Muljo,S.A., Zhu,J. and Zhao,K. (2013) Expression and regulation of intergenic long noncoding rnas during T cell development and differentiation. *Nat. Immunol.*, **14**, 1190–1198.
46. Su,A.I., Wiltshire,T., Batalov,S., Lapp,H., Ching,K.A., Block,D., Zhang,J., Soden,R., Hayakawa,M., Kreiman,G. *et al.* (2004) A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 6062–6067.
47. Heinz,S., Benner,C., Spann,N., Bertolino,E., Lin,Y.C., Laslo,P., Cheng,J.X., Murre,C., Singh,H. and Glass,C.K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell*, **38**, 576–589.
48. Khan,A., Fornes,O., Stigliani,A., Gheorghe,M., Castro-Mondragon,J.A., van der Lee,R., Bessy,A., Chèneby,J., Kulkarni,S.R., Tan,G. *et al.* (2018) JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res.*, **46**, D260–D266.
49. Castro-Mondragon,J.A., Jaeger,S., Thieffry,D., Thomas-Chollier,M. and van Helden,J. (2017) RSAT matrix-clustering: dynamic exploration and redundancy reduction of transcription factor binding motif collections. *Nucleic Acids Res.*, **45**, e119.
50. Ferré,Q., Charbonnier,G., Sadouni,N., Lopez,F., Kermezli,Y., Spicuglia,S., Capponi,C., Ghattas,B. and Puthier,D. (2020) OLOGRAM: determining significance of total overlap length between genomic regions sets. *Bioinformatics*, **36**, 1920–1922.
51. Lopez,F., Charbonnier,G., Kermezli,Y., Belhocine,M., Ferre,Q., Zweig,N., Aribi,M., Gonzalez,A., Spicuglia,S. and Puthier,D. (2019) Explore, edit and leverage genomic annotations using Python GTF toolkit. *Bioinformatics*.
52. Chen,N. (2004) Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinformatics*, **Chapter 4**, Unit 4.10.
53. Jurka,J., Kapitonov,V.V., Pavlicek,A., Klonowski,P., Kohany,O. and Walichiewicz,J. (2005) Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.*, **110**, 462–467.
54. Willems,T., Zielinski,D., Yuan,J., Gordon,A., Gymrek,M. and Erlich,Y. (2017) Genome-wide profiling of heritable and de novo STR variations. *Nat. Methods*, **14**, 590–592.
55. Yu,G. (2020) Gene ontology semantic similarity analysis using GOSemSim. *Methods Mol. Biol.*, **2117**, 207–215.
56. Hu,G., Cui,K., Fang,D., Hirose,S., Wang,X., Wangsa,D., Jin,W., Ried,T., Liu,P., Zhu,J. *et al.* (2018) Transformation of accessible chromatin and 3D nucleome underlies lineage commitment of early T cells. *Immunity*, **48**, 227–242.
57. Kruse,K., Hug,C.B. and Vaquerizas,J.M. (2020) FAN-C: a feature-rich framework for the analysis and visualisation of chromosome conformation capture data. *Genome Biol.*, **21**, 303.
58. Crane,E., Bian,Q., McCord,R.P., Lajoie,B.R., Wheeler,B.S., Ralston,E.J., Uzawa,S., Dekker,J. and Meyer,B.J. (2015) Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature*, **523**, 240–244.
59. Naito,Y., Hino,K., Bono,H. and Ui-Tei,K. (2015) CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics*, **31**, 1120–1123.
60. Mali,P., Yang,L., Esvelt,K.M., Aach,J., Guell,M., DiCarlo,J.E., Norville,J.E. and Church,G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823–826.
61. Donda,A., Schulz,M., Burki,K., De Libero,G. and Uematsu,Y. (1996) Identification and characterization of a human CD4 silencer. *Eur. J. Immunol.*, **26**, 493–500.
62. Maqbool,M.A., Pioger,L., El Aabidine,A.Z., Karasu,N., Molitor,A.M., Dao,L.T.M., Charbonnier,G., van Laethem,F., Fenouil,R., Koch,F. *et al.* (2020) Alternative enhancer usage and targeted polycomb marking hallmark promoter choice during T cell differentiation. *Cell Rep.*, **32**, 108048.
63. Zabidi,M.A., Arnold,C.D., Schernhuber,K., Pagani,M., Rath,M., Frank,O. and Stark,A. (2015) Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. *Nature*, **518**, 556–559.
64. Weiner,A., Lara-Astiaso,D., Krupalnik,V., Gafni,O., David,E., Winter,D.R., Hanna,J.H. and Amit,I. (2016) Co-ChIP enables genome-wide mapping of histone mark co-occurrence at single-molecule resolution. *Nat. Biotechnol.*, **34**, 953–961.
65. Schoenherr,C.J. and Anderson,D.J. (1995) The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. *Science*, **267**, 1360–1363.
66. Chong,J.A., Tapia-Ram°rez,J., Kim,S., Toledo-Aral,J.J., Zheng,Y., Boutros,M.C., Altshuller,Y.M., Frohman,M.A., Kraner,S.D. and Mandel,G. (1995) REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons. *Cell*, **80**, 949–957.
67. Coulson,J.M. (2005) Transcriptional regulation: cancer, neurons and the REST. *Curr. Biol.*, **15**, R665–R668.
68. Wu,J., Song,Y., Bakker,A.B., Bauer,S., Spies,T., Lanier,L.L. and Phillips,J.H. (1999) An activating immunoreceptor complex formed by NKG2D and DAP10. *Science*, **285**, 730–732.
69. Diefenbach,A., Tomasello,E., Lucas,M., Jamieson,A.M., Hsia,J.K., Vivier,E. and Raulet,D.H. (2002) Selective associations with signaling proteins determine stimulatory versus costimulatory activity of NKG2D. *Nat. Immunol.*, **3**, 1142–1149.
70. Schuster,M., Annemann,M., Plaza-Sirvent,C. and Schmitz,I. (2013) Atypical IkappaB proteins - nuclear modulators of NF-kappaB signaling. *Cell Commun. Signal.*, **11**, 23.
71. Schuster,M., Glauben,R., Plaza-Sirvent,C., Schreiber,L., Annemann,M., Floess,S., Kuhl,A.A., Clayton,L.K., Sparwasser,T., Schulze-Osthoff,K. *et al.* (2012) IkappaB(NS) protein mediates regulatory T cell development via induction of the Foxp3 transcription factor. *Immunity*, **37**, 998–1008.
72. Ebihara,T. and Taniuchi,I. (2019) Transcription factors in the development and function of group 2 innate lymphoid cells. *Int. J. Mol. Sci.*, **20**, 1377.
73. Anguita,E., Candel,F.J., Chaparro,A. and Roldan-Etcheverry,J.J. (2017) Transcription factor GFI1B in health and disease. *Front. Oncol.*, **7**, 54.
74. Chiang,C. and Ayyanathan,K. (2013) Snail/Gfi-1 (SNAG) family zinc finger proteins in transcription regulation, chromatin dynamics, cell signaling, development, and disease. *Cytokine Growth Factor Rev.*, **24**, 123–131.
75. Moroy,T. and Khandanpour,C. (2011) Growth factor independence 1 (Gfi1) as a regulator of lymphocyte development and activation. *Semin. Immunol.*, **23**, 368–378.
76. Bagshaw,A.T.M. (2017) Functional mechanisms of microsatellite DNA in eukaryotic genomes. *Genome Biol Evol*, **9**, 2428–2443.
77. Barakat,T.S., Halbritter,F., Zhang,M., Rendeiro,A.F., Perenthaler,E., Bock,C. and Chambers,I. (2018) Functional dissection of the enhancer repertoire in Human embryonic stem cells. *Cell Stem Cell*, **23**, 276–288.
78. Glinsky,G.V. (2015) Transposable elements and DNA methylation create in embryonic stem cells Human-specific regulatory sequences associated with distal enhancers and noncoding rnas. *Genome Biol Evol*, **7**, 1432–1454.
79. Kunarso,G., Chia,N.Y., Jeyakani,J., Hwang,C., Lu,X., Chan,Y.S., Ng,H.H. and Bourque,G. (2010) Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nat. Genet.*, **42**, 631–634.
80. Wang,X., He,L., Goggin,S.M., Saadat,A., Wang,L., Sinnott-Armstrong,N., Claussnitzer,M. and Kellis,M. (2018) High-resolution genome-wide functional dissection of transcriptional regulatory regions and nucleotides in human. *Nat. Commun.*, **9**, 5380.
81. Adoue,V., Binet,B., Malbec,A., Fourquet,J., Romagnoli,P., van Meerwijk,J.P.M., Amigorena,S. and Joffre,O.P. (2019) The histone

methyltransferase SETDB1 controls T helper cell lineage integrity by repressing endogenous retroviruses. *Immunity*, **50**, 629–644.
82. Kent,W.J., Sugnet,C.W., Furey,T.S., Roskin,K.M., Pringle,T.H., Zahler,A.M. and Haussler,D. (2002) The human genome browser at UCSC. *Genome Res.*, **12**, 996–1006.
83. Jørgensen,S., Schotta,G. and Sørensen,C.S. (2013) Histone H4 lysine 20 methylation: key player in epigenetic regulation of genomic integrity. *Nucleic Acids Res.*, **41**, 2797–2806.
84. Grapotte,M., Saraswat,M., Bessiere,C., Menichelli,C., Ramilowski,J.A., Severin,J., Hayashizaki,Y., Itoh,M., Tagami,M., Murata,M *et al.* (2021) Discovery of widespread transcription initiation at microsatellites predictable by sequence-based deep neural network. *Nat. Commun.*, **12**, 3297.
85. van Arensbergen,J., FitzPatrick,V.D., de Haas,M., Pagie,L., Sluimer,J., Bussemaker,H.J. and van Steensel,B. (2017) Genome-wide mapping of autonomous promoter activity in human cells. *Nat. Biotechnol.*, **35**, 145–153.
86. Weingarten-Gabbay,S., Nir,R., Lubliner,S., Sharon,E., Kalma,Y., Weinberger,A. and Segal,E. (2019) Systematic interrogation of human promoters. *Genome Res.*, **29**, 171–183.
87. Liu,M., Maurano,M.T., Wang,H., Qi,H., Song,C.Z., Navas,P.A., Emery,D.W., Stamatoyannopoulos,J.A. and Stamatoyannopoulos,G. (2015) Genomic discovery of potent chromatin insulators for human gene therapy. *Nat. Biotechnol.*, **33**, 198–203.
88. Wei,X., Xiang,Y., Peters,D.T., Marius,C., Sun,T., Shan,R., Ou,J., Lin,X., Yue,F., Li,W. *et al.* (2022) HiCAR is a robust and sensitive method to analyze open-chromatin-associated genome organization. *Mol. Cell*, **82**, 1225–1238.
89. Dao,L.T.M., Galindo-Albarran,A.O., Castro-Mondragon,J.A., Andrieu-Soler,C., Medina-Rivera,A., Souaid,C., Charbonnier,G., Griffon,A., Vanhille,L., Stephen,T. *et al.* (2017) Genome-wide characterization of mammalian promoters with distal enhancer functions. *Nat. Genet.*, **49**, 1073–1081.
90. Diao,Y., Fang,R., Li,B., Meng,Z., Yu,J., Qiu,Y., Lin,K.C., Huang,H., Liu,T., Marina,R.J *et al.* (2017) A tiling-deletion-based genetic screen for cis-regulatory element identification in mammalian cells. *Nat. Methods*, **14**, 629–635.
91. Engreitz,J.M., Haines,J.E., Perez,E.M., Munson,G., Chen,J., Kane,M., McDonel,P.E., Guttman,M. and Lander,E.S. (2016) Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature*, **539**, 452–455.
92. Andersson,R. (2015) Promoter or enhancer, what's the difference? Deconstruction of established distinctions and presentation of a unifying model. *Bioessays*, **37**, 314–323.
93. Cai,Y., Zhang,Y., Loh,Y.P., Tng,J.Q., Lim,M.C., Cao,Z., Raju,A., Lieberman Aiden,E., Li,S., Manikandan,L. *et al.* (2021) H3K27me3-rich genomic regions can function as silencers to repress gene expression via chromatin interactions. *Nat. Commun.*, **12**, 719.
94. Huang,D., Petrykowska,H.M., Miller,B.F., Elnitski,L. and Ovcharenko,I. (2019) Identification of human silencers by correlating cross-tissue epigenetic profiles and gene expression. *Genome Res.*, **29**, 657–667.
95. Ooi,L. and Wood,I.C. (2007) Chromatin crosstalk in development and disease: lessons from REST. *Nat. Rev. Genet.*, **8**, 544–554.
96. Ernst,J., Melnikov,A., Zhang,X., Wang,L., Rogov,P., Mikkelsen,T.S. and Kellis,M. (2016) Genome-scale high-resolution mapping of activating and repressive nucleotides in regulatory regions. *Nat. Biotechnol.*, **34**, 1180–1190.
97. Kuwahara,K., Saito,Y., Takano,M., Arai,Y., Yasuno,S., Nakagawa,Y., Takahashi,N., Adachi,Y., Takemura,G., Horie,M. *et al.* (2003) NRSF regulates the fetal cardiac gene program and maintains normal cardiac structure and function. *EMBO J.*, **22**, 6310–6321.
98. Martin,D., Tawadros,T., Meylan,L., Abderrahmani,A., Condorelli,D.F., Waeber,G. and Haefliger,J.A. (2003) Critical role of the transcriptional repressor neuron-restrictive silencer factor in the specific control of connexin36 in insulin-producing cell lines. *J. Biol. Chem.*, **278**, 53082–53089.
99. Rockowitz,S., Lien,W.H., Pedrosa,E., Wei,G., Lin,M., Zhao,K., Lachman,H.M., Fuchs,E. and Zheng,D. (2014) Comparison of REST cistromes across human cell types reveals common and context-specific functions. *PLoS Comput. Biol.*, **10**, e1003671.
100. Collins,A., Littman,D.R. and Taniuchi,I. (2009) RUNX proteins in transcription factor networks that regulate T-cell lineage choice. *Nat. Rev. Immunol.*, **9**, 106–115.
101. Cain,B. and Gebelein,B. (2021) Mechanisms underlying hox-mediated transcriptional outcomes. *Front. Cell Dev. Biol.*, **9**, 787339.
102. Hill,C.S. (2016) Transcriptional control by the SMADs. *Cold Spring Harb. Perspect. Biol.*, **8**, 10.
103. Jiang,H. and Peterlin,B.M. (2008) Differential chromatin looping regulates CD4 expression in immature thymocytes. *Mol. Cell. Biol.*, **28**, 907–912.
104. Wildt,K.F., Sun,G., Grueter,B., Fischer,M., Zamisch,M., Ehlers,M. and Bosselut,R. (2007) The transcription factor Zbtb7b promotes CD4 expression by antagonizing runx-mediated activation of the CD4 silencer. *J. Immunol.*, **179**, 4405–4414.
105. Jiang,H., Zhang,F., Kurosu,T. and Peterlin,B.M. (2005) Runx1 Binds positive transcription elongation factor b and represses transcriptional elongation by RNA polymerase II: possible mechanism of CD4 silencing. *Mol. Cell. Biol.*, **25**, 10675–10683.
106. Telfer,J.C., Hedblom,E.E., Anderson,M.K., Laurent,M.N. and Rothenberg,E.V. (2004) Localization of the domains in runx transcription factors required for the repression of CD4 in thymocytes. *J. Immunol.*, **172**, 4359–4370.
107. Cieslak,A., Charbonnier,G., Tesio,M., Mathieu,E.L., Belhocine,M., Touzart,A., Smith,C., Hypolite,G., Andrieu,G.P., Martens,J.H.A. *et al.* (2020) Blueprint of human thymopoiesis reveals molecular mechanisms of stage-specific TCR enhancer activation. *J. Exp. Med.*, **217**, e20192360.
108. Gu,A.D., Wang,Y., Lin,L., Zhang,S.S. and Wan,Y.Y. (2012) Requirements of transcription factor smad-dependent and -independent TGF-β signaling to control discrete T-cell functions. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 905–910.
109. Takimoto,T., Wakabayashi,Y., Sekiya,T., Inoue,N., Morita,R., Ichiyama,K., Takahashi,R., Asakawa,M., Muto,G., Mori,T. *et al.* (2010) Smad2 and Smad3 are redundantly essential for the TGF-beta-mediated regulation of regulatory T plasticity and Th1 development. *J. Immunol.*, **185**, 842–855.
110. Gordy,C., Dzhagalov,I. and He,Y.W. (2009) Regulation of CD8(+) T cell functions by RARgamma. *Semin. Immunol.*, **21**, 2–7.
111. Katsuoka,F. and Yamamoto,M. (2016) Small Maf proteins (MafF, MafG, MafK): history, structure and function. *Gene*, **586**, 197–205.
112. Guo,J., Xue,Z., Ma,R., Yi,W., Hui,Z., Guo,Y., Yao,Y., Cao,W., Wang,J., Ju,Z *et al.* (2020) The transcription factor Zfp281 sustains CD4. *Cell. Mol. Immunol.*, **17**, 1222–1232.
113. Balestrieri,C., Alfarano,G., Milan,M., Tosi,V., Prosperini,E., Nicoli,P., Palamidessi,A., Scita,G., Diaferia,G.R. and Natoli,G. (2018) Co-optation of tandem DNA repeats for the maintenance of mesenchymal identity. *Cell*, **173**, 1150–1164.
114. Yang,P., Wang,Y. and Macfarlan,T.S. (2017) The role of KRAB-zfps in transposable element repression and mammalian evolution. *Trends Genet.*, **33**, 871–881.
115. Ecco,G., Imbeault,M. and Trono,D. (2017) KRAB zinc finger proteins. *Development*, **144**, 2719–2729.
116. Helleboid,P.Y., Heusel,M., Duc,J., Piot,C., Thorball,C.W., Coluccio,A., Pontis,J., Imbeault,M., Turelli,P., Aebersold,R. *et al.* (2019) The interactome of KRAB zinc finger proteins reveals the evolutionary history of their functional diversification. *EMBO J.*, **38**, e101220.
117. Jang,S.M., Kauzlaric,A., Quivy,J.P., Pontis,J., Rauwel,B., Coluccio,A., Offner,S., Duc,J., Turelli,P., Almouzni,G. *et al.* (2018) KAP1 facilitates reinstatement of heterochromatin after DNA replication. *Nucleic Acids Res.*, **46**, 8788–8802.
118. Imbeault,M., Helleboid,P.Y. and Trono,D. (2017) KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. *Nature*, **543**, 550–554.
119. Karimi,M.M., Goyal,P., Maksakova,I.A., Bilenky,M., Leung,D., Tang,J.X., Shinkai,Y., Mager,D.L., Jones,S., Hirst,M. *et al.* (2011) DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell*, **8**, 676–687.
120. Gualdrini,F., Polletti,S., Simonatto,M., Prosperini,E., Pileri,F. and Natoli,G. (2022) H3K9 trimethylation in active chromatin restricts the usage of functional CTCF sites in SINE B2 repeats. *Genes Dev.***36**, 1–19.