MATERIAL AND METHODS

Mice and $CD4^+$ T cell isolation

CAST/EiJ male mice were maintained under specific pathogen-free conditions at the University of Cambridge, CRUK – Cambridge Institute under the auspices of a UK Home Office license. Inbred wild-type C57/BL6 mice were purchased from Charles River UK Ltd (Margate, United Kingdom). Animals were euthanized in accordance with Schedule 1 of the Animals (Scientific Procedures) Act 1986. Each animal used was macroscopically examined. Animals with lesions or phenotypic alterations in their internal organs were discarded.

Unstimulated naive $CD4^+$ T cells were purified from dissociated mouse spleens using EASY cell strainer (30 µm, Greiner BioOne), cell separation media (lympholyte, #CL5035) and the $CD4^+$ $CD62L^+$ T Cell Isolation Kit II (Miltenyi Biotec, #130-093-227). Flow cytometry confirmed that 96.4% of the isolated CD4+ T cells were naive in young B6 (Fig. S2D). Naive $CD4^+$ T cells formed a single, high-purity population in young animals. Old animals had a small population of $CD4^+$ T cells with slightly elevated CD44 levels, reduced CD62L expression, and attenuated activation dynamics (Fig. S2E-G); their removal did not impact our results (Fig. S7C).

Purified naive $CD4^+$ T cells were cultured in IMDM medium (GIBCO, #21980-032) supplemented with 10% Fetal Bovine Serum (Life Technology, #10500064), 1 µg/mL Penicillin/Streptomicin (Life Technology, #15070063), and 50 µM 2-mercaptoethanol (Gibco, #31350-010). Cells were seeded into 96-well plates coated for 1h at 37ºC with anti-CD3ε (1 µg/ml, clone: 145-2C11, eBioscience, #16-0031-82) and anti-CD28 (3 µg/ml, clone: 37.51, eBioscience, #16-0281-82) at a density of 80,000-120,000 cells/ml, and then cultured in a total volume of 100 µl media that did not contain cytokines or additional antibodies. We did not use additional cytokines to commit the naive CD4+ cells to specific helper cell lineages (*22*).

For clarity, we called purified unstimulated naive $CD4^+$ T cells *naive*, and stimulated cells *activated*.

All cells were cultured in a humidified incubator at 37°C, with 5% CO_2. Naive and activated $CD4^+$ T cells were then immediately collected and loaded on a 5–10 µm Auto Prep Integrated Fluidic Circuit (IFC; Fluidigm, San Francisco, CA) to capture single cells using the C1 Single cell Auto Prep System (Fluidigm). All the IFCs were visually inspected, and wells with multiple cells or cell debris were identified per instructions of the manufacturer (PN 101-2711 A1 White Paper). Upon cell capture, reverse transcription and cDNA amplification were performed using the SMARTer PCR cDNA Synthesis Kit (Clontech) and the Advantage 2 PCR Kit (Clontech) (*23*). ERCC spike-in RNA (Ambion) (1 µL diluted at 1:50,000) was added to the C1 lysis mix. All the capture sites were included for the RNA-seq library preparation, and wells identified above as multiple cells or containing debris were removed during computational analysis.

Flow cytometry

Naive $CD4^+$ T cells were purified from spleens of young and old C57/BL6 mice (see above). Purified cells were, directly or after 3h activation *in vitro* (see above), incubated with TruStainfcX (anti-mouse CD16/32, clone:93, BioLegend) before

staining with immunofluorescence conjugated antibodies against murine CD4 (clone: RM4-5, BioLegend), CD44 (clone: IM7, BioLegend), CD62L (clone: MEL-14, BioLegend), CD25 (clone: 3C7, BioLegend), CD69 (clone: H1.2F3, BioLegend), CD127 (clone: A7R34, BioLegend), and KLRG1 (clone: 2F1, BD Biosciences). Cell viability was determined using Fixable eFluor 780 viability dye (eBioscience). Data were acquired on a 5-laser Aria IIu SORP instrument (BD Biosciences) and data analysis was performed using FlowJo software (Tree Star).

Naive and effector memory $CD4^+$ T cells were purified from spleens of both young and old C57/BL6 mice by FACS. Briefly, spleens were harvested from both young and old animals and single cell suspensions were obtained by meshing through a cell strainer (70 μm). B cells were depleted from cell suspensions by MACS using CD19 microbeads (Miltenyi Biotec, #130-052-201) and red blood cells were lysed with RBC lysis buffer (Biolegend, #B205551). The enriched cell fraction was then stained with Fixable eFluor 780 viability dye (eBioscience) following by Fc receptor blocking with TruStain fcX™ (clone: 39, Biolegend) and subsequent staining with a panel of fluorescence-conjugated antibodies against CD4 (clone: RM4-5, BioLegend), CD44 (clone: IM7, BioLegend), CD62L (clone: MEL-14, BioLegend), CD24 (clone: M1/69, BioLegend), Qa2 (clone: 695H1-9-9, BioLegend), CD69 (clone: H1.2F3, BioLegend) and PD-1 (clone: RMP1-30, BioLegend). Stained cells were immediately sorted using a 5-laser Aria IIu SORP instrument (BD Biosciences) with the stringent gating strategy described in Fig. S8G.

RNA-seq library preparation and sequencing

Single Cell RNA-seq libraries were prepared using standard Fluidigm protocol (# PN 100-7168 K1) based on SMARTer chemistry and Illumina Nextera XT (Illumina) using paired-end 125bp sequencing on Illumina HiSeq2500. Each RNA-seq library was sequenced to a typical depth of 1.3 M reads on average.

To account for potential batch effects, for each experimental condition, two biological replicates were prepared using independent C1 IFCs.

Read alignment to reference genomes

Read alignment to reference genomes was performed using *gsnap* with default parameters, while supplying splice-site positions. Samples taken from B6 were mapped against the mouse reference GRCm38. CAST samples were aligned against the *Mus musculus castaneus de novo* genome assembly [ftp://ftp-mouse.sanger.ac.uk/REL-1509-Assembly/], which was used under an advance access agreement [ftp://ftp-mouse.sanger.ac.uk/REL-1509-Assembly/README]. Gene annotation for B6 was taken from the GRCm38 reference; gene annotation for CAST was taken from the newly constructed *Mus Castaneus* assembly [http://hgwdev.cse.ucsc.edu/~ifiddes/mouse_genomes_data/, version 0.2]. Additionally, since mitochondrial genes and certain immune genes (e.g., CD28) are absent from the most recent *Mus Castaneus* annotation, and since high mitochondrial gene expression is a well-established signature of low-quality single-cell transcriptome profiles (*24*), we also mapped CAST reads against GRCm38 and used the B6 annotation solely for these genes. Gene expression counts were obtained using HTSeq with default options. Only genes with orthologs in both

species were considered for downstream analysis. See section on “Divergence in expression between B6 and CAST” for further information.

Quality Control

Low-quality cells were filtered using the following quality control criteria, developed by previous single-cell RNA-sequencing studies (*25*) (*26*) (*27*) (*28*).

First, visual inspection of capture sites revealed empty wells and wells containing multiple cells, which were removed (Fig. S1A and S1B). Subsequently, the percentage of reads mapping to annotated genomic regions was compared to the percentage of reads mapping to ERCC spike-ins. Cells with low genomic read (< 20%) and/or high ERCC read (> 50%) numbers were excluded (Fig. S1C). Additionally, cells with few mapped reads (< 1,000,000) or in which > 3000 or < 1250 genes were detected were removed (Fig. S1D and S1E).

Next, cells with more than 10% or less than 0.5% of mitochondrial reads were excluded. Further, known markers of lymphocytes were used to filter cells - $CD19^{+}$/H2-Aa^{+} B cells as well as $CD8^{+}$ T cells were removed (Fig. S1H). Finally, in stimulated conditions, non-activated T cells were computationally removed (Fig. S1I).

Cell cycle analysis

Cell cycle states were established using *cyclone* (*27*) implemented in *scran* (https://bioconductor.org/packages/devel/bioc/html/scran.html). MACS purified CD4+ T cells were classified into G1, S and G2M phase, based on their normalized gene counts (Fig. S2A).

T Cell Receptor reconstruction

Recombined TCR sequences were reconstructed from single-cell RNA-seq data using the TraCeR algorithm as previously described (*29*). Briefly, reads originating from TCR mRNA were extracted from the sequence data for each cell and assembled into contigs representing the full-length TCR sequences expressed within each cell. Clonally-related cells were found by finding instances where more than one cell had identical TCR sequences (Fig. S2B).

Normalization

Read counts were normalized using the BASiCS package (*28*) incorporating spike-in reads for technical noise estimation. Prior to normalization, genes not expressed in at least 3 cells (rpm > 20) were filtered out. Similarly, ERCC spike-ins were removed if not detected in the data set. Posterior estimates were computed using a Markov chain Monte Carlo (MCMC) simulation with 40,000 iterations.

For cluster analysis, normalized and denoised expression rates were calculated using the *BASiCS_DenoisedRates* function. After variance decomposition, genes where less than 90% of the variability in expression levels across cells was attributed to biological factors (i.e., lowly-variable genes) were removed. For data visualization, normalized and denoised counts were computed using *BASiCS_DenoisedCounts* function.

Differential expression and differential variability analysis

We used BASiCS to detect differentially expressed and differentially variable genes. For changes in mean expression between populations, we first excluded genes that had an average posterior mean expression rate < 50 in each population. Subsequently, we applied the differential expression test developed within BASiCS, using a threshold of $\log_2$FC > 2 with an FDR cutoff of 0.05. Differential variability was tested on a mean $\log_2$FC = 0 and a variability $\log_2$FC > 0.4 at an FDR of 0.05. Only genes with no change in mean expression were considered for interpreting changes in variability.

Gene ontology enrichment analysis

We used DAVID (*30*) to test for gene ontology (GO) enrichment in differentially expressed genes. All tested genes in the conditions under study were used as background. Bonferroni corrected p-values (adjusted p-values) were used to visualize GO enrichment. The statistical significance threshold was set to adjusted p-value < 0.1.

Analysis of $CD4^+$ T cell subsets

Low quality cells were removed using established criteria (Fig. S1A-G). scRNA-sequencing identified cell-specific marker genes, and we removed residual B-cells, $CD8^+$ T cells, and (in activating conditions) non-activated T cells from our analyses (Fig. S1H, I). Even when activated, CD4+ T cells are in G1 phase of cell cycle as expected (Fig. S2A). Aged CD4+ T cells showed no clonal expansions (Fig. S2B) or differences in cell size (Fig. S2C) that could impact analysis of gene expression variability (*29*).

Upon T cell receptor (TCR) activation in the presence of particular cytokines, naive CD4+ T cells differentiate into several lineages of functionally different T helper cells (*22, 31*). We do not detect any early differentiation in naive and activated CD4+ T cell subsets and, as expected, we found Gata3 but not Th2 cytokines expressed in the majority of cells (*32*). Interestingly, the Th1-related genes Tbx21 and Ifng were up-regulated but not co-expressed, in a small population of activated CD4+ T cells of old animals. This is consistent with a known Th1 bias in CD4+ T cell responses in old mice (*33*) and humans (*34*)(Fig. S2H). Furthermore, we did not detect any difference in TCR components/signaling nor signs of T cell exhaustion (*35*), especially in cells isolated from old animals (Fig. S2I). We also ruled out species-specific differences in commitment towards T helper cell lineages (Fig. S2J).

Divergence in expression between B6 and CAST

After data quality control and quantification of gene expression levels, we used t-distributed stochastic neighborhood embedding (tSNE) to visualize cell clusters based on the denoised expression rates. Naive cells from young animals cluster based on the species from which they were extracted (Fig. 1B). To rule out the possibility that this clustering is driven by potential artifacts in the new *Mus musculus Castaneus* genome assembly, we also mapped reads from young CAST samples to the GRCm38 genome. All results presented were unaltered when the GRCm38 reference and annotation was used for both (Fig. S3).

Differential expression and variability analysis was performed using the BASiCS package implemented in R. Markov chain Monte Carlo (MCMC) simulations were used to infer posterior parameter distributions for all cells in the following conditions: naive, young B6; stimulated, young B6; naive, old B6; stimulated, old B6; naive, young CAST; stimulated, young CAST; naive, old CAST; stimulated, old CAST.

To estimate which differentially expressed genes may arise due to errors in the new CAST genome assembly, we performed the same differential expression analysis on CAST samples by mapping these experiments onto both GRCm38 and CAST. Roughly 5% of all tested genes are detected as differentially expressed even though the samples being compared are identical and only mapped to different genomes (Fig. S3C). Comparing this set of genes to the set of species-specific genes, we find that they make up 10% of differentially expressed genes between the two species, as expected (*36*). We performed a similar analysis for B6 samples. This approach allowed us to remove from our analyses the genes showing differences in expression, which may be driven by the quality of the reference genome.

Characterizing the activation process of $CD4^+$ T cells on a single cell level

Single cells extracted from young animals of B6 were activated *in vitro* for 3 hours with CD3ε/CD28 antibody stimulation. After QC and filtering, cells cluster based on their activation state using tSNE (Fig. 2A). To test whether this separation of cells into two groups is dictated by the use of tSNE, we also applied Principal Components Analysis (PCA) to the same data and observed the same result when projecting the cells onto the first two PCs (Fig. S4A).

We next sought to investigate the behavior of up- and down-regulated genes across the population of naive or activated cells. To this end, we calculated, for each gene, the percentage of cells in which it was expressed (> 0 transcript counts). Genes whose expression is down-regulated after activation are expressed in 18% of naive $CD4^+$ T cells isolated from B6, while genes that are up-regulated are expressed in 36% of activated cells (median fraction of cells in which gene expression was detected, 600 genes were randomly selected in each condition) (Fig. 2D). Before activation, up-regulated genes are only expressed on average in 5% of naive cells while after activation, down-regulated genes are expressed in 4% of activated cells (Fig. S4D). To investigate whether genes that are more variably expressed in the naive than the activated condition showed coordinated patterns of expression potentially associated with cryptic substructure, we identified genes with statistically higher expression variability in the naive population compared to activated cells (variability $\log_2 FC > 0.4$, FDR < 0.05, no change in mean expression). Genes with

high variability and high pairwise correlation (Spearman's rho > 0.8) in the naive population can be used to identify possible sub-populations of $CD4^+$ T cells. A hierarchical clustering analysis did not show any signs of this substructure (Fig. S4C).

Finally, it has previously been observed that covariance between cells due to unobserved factors such as the cell cycle can mask potentially interesting biological signals (*37*) In our dataset, ribosomal biogenesis is the strongest mediator of CD4+ T cell function upon activation. Importantly, controlling for this factor using scLVM (*26*) and re-performing the analysis on the corrected data did not reveal concealed cellular processes (Fig. S4F and S4G).

We performed the same analysis for $CD4^+$ T cells extracted from CAST and detect similar activation dynamics (Fig. S5).

Conservation of the activation program

To find genes that form the core activation program, we test for differential expression between the naive and activated state separately in B6 and CAST ($log_2FC > 2$, FDR < 0.05). Genes that are up-regulated after activation similarly in B6 and CAST form the shared activation program. Up-regulated genes that are differentially expressed in activated cells between the two species form the set of species-specific response genes; differential expression analysis was performed on genes up-regulated in either B6 or CAST ($log_2FC > 2$, FDR < 0.05). To identify the set of species-specific response genes that might be caused by mapping artifacts, we quantified gene expression in activated CAST and B6 samples based on the CAST and GRCm38 genome as described above. Genes in the activated state that show differential mapping between the two genomes were excluded from the species-specific lists of response genes. We find 225 genes strongly up-regulated similarly in B6 and CAST while 1208 genes in total are up-regulated across both species. Out of these genes, 171 are detected as differentially expressed between the two species forming the set of species-specific genes. For visualization purposes, we randomly selected 100 out of the 225 shared response genes, 43 of the 96 B6 specific and 33 of the 75 CAST specific response genes (Fig. 3B).

Aging does not show a conserved effect on gene expression levels

For each species, we identified differentially expressed genes between young and old animals separately for naive and activated $CD4^+$ T cells (Fig. S6C and S6D). For these comparisons, we computed the fraction of cells in which each differentially expressed gene was expressed. The distribution of these values was added as inlets to the plots (x-axis ranging from 0% to 100% of cells).

To test whether similar genes are differentially expressed during aging in both species, we calculated the Jaccard index, which measures the overlap between sets of elements, separately for up- and down-regulated genes (Fig. S6E and S6F).

Variability analysis on the conserved activation program during aging

Focusing on the conserved activation program in activated $CD4^+$ T cells, we calculated the change in mean expression of activated cells between young and old animals (Fig. 4A). Only genes with no changes in mean expression between young and old animals were considered when interpreting changes in transcriptional variability (Fig. 4B).

As described above, we calculated the fraction of cells in which genes in the shared activation program are expressed and compared these fractions between activated cells isolated from old or young animals in both species. We used a binomial test to evaluate which genes significantly change the percentage of cells in which they are expressed (Bonferroni corrected p-values < 0.1). Gene expression in activated cells isolated from old animals was used as the Null-distribution (Fig. 4C).

To test whether the global shift in variability is caused by different cell population structures between old and young B6 animals, we removed cells expressing marker genes that were inconsistent with their activation state as well as cells with a possible Th1 differentiation bias (*Ifng* expressing). Based on the library size adjusted counts, we removed activated cells with low *Cd69* (< 300 counts), high *Sell* (> 10 counts), low *Trac* (< 100 counts), low *Il2ra* (< 100 counts) and *Ifng* (> 0 counts) expression (Fig. S7B). The remaining 37 and 26 activated cells in young and old B6 animals showed the same shift in transcriptional variability compared to the non-filtered data (Fig. S7C).

Thymic output and maturation of recent thymic emigrants (RTEs) contributes significantly to the maintenance of the naive $CD4^+$ T cell pool in the periphery and is affected by aging (*38-40*). To exclude different proportions of RTEs within the naive $CD4^+$ T cell pool as a confounding factor, we characterized CD4 single positive (SP) thymocytes and splenic naive $CD4^+$ T cell by flow cytometry using CD24 and Qa2 – a marker combination used for identification of RTEs in peripheral tissues (*38, 39*). The majority of CD4 SP thymocytes showed a $CD24^{hi}$ $Qa2^{lo}$ phenotype (Fig. S8A). In contrast, we detected only a very small population of RTEs ($CD24^{hi}$ $Qa2^{lo}$, ~2%) within splenic naive $CD4^+$ T cell pool of young and old mice (Fig. S8A and S8B) (*39*), suggesting minimal contamination of RTEs in naive $CD4^+$ T cells purified by MACS.

Accordingly, using a stringent gating strategy (Fig. S8D) we isolated pure naïve, resting $CD4^+$ T cells by FACS (thereby excluding RTEs). Upon analyzing these cells using the scRNA-seq experimental and analysis strategy applied to the MACS-purified naïve population of cells, we observed an increase in transcriptional heterogeneity during aging as described in the main text (Fig. S9A).

Next, we examined whether the age-mediated increase in cell-to-cell variability is conserved across different memory subsets of $CD4^+$ T cells. As expected, a significantly higher proportion of effector memory $CD4^+$ T cells in old animals expressed the activation markers CD69 and PD-1, indicating a larger fraction of cells is already in an activated state (Fig. S8C). To avoid interference in cell-to-cell variability by endogenously activated cells, cells stained positive for CD69 and/or PD-1 were excluded during the sort of effector memory $CD4^+$ T cells by FACS. Similarly to naive $CD4^+$ T cells, effector memory $CD4^+$ T cells increase cell-to-cell transcriptional variability during aging (Fig. S9B)

SUPPLEMENTARY FIGURES

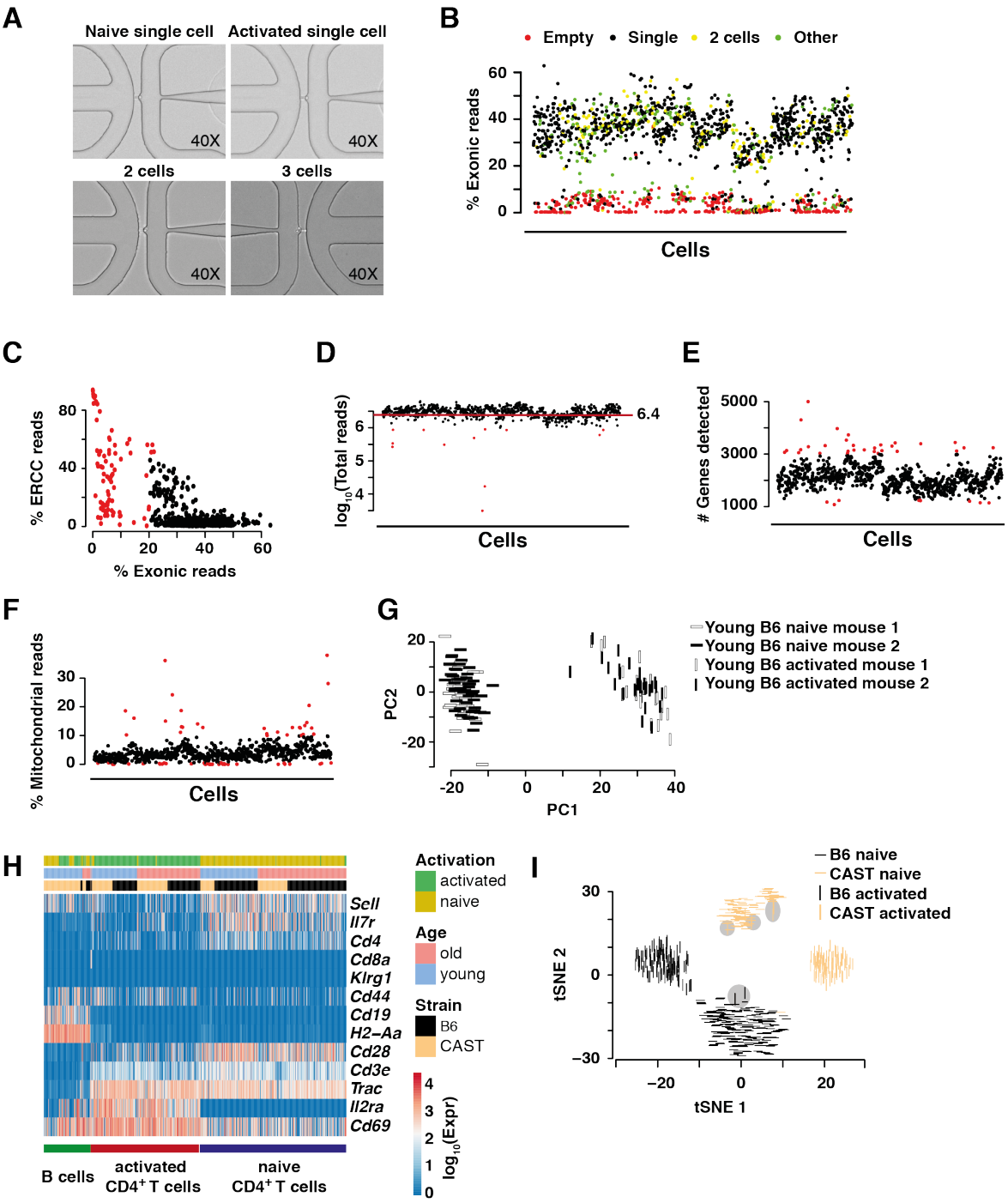

Fig. S1.
Quality control and filtering of raw read counts.
(A) Visual inspection of captured cells at 40x magnification in Integrated Fluidic Circuit (C1, Fluidigm) allows manual removal of empty capture sites, and capture sites holding multiple cells or debris;

(B) Percentage of reads mapping to exonic regions displayed for naive and activated $CD4^+$ T cells. Black dots: single cells; yellow dots: 2 cells, red dots: empty wells, green dots: debris, multiple cells, etc.;
(C) Removal of cells with less than 20% of mapped exonic reads and more than 50% of ERCC spike-in reads (red dots);
(D) Cells with less than 1 million mapped reads were excluded from downstream analysis (red dots);
(E) Cells with more than 3000 or less than 1250 genes detected were excluded in the analysis (red dots);
(F) Cells with more than 10% or less than 0.5% of mitochondrial reads were excluded for downstream analysis (red dots);
(G) Naive and activated cells isolated from young B6 animals (replicates) were colored batch-specifically. 4 batches from 2 mice: naive and activated from mouse 1 (white bars), naive and activated from mouse 2 (black bars). Naive condition is represented in horizontal bars and activated condition in vertical bars;
(H) Data set was filtered for immune markers to exclude B cell and $CD8^+$ T cell contamination. Cells in columns were labeled based on their activation state (naive in beige, activated in green), their age (old in red, young in blue), and the strain of the animals (B6 in black, CAST in yellow);
(I) Unbiased tSNE clustering allowed the removal of not fully activated cells (indicated in grey circles). Cells were labeled based on their activation state (naive: horizontal bar, activated: vertical bar) and the strain of the animals (B6 in black, CAST in yellow).

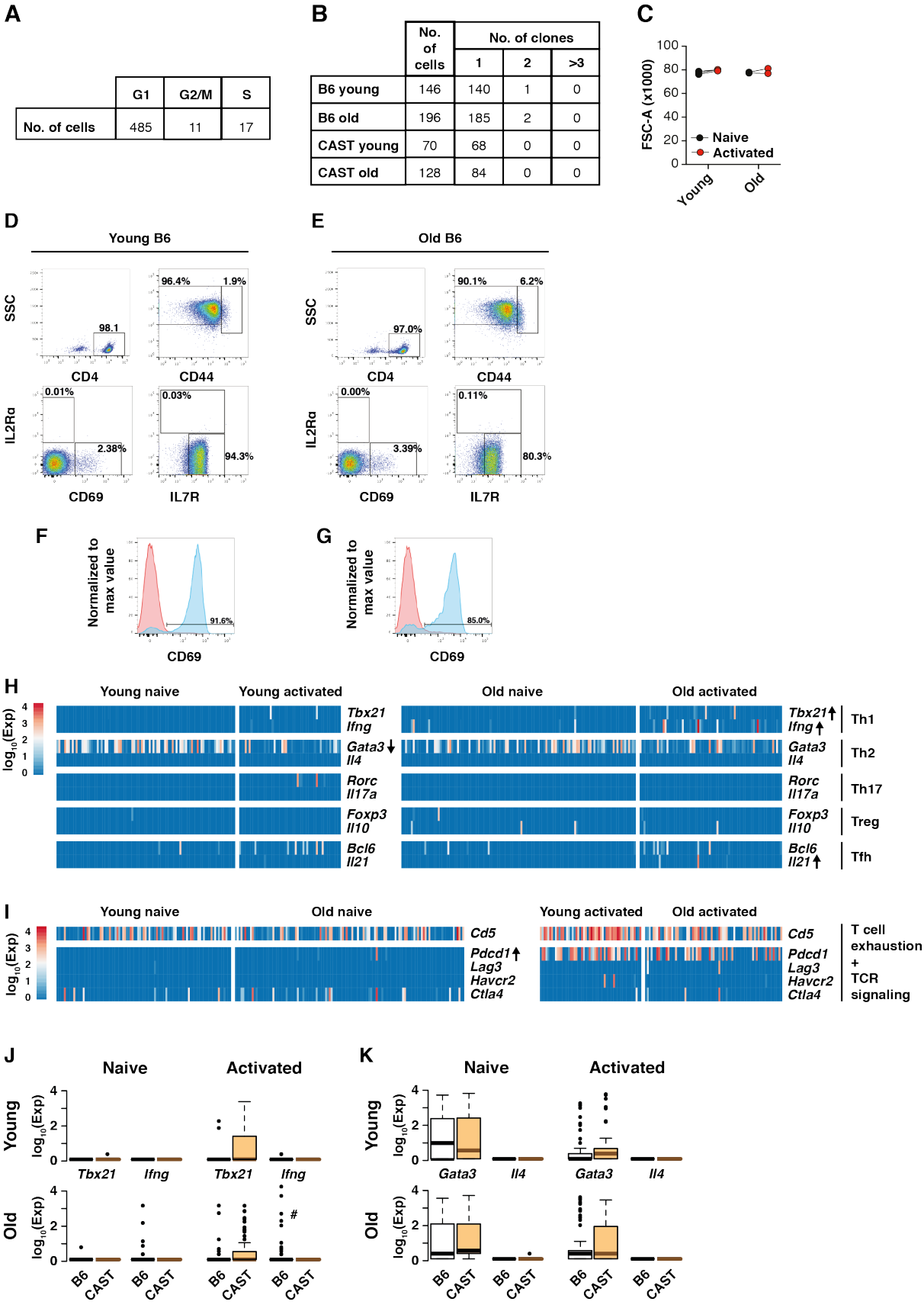

Fig. S2. Characterization of isolated CD4$^+$ T cells.

Fig. S2. Characterization of isolated CD4$^+$ T cells.
(A) Cyclone was used to classify individual naive and activated CD4$^+$ T cells into the cell cycle phases G1, G2/M and S;
(B) TraCeR constructed T cell receptor sequences from scRNA-seq data to analyze clonal diversity in naive and activated CD4$^+$ T cells;
(C) Cell sizes were estimated for naive and activated CD4$^+$ T cells in young (n=3) and old (n=2) B6 animals measured by forward scatter (FSC-A) using flow cytometry. Lines connect cells isolated from the same mouse in both conditions;
(D) CD4$^+$ T cells were purified from spleens of young B6 mice and stained with antibodies against CD4, CD62L, CD44, CD69, IL2Rα (CD25), CD127, and KLRG1 as well as viability dye. FACS plots shown are gated on single live cells (top left panel) and single live CD4$^+$ T cells (other panels), and percentages shown relate to total of gated cells. A representative FACS plot from one of five young mice is shown;
(E) CD4$^+$ T cells were purified from spleens of aged B6 mice and stained with antibodies against CD4, CD62L, CD44, CD69, IL2Rα (CD25), CD127, and KLRG1 as well as viability dye. FACS plots shown are gated on single live cells (top left panel) and single live CD4$^+$ cells (other panels) and % shown relate to total of gated cells. A representative FACS plot from one of two aged mice is shown;
(F) Naive CD4$^+$ T cells were purified from spleens of five young B6 mice, and were left naive or were activated with plate-bound antibody against CD3ε and CD28 for 3 hours. Cells were stained with antibodies against CD4, CD69, and viability dye. Representative histograms are shown and were gated on single live CD4$^+$ T cells: naive (red) and activated (blue) cells are shown;
(G) CD4$^+$ T cells were purified from spleens of two aged B6 mice and were left naive or were activated with plate-bound antibody against CD3ε and CD28 for 3 hours. Cells were stained with antibodies against CD4, CD69, and viability dye. Representative histograms are shown and were gated on single live CD4$^+$ T cells: naive (red) and activated (blue) cells are shown;
(H) Characterization of possible differentiation processes leading to T helper 1 (Th1), T helper 2 (Th2), T helper 17 (Th17), regulatory T (Treg) and follicular helper T (Tfh) cell lineages. For each lineage the major regulatory transcription factor (upper row) and an effector cytokine (lower row) is shown. Statistical differential expression testing was performed between activated and naive cells from young B6 animals (left panel) and between activated and naive cells from old B6 animals (right panel). Upward arrow: up-regulation of expression ($\log_2$FC > 2, FDR < 0.05) after activation, Downward arrow: down-regulation of expression ($\log_2$FC > 2, FDR < 0.05) after activation;
(I) Heatmap showing T cell exhaustion (*Pdcd1, Lag3, Havcr2, Ctla4*) and TCR activation markers (*Cd5*). Statistical differential expression testing was performed in naive cells between young and old B6 animals (left panel) and in activated cells between young and old B6 animals (right panel). Upward arrow: up-regulation of expression ($\log_2$FC > 2, FDR < 0.05) during aging;
(J) Th1 lineage marker (*Tbx21, Ifng*) expression was compared between B6 and CAST in following conditions: naive cells from young animals (upper left panel), naive cells from old animals (lower left panel), activated cells from young animals (upper right panel), activated cells from old animals (lower right panel). #: statistically significant differential expression ($\log_2$FC > 2, FDR < 0.05);
(K) Th2 lineage marker (*Gata3, Il4*) expression was compared between B6 and CAST in following conditions: naive cells from young animals (upper left panel), naive cells from old animals (lower left panel), activated cells from young animals (upper right panel), activated cells from old animals (lower right panel).

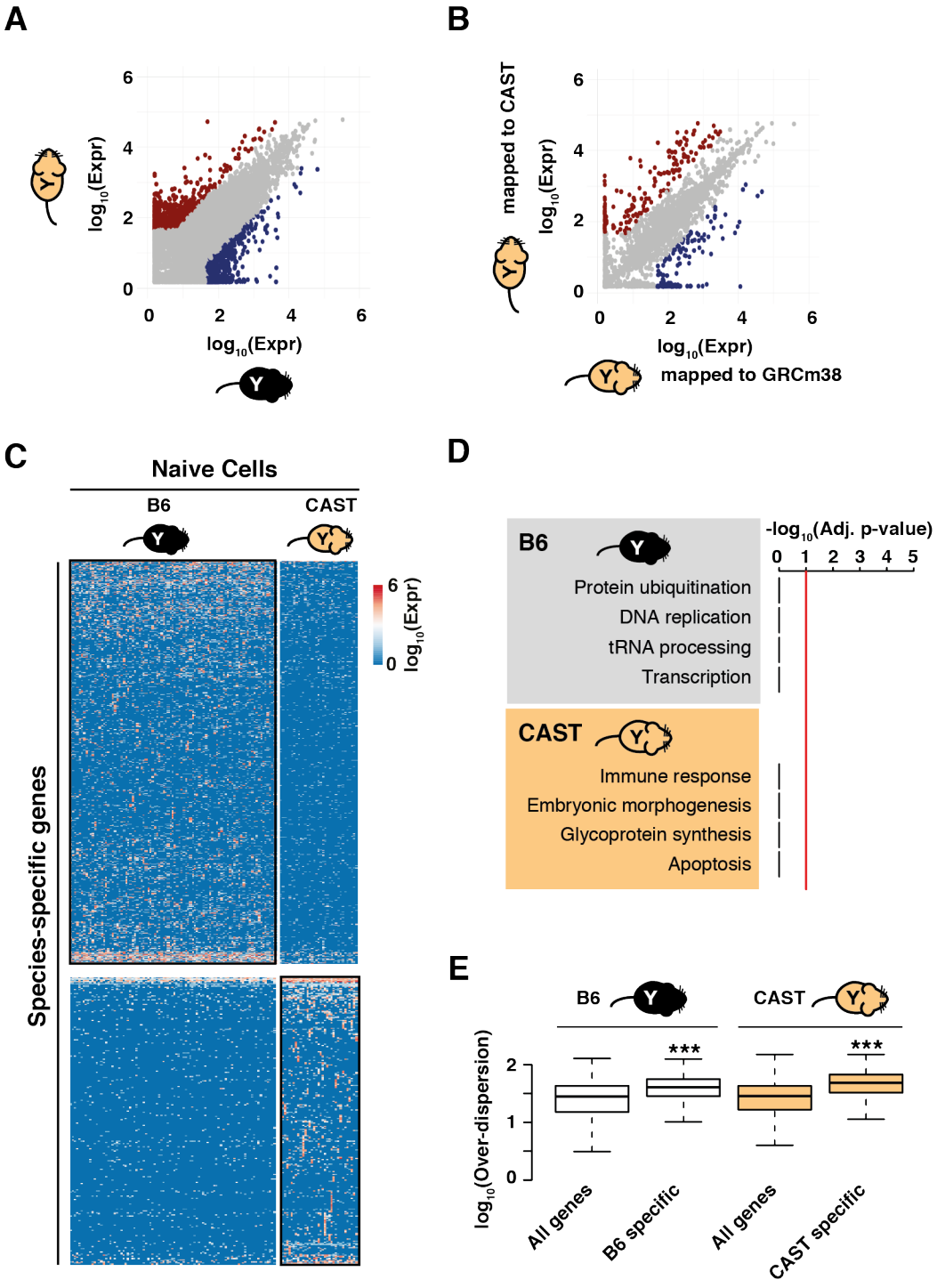

Fig. S3.
Characterization of species-specific expression profiles.
(A) Species-specific gene expression in B6 (blue) or CAST (red). Average gene expression using posterior estimation, threshold of means > 50, $log_2FC > 2$, FDR < 0.05;
(B) Mean normalized transcript counts of mapped reads from CAST cells (young, naive) using the GRCm38 genome (x-axis) or the CAST genome (y-axis) as reference. Differentially mapped genes were removed from downstream analysis.

Average gene expression using posterior estimation, threshold of means > 50, $log_2FC > 2$, FDR < 0.05;
(C) Complete heatmap showing all species-specific genes expressed in all naive $CD4^+$ T cells isolated from young B6 and CAST;
(D) Bar plots of functional gene categories detected in species-specific gene lists (Bonferroni multiple testing corrected p-values, red line marking 0.1; Material and Methods);
(E) Variability of species-specific genes was compared to variability of all genes expressed in B6 (left) and CAST (right). Mann-Whitney-Wilcoxon test; $p<10^{-10}$.

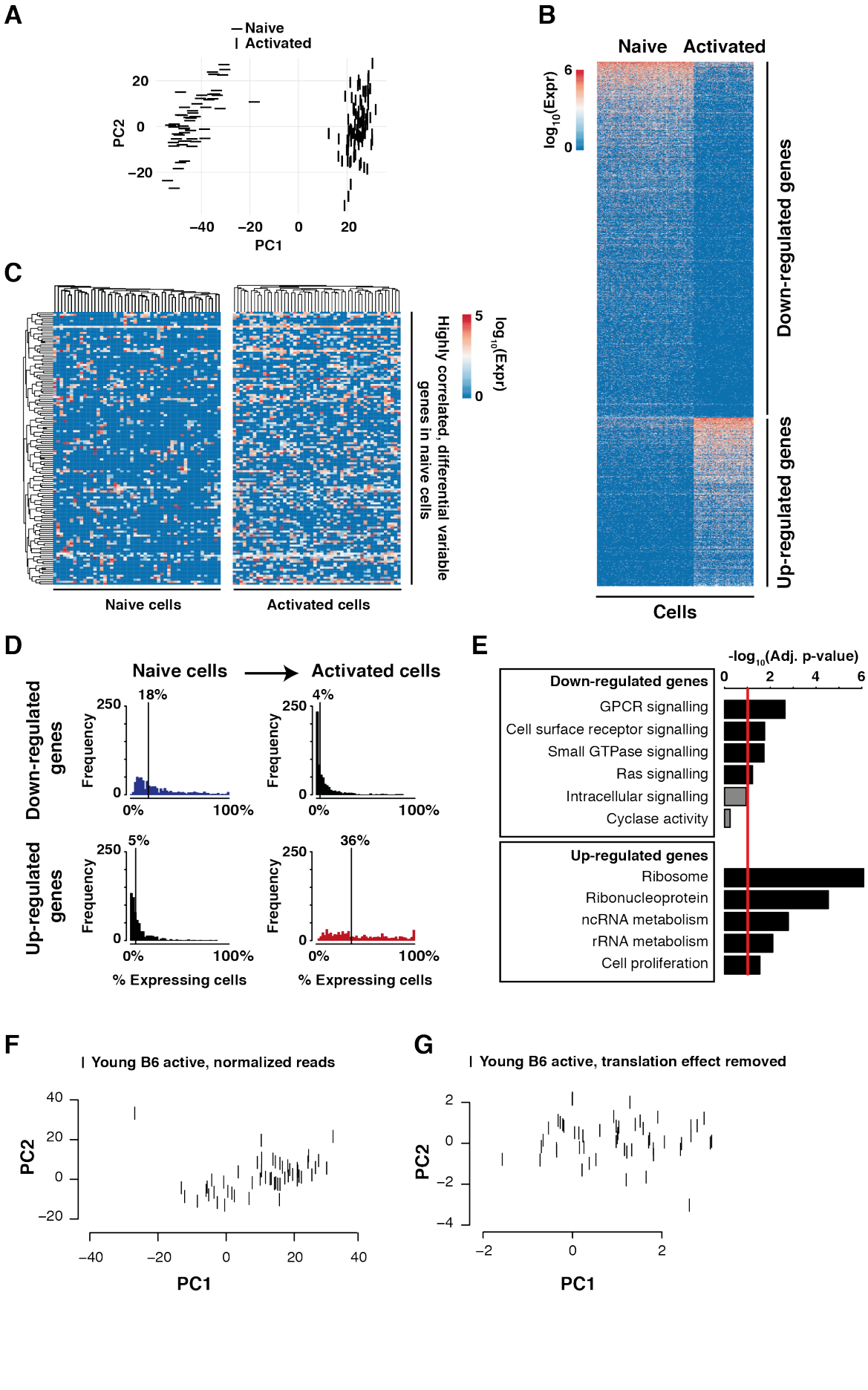
A
Naive
Activated
PC2
PC1
B
Naive
Activated
log10(Expr)
Down-regulated genes
Up-regulated genes
Cells
C
Highly correlated, differential variable genes in naive cells
log10(Expr)
Naive cells
Activated cells
D
Naive cells
Activated cells
Down-regulated genes
Up-regulated genes
Frequency
% Expressing cells
E
-log10(Adj. p-value)
Down-regulated genes
GPCR signalling
Cell surface receptor signalling
Small GTPase signalling
Ras signalling
Intracellular signalling
Cyclase activity
Up-regulated genes
Ribosome
Ribonucleoprotein
ncRNA metabolism
rRNA metabolism
Cell proliferation
F
Young B6 active, normalized reads
PC2
PC1
G
Young B6 active, translation effect removed
PC2
PC1

Fig. S4.
Characterization of the early activation program in B6.
(A) Principal component analysis (PCA) separates naive and activated $CD4^+$ T cells along the first PC;
(B) Full heatmap showing genes down- and up-regulated after immune stimulation in naive and activated $CD4^+$ T cells;
(C) To detect possible sub-populations in naive or activated $CD4^+$ T cells, differentially variable genes ($log_2FC > 0.4$, $FDR < 0.05$) in naive cells with high gene-to-gene correlation (Spearman's rho > 0.8) were used for hierarchical cluster analysis;
(D) Fractions of cells in which down- or up-regulated genes are expressed (600 genes were randomly selected, histograms with 50 bins, median value is indicated). Upper panels: fraction of naive (left) and activated (right) cells in which down-regulated genes are expressed. Lower panels: fraction of naive (left) and activated (right) cells in which up-regulated genes are expressed;
(E) Bar plots of functional gene categories enriched in up- and down-regulated genes during activation of $CD4^+$ T cells in B6 (Bonferroni multiple testing corrected p-values, red line marking 0.1; Material and Methods);
(F) PCA of activated $CD4^+$ T cells of young B6 animals;
(G) PCA of activated $CD4^+$ T cells of young B6 animals after removing differences in the translation program as confounding factor.

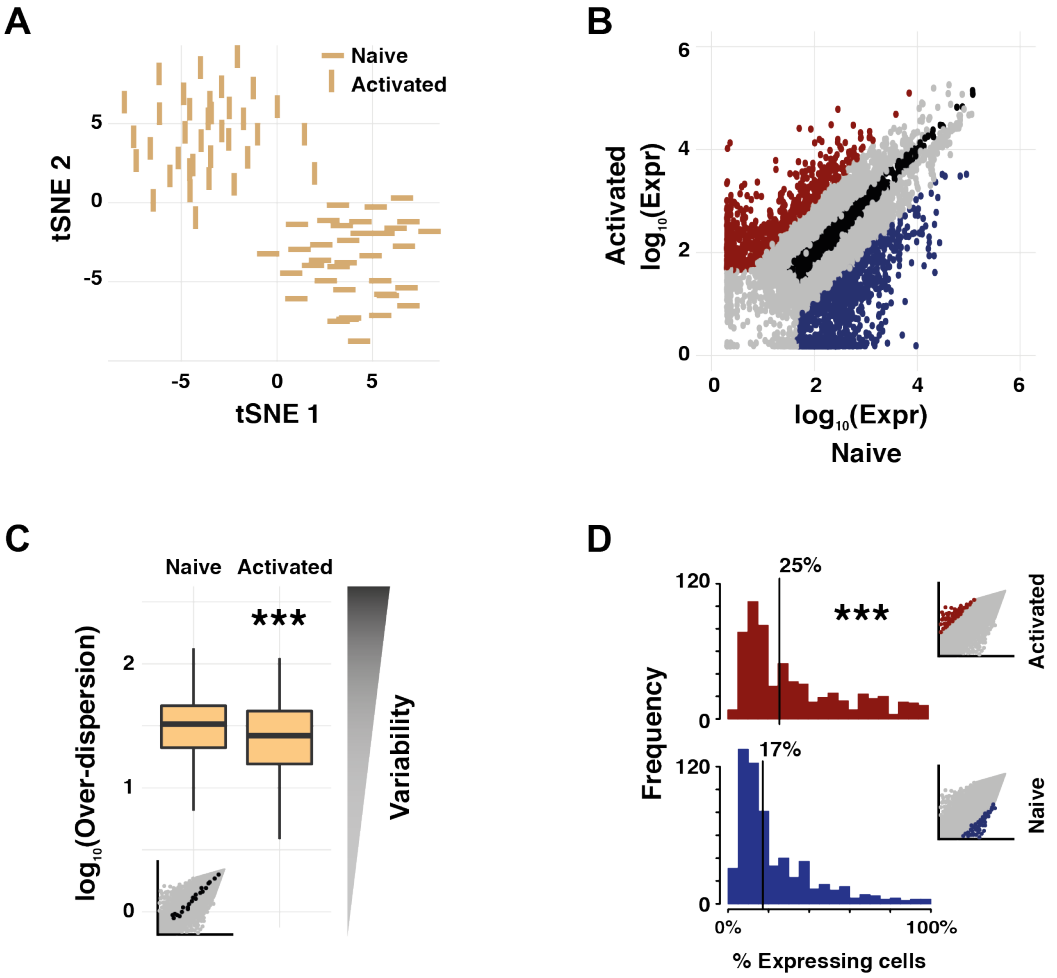

Fig. S5.
Activation of CD4$^+$ T cells induces a transcriptomic switch from stochastic to regulated gene expression in CAST.
(A) Activation of CD4$^+$ T cells from young CAST mice in anti-CD3ε/CD28 coated plates induces large-scale transcriptional changes, visualized using tSNE dimensionality reduction;
(B) Genes up-regulated (red) and down-regulated (blue) by immune stimulation in young CAST mice. Non-differentially expressed genes used in (C) are also shown (black). Average gene expression using posterior estimation, threshold of means > 50, $\log_2$FC > 2, FDR < 0.05;
(C) Genes with no overall gene expression differences during activation show decreased cell-to-cell variability in transcription (Mann-Whitney-Wilcoxon test, ***: $p<10^{-10}$);
(D) Up-regulated genes were expressed in a relatively large fraction of activated CD4$^+$ T cells after stimulation (median 25%); down-regulated genes were expressed in a smaller fraction of naive CD4$^+$ T cells (median 17%). 600 genes of each condition were randomly selected.

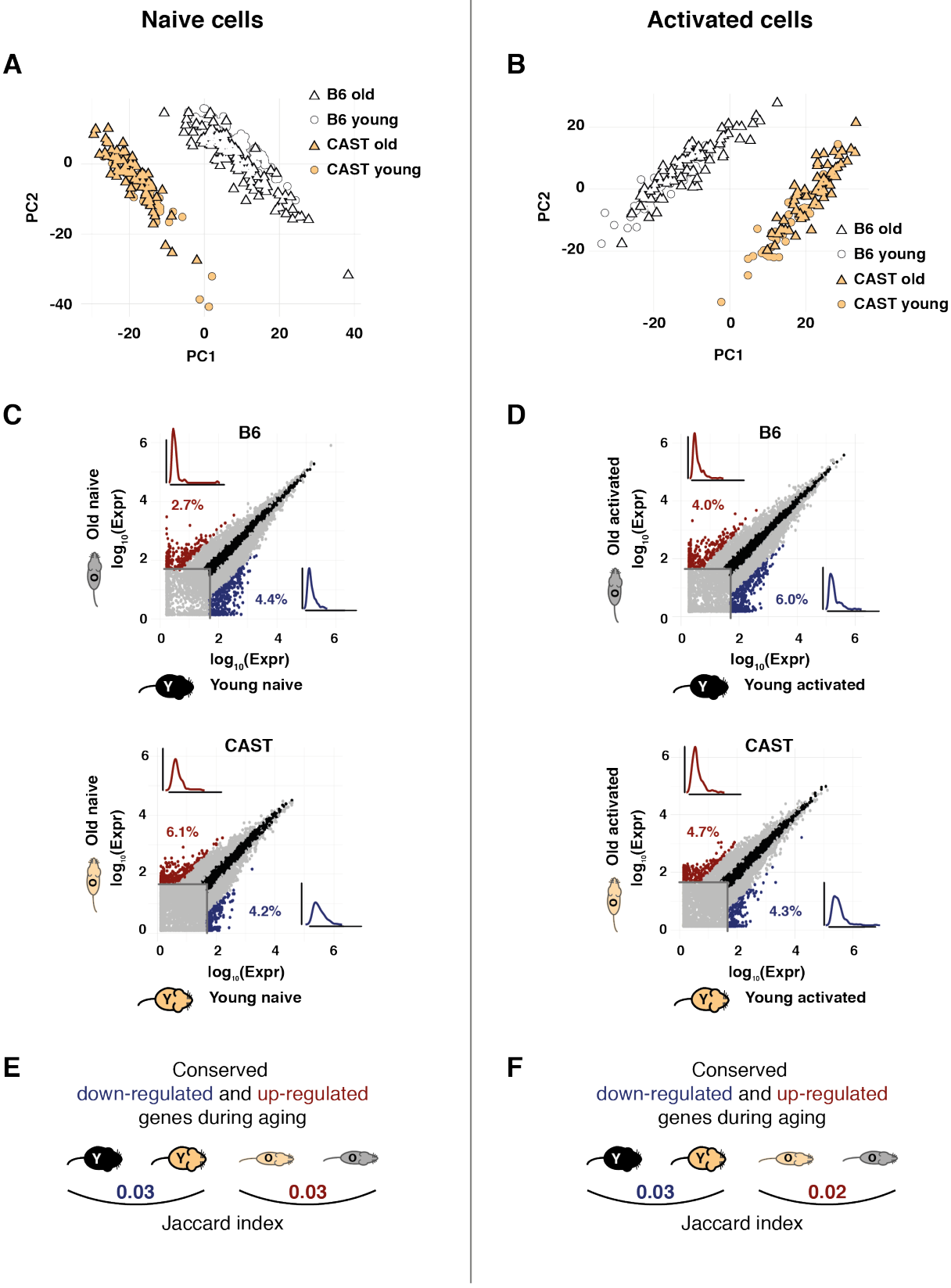
Naive cells
Activated cells
A
B
B6 old
B6 young
CAST old
CAST young
PC1
PC2
C
D
B6
CAST
Old naive
Old activated
log10(Expr)
2.7%
4.4%
6.1%
4.2%
4.0%
6.0%
4.7%
4.3%
Young naive
Young activated
E
F
Conserved
down-regulated and up-regulated
genes during aging
0.03
0.03
0.02
Jaccard index

Fig. S6.
Aging did not change the global expression profile of naive or activated $CD4^+$ T cells.
Principal component analysis reveals no separation between cells isolated from young or old animals in the naive **(A)** or activated **(B)** state;
(C) 7.1% of all tested genes in B6 and 10.3% of genes in CAST are differentially expressed in naive $CD4^+$ T cells between old (red) and young (blue) animals. Average gene expression using posterior estimation, threshold of means > 50, $\log_2$FC > 2, FDR < 0.05. Insets show distributions of fraction of cells in which these genes are expressed. X-axis: 0% - 100% of cells;
(D) 10% of all tested genes in B6 and 9% of genes in CAST are differentially expressed in activated $CD4^+$ T cells between old (red) and young (blue) animals. Average gene expression using posterior estimation, threshold of means > 50, $\log_2$FC > 2, FDR < 0.05. Insets show distributions of fraction of cells in which these genes are expressed. X-axis: 0% - 100% of cells;
(E) The overlap of aging-associated genes in naive cells was calculated using the Jaccard index between gene sets. Genes highly expressed in old animals (red) or genes highly expressed in young animals (blue) show little overlap (3%) between B6 and CAST;
(F) The overlap of aging-associated genes in activated cells was calculated using the Jaccard index between gene sets. Genes highly expressed in old animals (red) or genes highly expressed in young animals (blue) show little overlap (2-3%) between B6 and CAST.

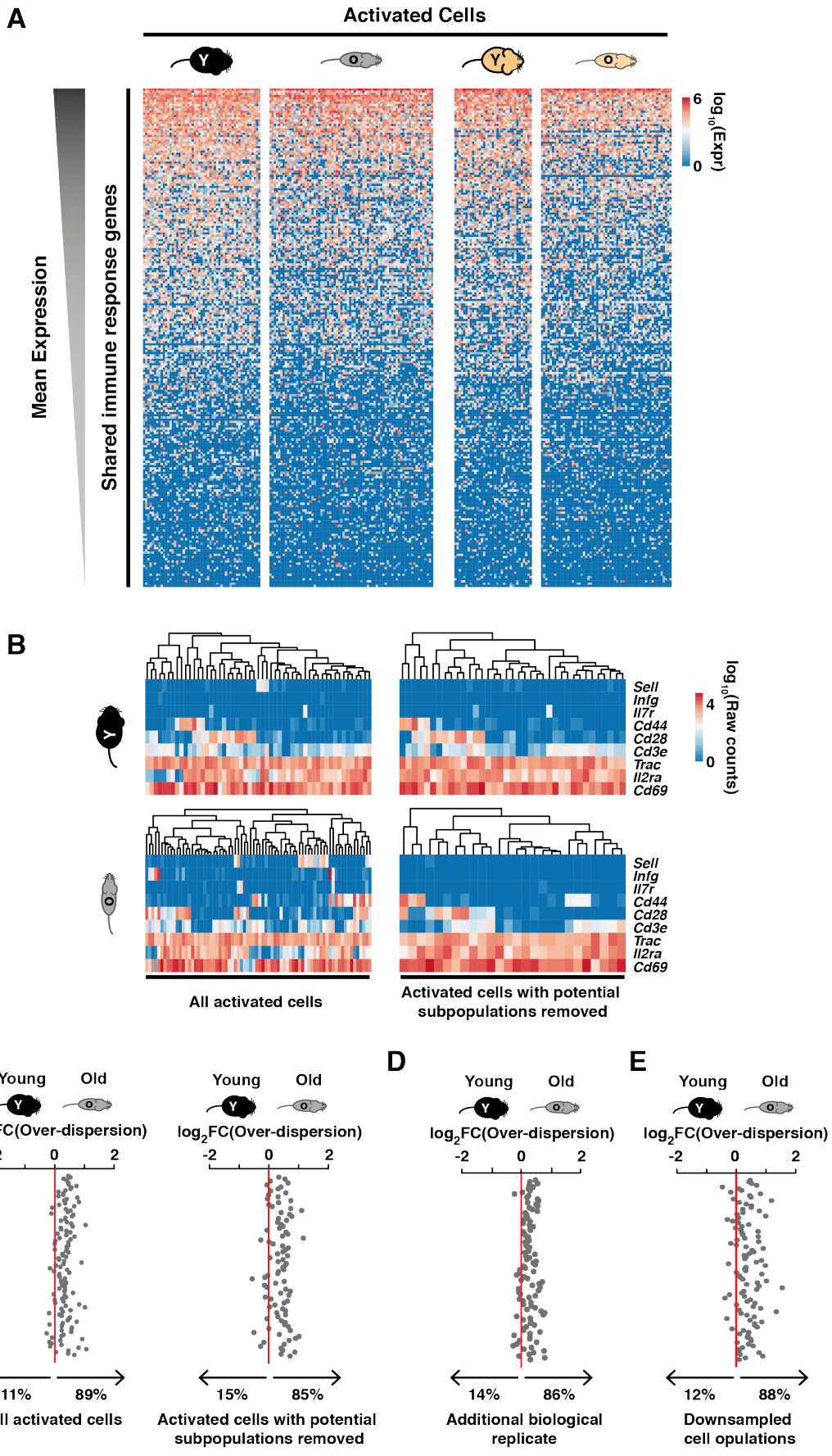
A
Activated Cells
Mean Expression
Shared immune response genes
log10(Expr)
6
0
B
Sell
Infg
Il7r
Cd44
Cd28
Cd3e
Trac
Il2ra
Cd69
log10(Raw counts)
4
0
All activated cells
Activated cells with potential subpopulations removed
C
Young
Old
log2FC(Over-dispersion)
-2
0
2
11%
89%
All activated cells
15%
85%
Activated cells with potential subpopulations removed
D
14%
86%
Additional biological replicate
E
12%
88%
Downsampled cell opulations

Fig. S7.
Aging destabilizes the shared activation program.
(A) Full heatmap showing all 225 genes of the shared activation program expressed in all activated cells from young and old CAST and B6. Genes were ordered based on their mean expression;
(B) Expression of $CD4^+$ T cell activation markers in all activated cells from young and old B6 before (left) and after (right) filtering based on *Ifng*, *Sell*, *Trac*, *Il2ra* and *Cd69* expression (Material and Methods);
(C) Changes in gene expression variability of activated cells between young and old animals are displayed before (left) and after (right) filtering based on *Ifng*, *Sell*, *Trac*, *Il2ra* and *Cd69* expression (Material and Methods);
(D) A biological replicate of 115 activated $CD4^+$ T cells from old B6 animals was generated and changes in gene expression variability were compared to activated $CD4^+$ T cells from young B6 animals;
(E) 30 out of all activated $CD4^+$ T cells from young or old B6 mice were randomly selected and changes in gene expression variability were compared between the downsampled populations of cells.

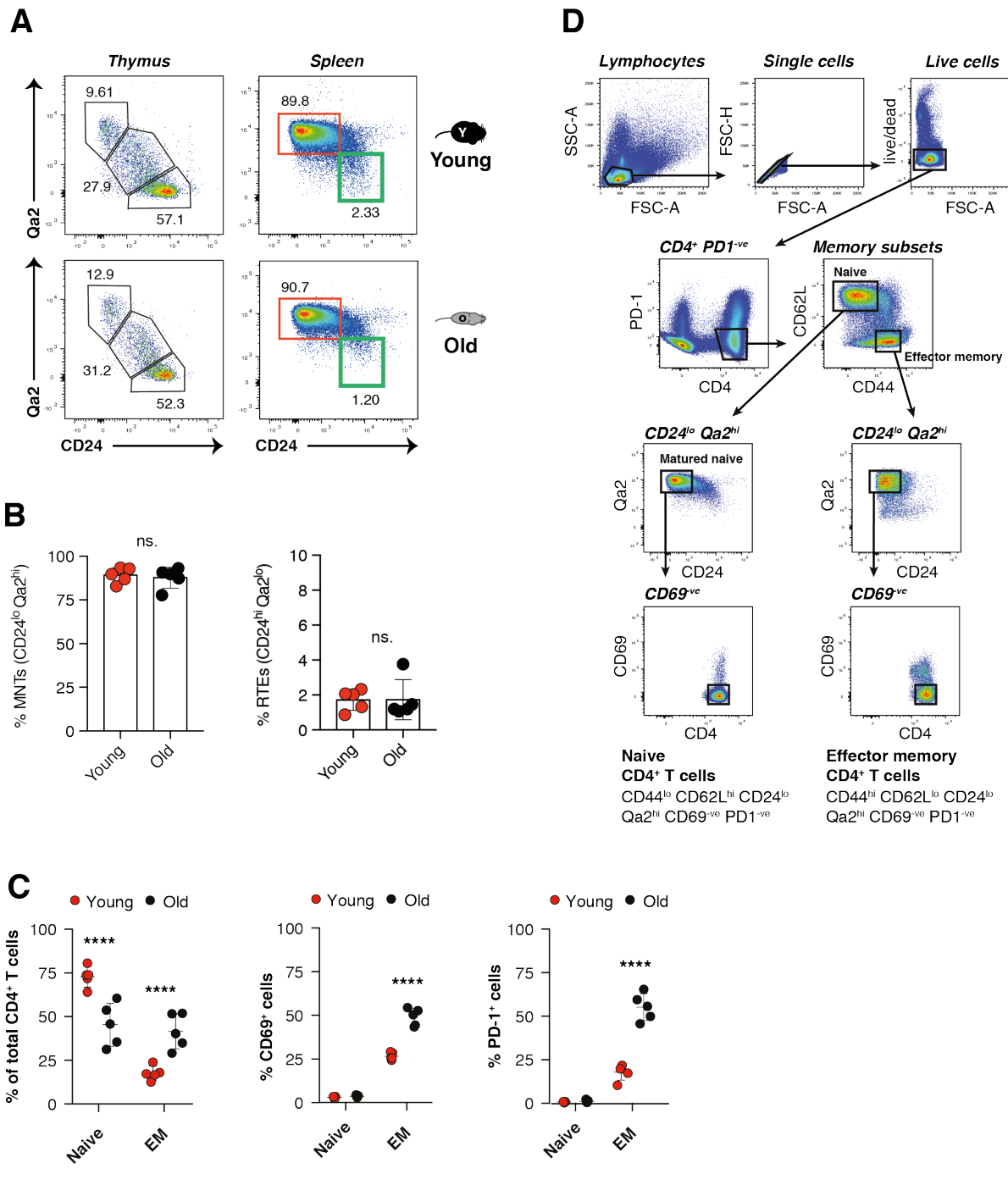

Fig. S8
Sorting strategy for highly pure naive and effector memory CD4$^+$ T cell populations.
(A) Thymus and spleen were collected from young and old B6 mice, dissociated into single cell suspensions, stained with viability dye and antibodies against CD4, CD8, CD24 and Qa2 for thymocytes or antibodies against CD4, CD44, CD62L, CD24 and Qa2 for splenocytes. FACS plots shown are gated on single, live cells and either CD4$^+$ CD8$^-$ CD4 single positive (SP) thymocytes (thymus) or CD44^{lo} CD62L^{hi} naive CD4$^+$ T cells (spleen). Percentages relate to total gated cells. Representative FACS plots of 5 young and 5 old mice are shown. In the thymus the majority of CD4 SP express markers of recent thymic emigrants (RTEs, CD24^{hi} Qa2^{lo}) while in the spleen naive CD4$^+$ T cells are comprised mainly of mature naive cells (red box).

(B) Analysis of flow cytometry data from (A). Significance of difference was calculated by Mann-Whitney test (ns = not significant).
(C) Spleen were collected from young and old B6 mice, dissociated into single cell suspensions, stained with viability dye and antibodies against CD4, CD44, CD62L, CD24, Qa2, CD69, and PD-1. FACS plots were gated on single, live, $CD4^+$ T cells and subsequently on either $CD44^{lo}$ $CD62L^{hi}$ (Naive) or $CD44^{hi}$ $CD62L^{lo}$ (Effector Memory, EM) subsets. Expression of CD69 and PD-1 was analyzed in these subsets. Results shown are pooled from 10 independent experiments with spleens harvested from 5 young and 5 old mice. Significance of difference was calculated by two-way ANOVA (**$p \leq 0.01$; ****$p \leq 0.0001$)
(D) Gating Strategy: lymphocytes were gated by the use of forward scatter (FSC-A) and side scatter (SSC-A). Cell doublets were excluded according to area and height of forward scatter (FSC-A/FSC-H). Dead cells were removed using viability dye. $PD\text{-}1^+$ $CD4^+$ T cells were excluded and $PD\text{-}1^{-ve}$ $CD4^+$ T cells were t further separated into naive and EM $CD4^+$ T cell subsets according to their CD44 and CD62L expression. Cells with a mature $CD24^{lo}$ $Qa2^{hi}$ phenotype were then gated from naive and EM subsets and $CD69^+$ cells were removed.

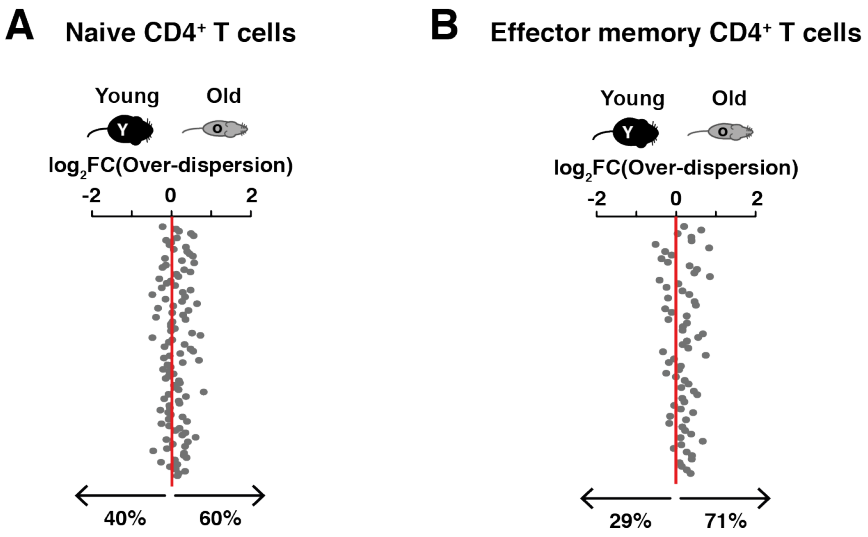

Fig. S9
Changes in expression variability in naive and effector memory cells between young and old animals after activation.
(A) Activated, FACS-purified naive CD4+ T cells from old animals showed higher transcriptional variability compared to young animals;
(B) Activated, FACS-purified effector memory CD4+ T cells from old animals showed higher transcriptional variability compared to young animals.

Table S1. (Uploaded xls file)
List of species-specific genes in B6 (sheet 1) and CAST (sheet 2). Average mean expression for each gene in B6 or CAST is quantified. Differential expression was inferred using the posterior distributions of gene expression. Threshold of means > 50, $\log_2$FC > 2, FDR < 0.05.

Table S2. (Uploaded xls file)
List of down-regulated (sheet 1) and up-regulated (sheet 2) genes upon activation of CD4$^+$ T cells isolated from B6. Differential expression was inferred using the posterior distributions of gene expression. Threshold of means > 50, $\log_2$FC > 2, FDR < 0.05. BASiCS test parameters are indicated for each differentially expressed gene.

Table S3. (Uploaded xls file)
List of shared (sheet 1), B6-specific (sheet 2), and CAST-specific (sheet 3) activation genes. Differential expression was tested between naive and activated cells to detect shared genes, and in activated cells between B6 and CAST to find species-specific genes. Threshold of means > 50, $\log_2$FC > 2, FDR < 0.05. Average expression levels are quantified for the following conditions: B6 naive, CAST naive, B6 activated and CAST activated.

REFERENCES

22. J. Zhu, H. Yamane, W. E. Paul, Differentiation of effector CD4 T cell populations. *Annu Rev Immunol* **28**, 445-489 (2010).
23. D. Ramskold *et al.*, Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol* **30**, 777-782 (2012).
24. T. Ilicic *et al.*, Classification of low quality cells from single-cell RNA-seq data. *Genome Biol* **17**, 29 (2016).
25. P. Brennecke *et al.*, Accounting for technical noise in single-cell RNA-seq experiments. *Nat Methods* **10**, 1093-1095 (2013).
26. F. Buettner *et al.*, Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat Biotechnol* **33**, 155-160 (2015).
27. A. Scialdone *et al.*, Computational assignment of cell-cycle stage from single-cell transcriptome data. *Methods* **85**, 54-61 (2015).
28. C. A. Vallejos, J. C. Marioni, S. Richardson, BASiCS: Bayesian Analysis of Single-Cell Sequencing Data. *PLoS Comput Biol* **11**, e1004333 (2015).
29. M. J. Stubbington *et al.*, T cell fate and clonality inference from single-cell transcriptomes. *Nat Methods*, (2016).
30. D. W. Huang, B. T. Sherman, R. A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57 (2009).
31. S. L. Swain, K. K. McKinstry, T. M. Strutt, Expanding roles for CD4(+) T cells in immunity to viruses. *Nat Rev Immunol* **12**, 136-148 (2012).
32. I. C. Ho, T. S. Tai, S. Y. Pai, GATA3 and the T-cell lineage: essential functions before and after T-helper-2-cell differentiation. *Nat Rev Immunol* **9**, 125-135 (2009).
33. W. Zhang, V. Brahmakshatriya, S. L. Swain, CD4 T cell defects in the aged: causes, consequences and strategies to circumvent. *Exp Gerontol* **54**, 67-70 (2014).
34. S. Sakata-Kaneko, Y. Wakatsuki, Y. Matsunaga, T. Usui, T. Kita, Altered Th1/Th2 commitment in human CD4+ T cells with ageing. *Clin Exp Immunol* **120**, 267-273 (2000).
35. E. J. Wherry, T cell exhaustion. *Nat Immunol* **12**, 492-499 (2011).
36. A. Goncalves *et al.*, Extensive compensatory cis-trans regulation in the evolution of mouse gene expression. *Genome Res* **22**, 2376-2384 (2012).
37. O. Stegle, S. A. Teichmann, J. C. Marioni, Computational and analytical challenges in single-cell transcriptomics. *Nat Rev Genet* **16**, 133-145 (2015).
38. T. E. Boursalian, J. Golob, D. M. Soper, C. J. Cooper, P. J. Fink, Continued maturation of thymic emigrants in the periphery. *Nat Immunol* **5**, 418-425 (2004).
39. J. S. Hale, T. E. Boursalian, G. L. Turk, P. J. Fink, Thymic output in aged mice. *Proc Natl Acad Sci U S A* **103**, 8447-8452 (2006).
40. P. J. Fink, The biology of recent thymic emigrants. *Annu Rev Immunol* **31**, 31-50 (2013).