Author: Dijkgraaf, F.E.
Title: T cells in focus: Formation and function of tissue-resident memory
Issue date: 2021-01-12

Chapter 5

Tissue-resident CD8$^+$ memory T cell formation is clonally imprinted prior to tissue entry and fixed upon antigen re-encounter

Lianne Kok[1,*], Feline E Dijkgraaf[1,*], Jos Urbanus[1], Kaspar Bresser[1], David W Vredevoogd[1], Rebeca F Cardoso[1, 2], Leila Perie[3], Joost B Beltman[4], and Ton N Schumacher[1,5,#]

Published in revised format in: *Journal of Experimental Medicine,* 2020 Oct 5;217(10):e20191711. doi: 10.1084/jem.20191711.

[1] Division of Molecular Oncology & Immunology, Oncode Institute, The Netherlands Cancer Institute, Amsterdam, The Netherlands

[2] Present address: Immunology and Allergy Unit, Department of Medicine Solna, Karolinska Institute and University Hospital, Stockholm, Sweden

[3] Institut Curie, PSL Research University, CNRS UMR168, 11 rue Pierre et Marie Curie, 75005 Paris, France.

[4] Division of Drug Discovery & Safety, Leiden Academic Centre for Drug Research, Leiden University, Leiden, The Netherlands

[5] Department of Immunohematology and Blood Transfusion, Leiden University Medical Center, Leiden, The Netherlands.

* These authors contributed equally to this work

\# To whom correspondence should be addressed: t.schumacher@nki.nl

ABSTRACT

An increasing body of evidence emphasizes the crucial role of tissue-resident memory T cells (T_{RM}) in the defense against recurring pathogens and malignant neoplasms. However, little is known about the origin of these cells and their kinship to other $CD8^+$ T cell compartments. To address this issue, we traced the output of individual $CD8^+$ T cells to the T_{RM}, T circulating memory (T_{CIRCM}) and the T effector (T_{EFF}) pool by lineage analysis. We demonstrate that, while individual T cell clones contribute proportionally to systemic and local immunity during the effector phase, a subset of T cell clones is biased to form the tissue-resident memory T cell pool that arises following antigen clearance. Notably, this preferential T_{RM} formation is a clone intrinsic property rather than a stochastic process. Our data indicate that the capacity of T_{RM} formation is imprinted at the clonal level prior to tissue entry, and is preserved upon subsequent antigen encounter.

INTRODUCTION

Upon local infection, antigen-specific naïve CD8$^+$ T cells undergo rapid clonal expansion to generate a large pool of effector T cells (T_{EFF}) that are present in the circulation and at the affected peripheral site. Following pathogen clearance, this effector cell population contracts to form a small pool of memory T cells in the blood and secondary lymphoid organs (T_{CIRCM}), but also at the site of pathogen entry (Steinert et al., 2015). The latter population of tissue-resident memory T cells (T_{RM}) has been shown to be important for local control of reinfection in tissues such as skin, intestine and lung (Gebhardt et al., 2009, Ariotti et al., 2012, Masopust et al., 2010, Turner et al., 2014, Mueller and Mackay, 2016), and can be distinguished from its circulating counterpart by increased expression of markers such as CD103 and CD69 (Mackay et al., 2013, Mueller and Mackay, 2016).

A number of studies have provided evidence that certain subsets of T_{EFF} cells possess an enhanced capacity to differentiate into T_{RM}. Specifically, T_{EFF} located in inflamed tissues that express CD69, CD103 or CD127, but lack KLRG1 expression are considered to have a superior capacity to give rise to T_{RM} (Sheridan et al., 2014, Mackay et al., 2013, Herndler-Brandstetter et al., 2018). Furthermore, those T_{EFF} in peripheral tissues that are prone to differentiate into T_{RM} display a unique phenotype that differs from the transcriptional profile associated with T_{CIRCM} formation (Milner et al., 2017). While these studies have established that the propensity to generate T_{RM} is unequally distributed over the effector pool, prior work has also demonstrated that T_{RM} and T_{CIRCM} share a common clonal origin (Gaide et al., 2015). Thus, differences in T_{RM} formation capacity do not appear imprinted in naïve CD8$^+$ T cells, but a diversification in T_{RM} potential is evident in the effector T cell pool. Importantly, at which point the progeny of naïve T cells is instilled with T_{RM} forming capacity, how this trait is distributed over the pool of responding naïve T cell clones, and whether this capacity is stably imprinted has not been established. To address these issues, we tracked the offspring of individual naïve T cells responding to local skin vaccination or infection by means of genetic barcoding. Using such lineage tracing tool, we provide evidence that, while independent T cell clones possess an equal capacity to enter inflamed tissue during the effector phase, a subset of T cell clones possesses a heightened capacity to subsequently form resident T cell memory. Moreover, this clone intrinsic propensity to generate T_{RM} is acquired prior to tissue entry and is fixed upon secondary antigen encounter.

RESULTS

Individual T cell clones contribute proportionally to the systemic and skin effector T cell response

To evaluate how individual naïve T cells contribute to the T_{RM} lineage, and how the T_{RM} population is developmentally related to the systemic $CD8^+$ T cell subsets, we set-out to track the progeny of individual naïve $CD8^+$ T cells within the T_{EFF}, T_{CIRCM} and T_{RM} compartment *in vivo* by cellular barcoding. To this purpose, we first generated a high diversity retroviral barcode library that comprises approximately 200,000 unique cellular identifiers, thereby enabling the tracking of many individual cells in parallel. Using this BC2.0 genetic labeling system, we subsequently generated naïve $CD8^+$ T cells that each carry a unique DNA barcode (Gerlach et al., 2010, Gerlach et al., 2013). Specifically, thymocytes were transduced with the BC2.0 library and injected intra-thymically into recipient mice, to allow maturation into barcode-labeled naive T cells. This experimental approach allows for the genetic labeling of naturally cycling T cell precursors, thereby avoiding a requirement for *in vitro* activation of naïve T cells. As shown previously, barcode-labeled T cells that are generated in this manner behave identical to unmanipulated naïve OT-I T cells, both in terms of T cell response kinetics and effector differentiation potential (Gerlach et al., 2010). To be able to examine T cell fate and T cell development into the T_{RM} lineage without TCR affinity as a confounder (Zehn et al., 2009), thymocytes were obtained from OT-I transgenic mice, of which all $CD8^+$ T cells carry the OT-I TCR specific for the $OVA_{257-264}$-H2-K^b complex (**Fig. 1 A**).

Following *in vivo* development of barcode-labeled thymocytes into mature naïve GFP^+ OT-I T cells, cells were harvested and physiologically relevant numbers (i.e. 500-1,000) of cells were transferred into wild type recipient mice. Subsequently, a local immune response was induced by vaccination of hind leg skin of recipient mice with DNA encoding the $OVA_{257-264}$ epitope (Bins et al., 2005, Oosterhuis et al., 2012, Ahrends et al., 2016) (**Fig. 1 A**). Local vaccination induced clonal expansion and subsequent contraction of the barcode-labeled OT-I T cell pool (**Fig. 1 B** and **Supplementary Fig. 1 A**). At late time points (>60 days) following vaccination, GFP^+ OT-I T cells remained detectable at low frequencies in both the circulation and at the site of skin vaccination (**Fig. 1 C**). Consistent with prior work, the large majority of the (barcode-labeled) T_{RM} cells harvested from the tissue site expressed the canonical tissue-residency markers CD69 and CD103 (**Fig. 1 C**).

Having validated that skin vaccination induces clonal expansion of naïve barcode labeled T cells and their differentiation into T_{EFF}, T_{CIRCM}, and T_{RM} cells, we aimed to assess whether individual naïve T cells differ in their capacity to yield T_{EFF} at distinct body sites. To this end, vaccinated recipient mice were sacrificed at the peak of the T_{EFF} expansion phase (d12), and blood, spleen, draining lymph nodes (dLN) and the affected skin tissue were collected, and clonal output was quantified by DNA barcode sequencing (**Fig. 1D**, top left). Barcode analysis of GFP^+ OT-I T cells present in the blood compartment at the

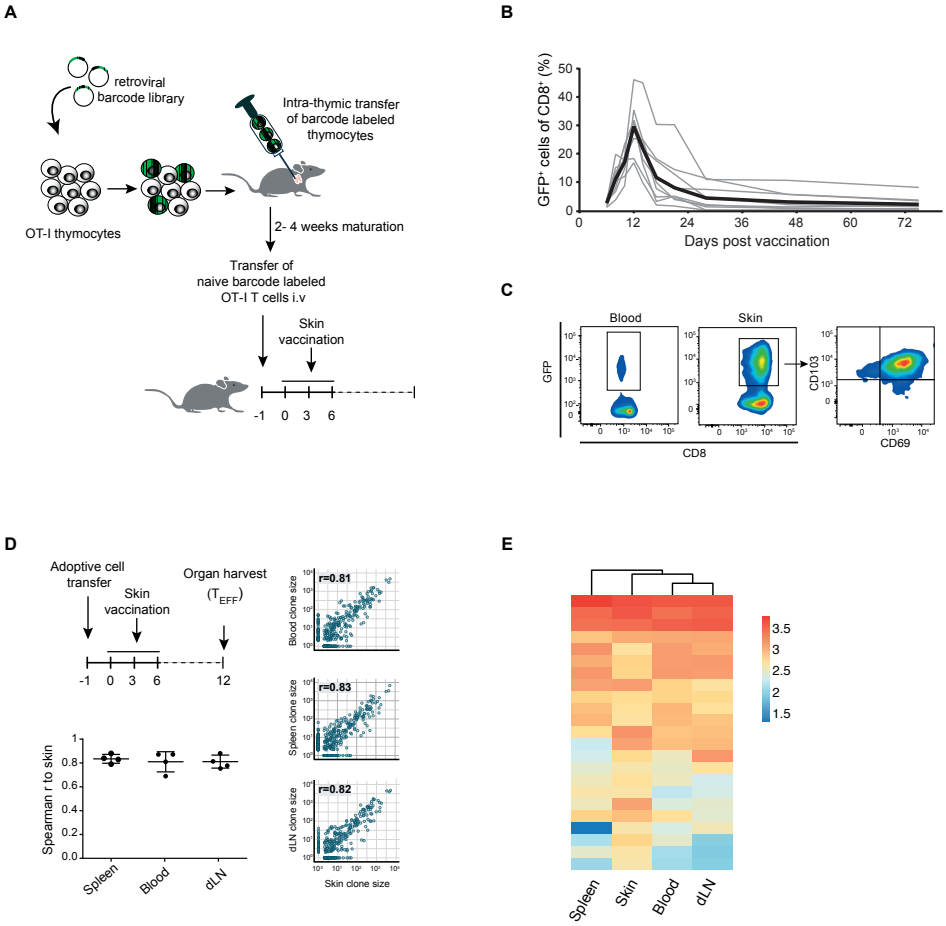


Figure 1 | Proportional contribution of individual T cell clones to the systemic and skin effector response. OT-I thymocytes were transduced with the barcode library and intra-thymically transferred into recipient mice. After maturation, barcode labeled GFP+ OT-I T cells were transferred into secondary recipients that were subsequently exposed to skin vaccination. **A**, Schematic overview of experimental set-up. **B**, Barcode labeled GFP+ OT-I T cell response to DNA vaccination in blood (n=11 mice, gray lines). Black line represents group average. **C**, Representative flow cytometry plots showing the presence of GFP+ memory T cells within CD8+ cells in blood and skin on d>60 after vaccination. **D**, **E**, Spleen, skin, dLNs and whole blood were collected from vaccinated recipient mice on d12 after start of vaccination. D, Analysis of the contribution of individual T cell clones to the spleen, blood and dLN effector stage T cell compartment, relative to the skin effector stage T cell compartment. Spearman correlation r was calculated over clones detected in both samples, D (left), Spearman correlations for individual mice (n=4), mean with whiskers representing SD. D (right), Dots represent individual clones, p values were <0.0005. E, Clonal output in all examined tissues of the 5% of largest clones detected in skin tissue. Heat map depicts log10-transformed clone sizes (read counts). D, E, Representative data of two independent experiments.

peak of the response showed that, similar to prior lineage tracing studies involving *Listeria monocytogenes*-OVA$_{257-264}$ infection (Gerlach et al., 2013, Buchholz et al., 2013), the capacity of individual naïve T cells to expand in response to DNA vaccination was highly variable, with ~7% of the clones producing ~50% of the total effector T cell pool. Comparison of clonal output in the sampled tissues showed that at the peak of the antigen-specific T cell response, the vast majority of clones contributed to the T cell pool at all the 4 examined locations (**Fig. 1D** bottom and right, **1E**, and controls in **Supplementary Fig. 1 B, C**). Furthermore, the relative sizes of individual T cell clones at these different sites were highly correlated (r>0.8), indicating that the progeny of different naïve T cells possesses a similar capacity to disseminate throughout the body during the effector T cell stage (**Fig. 1 D**). As a control, the high clonal overlap between T cell compartments in the skin and at other body sites was shown not to be explained by a potential contamination of skin samples with blood borne T cells (**Supplementary Fig. 1 D**). Thus, the ability to enter inflamed peripheral tissues is equally distributed over the progeny of responding T cell clones.

Clonal bias in tissue-resident memory T cell generation

Having established that individual T cell clones display a similar capacity to disseminate to the skin and lymphoid compartments during the effector phase, we next evaluated whether this equal distribution of clones persisted into memory. To quantify the output of individual clones in the two memory compartments, recipient mice received a local skin vaccination, T_{EFF} blood samples were drawn at day 12, and the skin-T_{RM} and T_{CIRM} populations from the same mice were isolated after memory formation (>d60, **Fig. 2A**). In line with prior work (Gaide et al., 2015), comparison of clone sizes in the two memory pools revealed that a large majority of naïve T cells (84.8%) contributed to both the T_{CIRCM} and T_{RM} cell lineage. Strikingly, however, the contribution of individual T cell clones to the T_{CIRCM} or T_{RM} pool was highly disparate, with a correlation in clone size of r=0.32 (**Fig. 2 B**, quality controls in **Supplementary Fig. 2 A, B**). Importantly, this low degree of overlap was not due to suboptimal sampling of the lower number of T cells in the memory phase, as shown by the high correlation (r>0.9, **Supplementary Fig. 2 A**) of technical replicates of either the skin-resident or the circulating memory T cell pool. Thus, while during the effector phase individual T cell clones contribute essentially equally to the T cell pool at different body sites, many clones preferentially contribute to either circulating or tissue-resident T cell memory following contraction. Furthermore, this disparity in memory clone distribution is also present upon natural infection, as shown by DNA barcode analysis of the circulating and tissue-resident memory T cell compartment upon localized herpes simplex virus (HSV-OVA$_{257-264}$) infection. Specifically, following HSV-OVA$_{257-264}$ infection, the average T cell clone preferentially contributed toward either the T_{RM} or the T_{CIRCM} compartment by a factor of 11.34-fold (r=0.25, **Fig. 2 C**). As a control, the average ratio between technical replicates was 2.19 (r=0.86). By the same token, in response to DNA vaccination, T cell clones showed a preferential contribution toward either

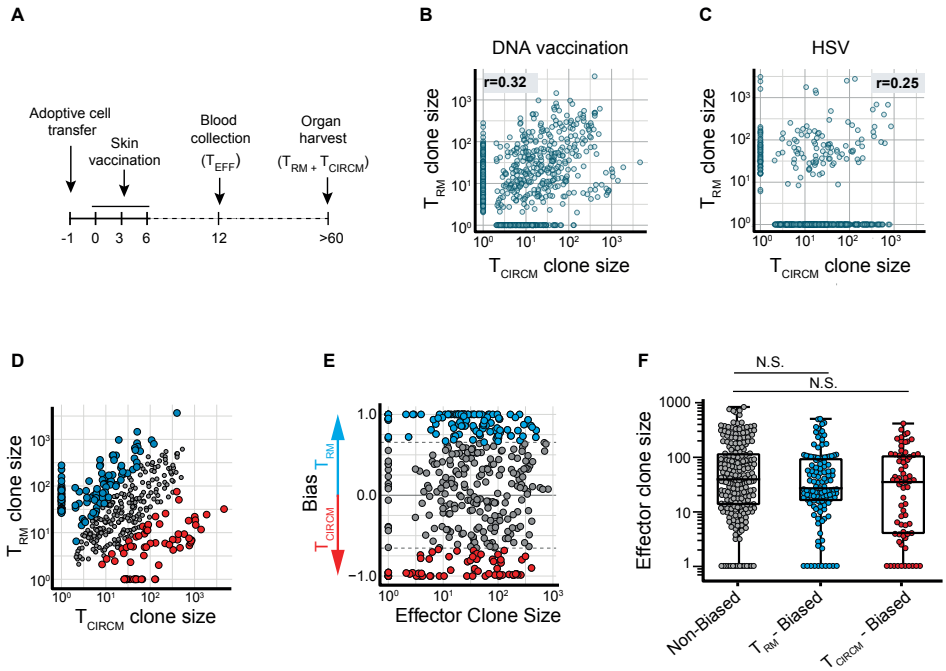


Figure 2 | Clonal bias in tissue-resident memory T cell generation. A, Representation of the experimental timeline. Barcode-labeled T_{RM} and T_{CIRCM} cells were isolated from the skin and from the circulatory compartment (spleen, LN and blood) of DNA vaccinated (**B**) or HSV-OVA$_{257-264}$ infected (**C**) recipient mice, and clonal output was quantified. B, Comparison of clonal contribution to the skin-T_{RM} and T_{CIRCM} compartment after DNA vaccination. C, Comparison of clonal contribution to the skin-T_{RM} and T_{CIRCM} compartment after HSV-OVA$_{257-264}$ infection. D, E, F, Clones responding to DNA vaccination were defined as T_{RM}-biased, T_{CIRCM}-biased, or non-biased based on their relative contribution to either memory compartment. **D**, Scatterplot similar to (B), highlighting T_{RM} biased (blue), T_{CIRCM} biased (red) and non-biased (gray) T cell clones. Small clones for which clone size measurements were less reliable were excluded from analysis and are not depicted. **E**, **F**, Comparison of effector stage burst size of non-biased (gray), T_{RM}-biased (blue), and T_{CIRCM}-biased (red) T cell clones. E, Values on Y-axis depict *(clone size T_{RM} - clone size T_{CIRCM}) / (clone size T_{RM} + clone size T_{CIRCM})*, and represents the degree of preferential contribution to T_{RM} or T_{CIRCM}. Dashed lines indicate bias threshold of 4.8-fold. F, Median with whiskers representing min/max, Kruskal-Wallis test with Dunn's multiple comparisons test, N.S., not significant. B, C, Spearman correlation r was calculated over all clones that contributed to both samples, B: P<0.0005, C: P=0.01. Data for four mice, representative of two individual experiments. B-F, Dots represent individual clones.

the circulating or skin-resident memory T cell compartment by a factor of 11.98 (factor of 1.66 between technical replicates).

Next, we examined whether the bias in T_{RM} and T_{CIRM} generation in response to DNA vaccination could be explained by differences in clonal T_{EFF} expansion. First, to exclude clones that could show clonal bias because of random sampling variation, clones that were exclusively observed in one of the two memory T cell compartments and that represented

<0.25% of that pool were removed (retaining 58.5% of barcodes and 97.2% of reads; pre-filtering: **Fig. 2 B**, post-filtering: **Fig. 2 D**, filtering strategy: **Supplementary Fig. 2 C**). Subsequently, to be able to identify biased clones, we defined a 'bias threshold' based on comparison of technical replicates, a setting in which clonal bias can by definition not occur (**Supplementary Fig. 2 C**). Application of the resulting threshld (a fold difference of >4.8) to the lineage tracing data revealed that close to 50% of T cell clones preferentially contributed to either memory T cell compartment, with 29.7% of clones being biased toward T_{RM} formation, and 16,9% biased toward T_{CIRCM} formation (**Fig. 2 D**). Notably, analysis of effector phase burst sizes of T_{RM}-biased, T_{CIRCM}-biased, and non-biased T cell clones showed that biased memory T cell generation was both observed for T_{EFF}-stage clones that had undergone massive or little expansion (**Fig. 2 E, F**). These results demonstrate that – independent of clonal burst size – a large fraction of T cell clones preferentially produces T_{RM} or T_{CIRCM}, indicating that T_{RM} and T_{CIRCM} are not only separated by location and phenotype, but also by descent.

Non-stochastic formation of tissue-resident and systemic T cell memory

Next, we wanted to understand whether the clonal bias observed in memory (**Fig. 2 B, D**) was due to remodeling of either the circulatory or the skin-resident compartment during T cell contraction. As clonal hierarchy is highly similar at different body sites during the effector phase (**Fig. 1 D, E**), we reasoned that the T_{EFF} pool in blood could be used as a "historical snapshot" of clonal distribution in all immune compartments before memory formation. Comparison of clone sizes between d12 effector blood and the two T cell memory compartments of the same mice demonstrated that the skin and spleen compartment in memory phase were substantially more disparate from the blood T_{EFF} compartment than in the effector phase (**Fig. 3 A-C** and **Supplementary Fig. 3**). Thus, during memory formation, both the skin-resident and the circulating T cell compartment undergo a substantial change in clonal hierarchy (**Fig. 3 A-C** and **Supplementary Fig. 3**), resulting in differential contributions of individual T cell clones to the two memory compartments (**Fig. 2 D**).

The observed divergence in clonal composition of T cell populations at the two sites could either arise through an intrinsic difference in cell fitness to survive in particular microenvironments, or through the stochastic engraftment of cells at the individual sites. To test the latter hypothesis, we simulated the generation of T_{RM} and T_{CIRCM} pools that were derived from a founder population with a size that equaled either the experimentally observed T memory pool (indicated as α, **Fig. 3 D**), 10% of the observed T memory pool, or the smallest possible founder pool (i.e. the number of individual clones observed in the memory pool, indicated as β, **Fig. 3 D**). Subsequently, the correlation in clone sizes between the simulated T memory pools and the experimentally observed T_{EFF} pool (indicated as Y, **Fig. 3 D**) were calculated and compared to the correlation between the experimentally observed T memory and T_{EFF} pool (indicated as X, **Fig. 3 D**). Note that only when Y approaches X, stochastic engraftment of T cells can explain the observed clonal bias in memory phase. Interestingly,

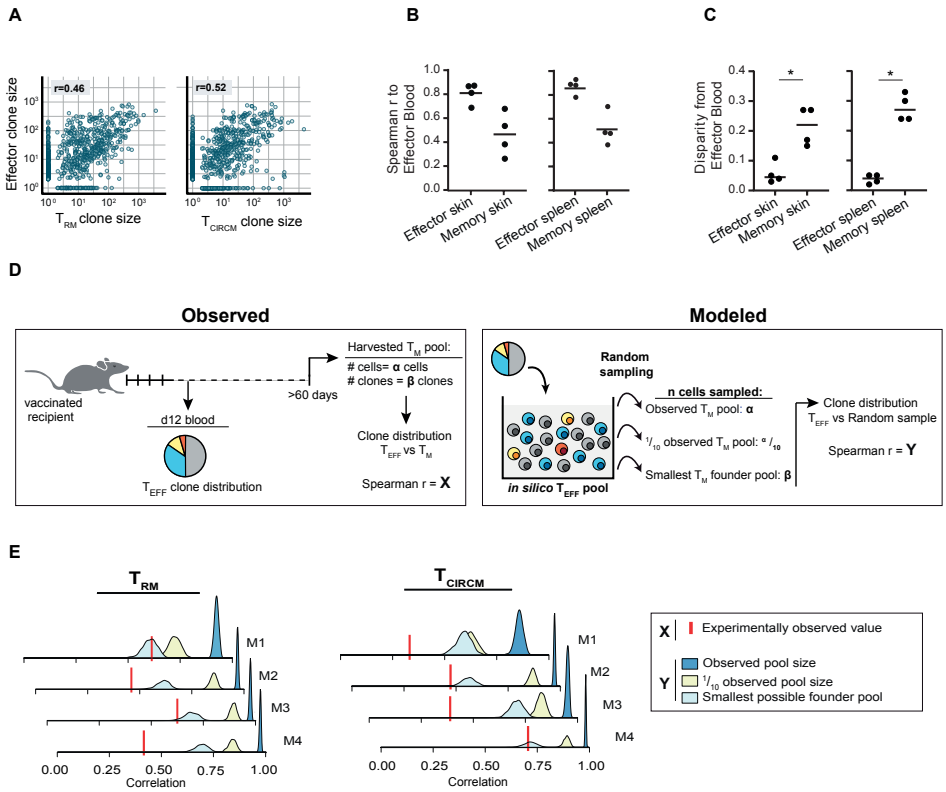


Figure 3 | Non-stochastic formation of tissue-resident and systemic T cell memory. A, Contribution of T cell clones to the T_{RM} (left) or the T_{CIRCM} pool (right), relative to the effector stage blood compartment. Spearman correlation r was calculated over T cell clones that were detected in both samples, n=4 mice. **B**, Spearman correlations of clone sizes in skin (left) and spleen (right) samples collected during effector (n=4 mice) and during memory phase (n=4 mice) to d12 effector blood. **C**, Clone size disparity of skin (left) and spleen (right) T cell pools in the effector and memory phase from the d12 effector blood T cell pool. See Supplementary Fig. 3 A on the definition of disparity. **D**, Illustration depicting the strategy used to assess whether stochasticity can explain the observed clonal skewing during memory formation. Based on observed clonal distribution in the T_{EFF} pool, a virtual pool of T_{EFF} cells is generated *in silico* from which cells are samples to form a randomly selected T_{RM} or T_{CIRCM} memory pool. The number of randomly sampled cells is either equal to the number of observed cells in the biological memory (T_M) pool (α), to 10% of the observed T_M pool (α/10), or to the number of observed clones in the biological T_M pool (β), which represents the smallest theoretically possible T_M founder pool. The Spearman correlation coefficient between the randomly sampled cell pool and the experimentally observed T_{EFF} is calculated (Y) and compared to the Spearman correlation coefficient between the experimentally observed T_M pools and the experimentally observed T_{EFF} pool (X). Only if Y approaches X, stochastic engraftment can explain the observed skewing in clonal output in the T_M pool. **E**, Stochastically formed T_{RM} (left) and T_{CIRCM} (right) pools were modeled 10,000 times *in silico,* as described in D, and the Spearman correlation between the modeled memory pools and the observed T_{EFF} pool was calculated (Y). Graphs indicate individual mice (n=4) and histograms represent the distribution of Spearman r. Red vertical line indicates the correlation between the clonal distribution of the T_{EFF} pool and the experimentally observed memory pool (X). Spearman r correlations were calculated over all clones detected either in the effector pool or the (modeled or experimental) memory pool. A, Dots represent individual clones. B, C, Dots represent individual mice. A, B, Spearman correlation r was calculated over clones that were detected in both samples A (left): P<0.0005, A (right): P<0.0005. C, Mann-Whitney U-test, * P<0.05. Representative data of two individual experiments.

this analysis demonstrated that stochastic engraftment of a founder population with the size of the observed T memory pool (α) or 1/10th of this size could not explain the observed skewing during T memory formation in any of the mice (**Fig. 3 E**). Furthermore, stochastic engraftment by the smallest possible founder pool was also insufficient to explain the skewing in the observed T cell memory pool in the majority of mice (**Fig. 3 E**). Collectively, these data indicate that the skewed composition of both the T_{RM} and T_{CIRCM} pool is unlikely to be explained by stochastic survival or engraftment, thereby suggesting the existence of intrinsic differences between T cell clones in their capacity to form systemic and tissue-resident T cell memory.

T_{RM} differentiation is a clone-imprinted attribute that is preserved upon secondary antigen encounter

To directly test whether individual T cell clones display an intrinsic difference in their ability to form circulating and resident T cell memory, two conceptually distinct strategies may be used. As a first approach, phenotypic marker expression by effector stage clones may be analyzed to understand whether cell phenotype predicts the capacity of clones to yield the two types of memory cells. CD27 and KLRG1 have previously been used to identify subsets of the systemic effector T cell pool that display an increased and reduced memory potential, respectively (Sarkar et al., 2008, Joshi et al., 2007, Obar and Lefrancois, 2010, Kaech and Wherry, 2007). In addition, CD69 expression by T_{EFF} in the skin has been associated with T_{RM} formation (Mackay et al., 2015a, Mackay et al., 2013). Comparison of CD27-negative and CD27-positive and of KLRG1-negative and KLRG1-positive effector cells in spleen revealed a highly similar clonal distribution between the different subsets (r>0.7, **Fig. 4 A**, left and middle). Similarly, CD69 expression on T_{EFF} isolated from the inflamed skin compartment was highly constant between T cell clones (r=0.88, **Fig. 4 A**, right). Thus, the expression of 3 previously established effector T cell markers that are considered indicators of differentiation potential on a population level, does not suffice to predict T_{RM} or T_{CIRCM} generation potential of individual clones.

As a second, fully unbiased strategy to determine whether the capacity to form a local memory T cell pool is a pre-determined and thus cell intrinsic property, it may also be tested whether a given cell pool reproducibly shows the same behavior. To this purpose, we aimed to generate two sites of skin-resident T cell memory, by parallel vaccination of the right and left hind leg skin of mice. If the development of T_{RM} would solely be determined by stochastic encounter of a micro-environmental signal, clone size distributions in the two anatomically separate skin sites would be expected to be disparate. Conversely, if T_{RM} fate-commitment were to be clonally imprinted, the two skin sites would be expected to show a similar clonal distribution. Comparison of the clonal composition of either the left or the right leg skin memory T cell compartment with that of the circulating memory compartment at >day 60 post-vaccination recapitulated the prior observation that a large fraction of naïve T cells

yield progeny that either preferentially forms systemic T cell memory or tissue-resident T cell memory ($T_{RM\text{-}LEFT}$ - T_{CIRCM}: r=0.37, P<0.0005 ; $T_{RM\text{-}RIGHT}$ - T_{CIRCM}: r=0.30, P<0.0005) (**Fig. 4 B**), with the average T cell clone differing more than 10-fold in contribution to the skin and the systemic memory compartment (average ratio $T_{RM\text{-}LEFT}$ - T_{CIRCM}: 10.14, average ratio $T_{RM\text{-}RIGHT}$ - T_{CIRCM}: 11.67, **Fig. 4 C**, right). Strikingly, comparison of the T_{RM} populations at the two spatially separated skin sites revealed a substantially higher degree of overlap (r=0.78, P<0.0005), with an average clone size ratio of 3.17 (**Fig. 4 C**). In order to compare the magnitude of this clone-intrinsic bias in T_{RM} formation relative to a bias of individual T cell

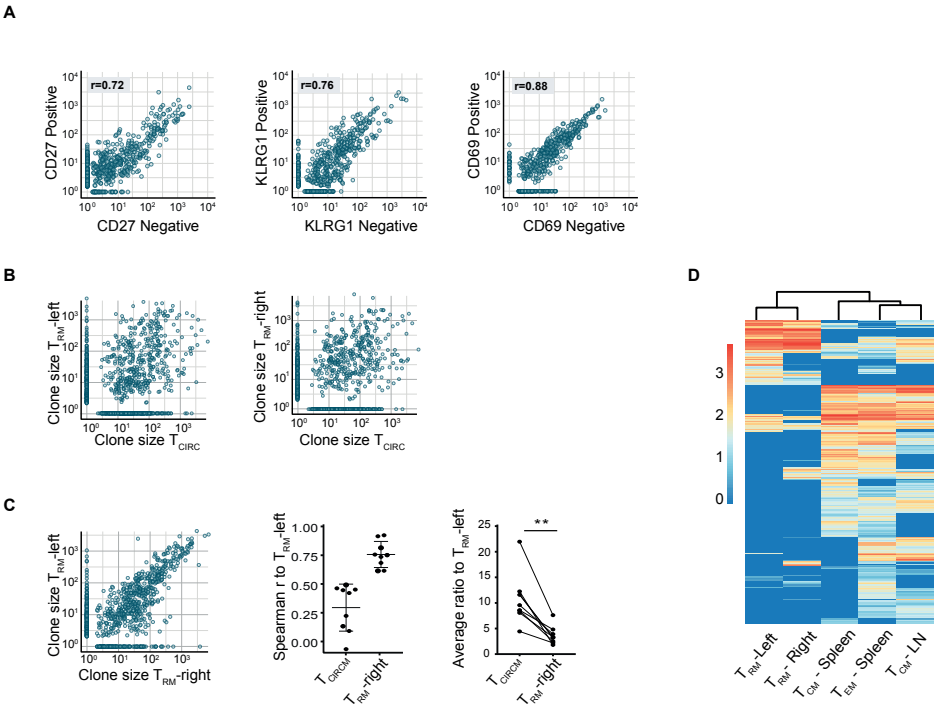


Figure 4 | T_{RM} differentiation is a clone-imprinted attribute. A, Analysis of clone sizes in the CD-27pos/neg (left) and KLRG1pos/neg (middle) T cell compartment from spleen 12 days after vaccination, and CD69pos/neg (right) T cell compartment from skin 15 days after vaccination. Each plot displays data of four individual mice derived from individual experiments. Dots represent individual clones. All plots: P<0.0005. **B**, **C**, Contribution of T cell clones to the T_{RM} pool present at two separate sites of primary vaccination ($T_{RM\text{-}LEFT}$, $T_{RM\text{-}RIGHT}$) relative to the T_{CIRCM} pool (B) and relative to each other (C, left). Dots represent individual clones. Data from nine mice from two independent experiments. C, (middle) Spearman correlations, with mean with whiskers representing SD, and average ratios (right) of individual mice, comparing the clonal composition of the $T_{RM\text{-}LEFT}$ compartment to the T_{CIRCM} and to the $T_{RM\text{-}RIGHT}$ compartment. Dots represent individual mice (n=9). **P<0.005, Wilcoxon signed-ranked test. **D**, Output of individual OT-I T cells to different T_{RM} and T_{CIRCM} pools, as indicated in the columns. Heat map depicts log10-transformed clone sizes (read counts), clustered using Euclidian distance. Data from 6 mice from two independent experiments. A, C: Spearman correlations are calculated over clones that contributed to both samples.

clones to yield either systemic central memory (T_{CM}) or effector memory (T_{EM}) T cell pools, we subsequently performed barcode lineage tracing of T_{RM} from the two anatomically separate skin compartments, of T_{CM} (defined as CD62L$^+$) from LN and spleen, and of T_{EM} (defined as CD62L$^-$) from spleen. Complete-linkage clustering analysis again showed the highly similar clonal composition of the memory T cells at the two spatially separated skin compartments (**Fig. 4 D**). In addition, this analysis revealed that these two T_{RM} compartments differ more strongly in clonal composition from all the 3 systemic memory T cell compartments than, for instance, splenic T_{EM} and LN T_{CM} differ from each other (**Fig. 4 D**). Thus, relative to differences in capacity to produce central memory or effector memory T cells, clonal imprinting of the capacity to yield tissue-resident T cell memory versus systemic T cell memory is profound.

Finally, to test whether the acquisition of T_{RM} generation potential is a stable property of CD8$^+$ T cells, recipients of barcode-labeled naïve OT-I T cells were subjected to a primary vaccination on the right hind leg, followed by a secondary vaccination on the left hind leg >60 days later (**Fig. 5 A**). In line with prior work (Jiang et al., 2012, Casey et al., 2012), low frequencies (on average 4-fold less than at the vaccinated site) of T_{RM} were detected at the

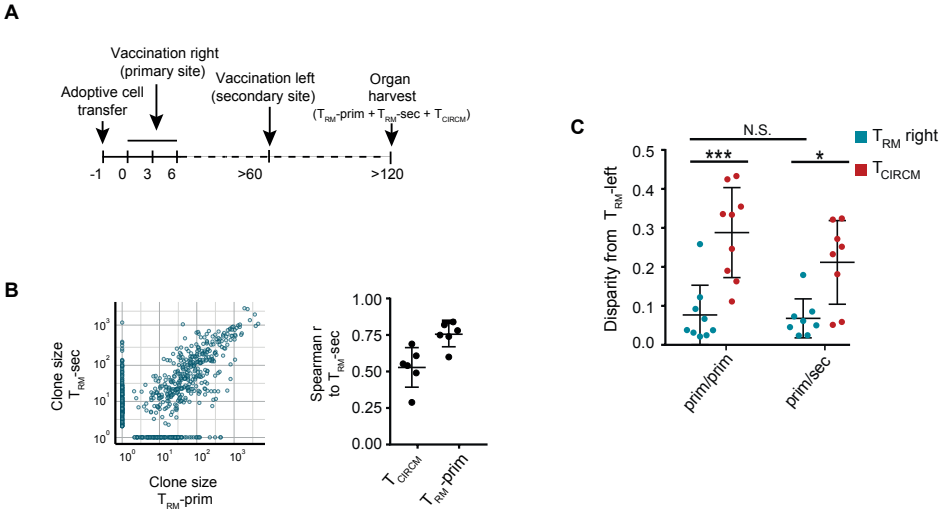


Figure 5 | The clone-intrinsic propensity to generate T_{RM} is preserved upon secondary antigen encounter. Recipient mice were vaccinated on the right hind leg (primary site) and >60d later on the left hind leg (secondary site), and clonal composition was assessed >60d after secondary vaccination. **A**, Schematic timeline of experiment. **B** (left), Contribution of T cell clones to the $T_{RM\text{-}SEC}$ pool relative to the $T_{RM\text{-}PRIM}$ pool. Dots represent individual clones. Right: Spearman correlations of 6 individual mice, mean with whiskers representing SD. **C**, Disparity between T_{RM} and T_{CIRCM} pool and between the two T_{RM} pools in case of simultaneous or staggered vaccination. Prim/prim indicates simultaneous vaccination, prim/sec indicates primary and secondary vaccination separated by >60d. N.S., not significant; *P<0.05, ***P<0.0005, Mann-Whitney U-test. Mean with whiskers representing SD. See Supplementary Fig. 3A on the definition of disparity. Dots represent individual mice. Prim/prim and prim/sec groups each consist of 9 mice. Data from three independent experiments.

initially unperturbed tissue site upon primary vaccination (**Supplementary Fig. 4 A**). Following secondary vaccination at this site, local memory T cell numbers increased to exceed those seen at the primary vaccination site, indicative of *de novo* T_{RM} formation induced by the secondary vaccine (**Supplementary Fig. 4 B**). Subsequently, barcode abundance was separately assessed at the primary and secondary vaccination site late (>60 days) after secondary vaccination, and was compared to barcode abundance in the circulating memory T cell pool at the same time point. This analysis revealed that the secondary T_{RM} pool was dissimilar to the T_{CIRCM} compartment in terms of clonal hierarchy (average r=0.5), but greatly resembled the T_{RM} pool generated at the primary site of vaccination (average r=0.73) (**Fig. 5 B**). Furthermore, disparity analysis (**Fig. 5 C** and **Supplementary Fig. 3 A**) revealed that the clonal composition of these two T_{RM} pools that were separated in time was equally similar as when two distinct T_{RM} pools were generated simultaneously, indicating that the capacity of individual T cell clones to yield T_{RM} is stable over time. Thus, these data reveal that the ability of effector phase T cells to form T_{RM} is differentially and permanently imprinted at a clonal level prior to skin entry.

DISCUSSION

The current data demonstrate that, while all naïve T cells yield progeny that disseminate equally well to inflamed skin and the systemic lymphoid compartments, a subset of T cell clones yield offspring with a heightened capacity to persist long term in peripheral tissues. The observation that tissue entry is equal between progeny derived from distinct clones implies that the selection of the T_{RM} privileged clones is not driven by an enhanced capacity of a subset of circulating effector stage clones to migrate into the inflamed tissue. At the same time, the observation that the clonal composition of T_{RM} pools that form at anatomically separate sites is highly similar indicates that the property to effectively produce T_{RM} is imprinted into T cells prior to tissue entry. Previous work of Masopust and colleagues has shown that, as based on homing receptor expression, T_{EFF} cells predominantly enter inflamed tissue during early stages of the T cell response (Masopust et al., 2010). This observation, together with the current observation of fate imprinting prior to tissue entry, suggest that the observed commitment to T_{RM} fate must already have occurred before the end of the effector phase. In line with this, the fact that the capacity to generate T_{RM} is unequally distributed over T cell clones implies that this property must be instilled prior to substantial clonal expansion. Earlier work has established a central role for local environmental cues, including TGFβ, IL-15 and cognate antigen (Muschaweckh et al., 2016, Mackay et al., 2015b), in promoting T_{RM} formation in peripheral tissues. The current data are consistent with a model in which a subset of T cell clones develop a heightened capacity to respond to such local cues, thereby promoting their differentiation into long-term persisting T_{RM} (**Supplementary Fig.**

5). Furthermore, while TCR affinity has been linked to T_{RM} generation potential (Frost et al., 2015), the current data indicate that differential imprinting of T_{RM} generation potential can occur independent of variations in TCR affinity.

By revealing an imprinted capacity to form tissue-resident T cell memory at anatomically separate sites, our data provide evidence for the existence of a T_{RM} precursor in the systemic immune compartment at an early stage following immunization. While the mechanisms that drive the divergence in memory differentiation potential on a clonal level remain to be elucidated, an extensive body of work has demonstrated that external factors, such as cytokines and ligands of co-stimulatory receptors, at the T cell priming site can influence the production of functional memory T cells (Parameswaran et al., 2005, Scholer et al., 2008, Ahrends et al., 2017, Mousavi et al., 2008, Hendriks et al., 2005, Agarwal et al., 2009, Cui and Kaech, 2010). In addition, cross-priming by Batf3$^+$ cDC1s (Iborra et al., 2016) and inhibition of mTOR activity (Araki et al., 2009, Sowell et al., 2014) have opposing roles in promoting T_{RM}-lineage over T_{CIRCM} cell fate commitment. Conceivably, differential exposure of individual T cell clones to these cues during the priming process forms the mechanistic basis for T_{RM} precursor formation. To evaluate the role of such signals, but also a potential contribution of the developmental origin of naïve T cells (Smith et al., 2018) and of stochastic variation in gene expression (Marchingo et al., 2016, Feinerman et al., 2008), the use of advanced lineage tracing technologies that allow the mapping of intra-clonal fate diversification over time should be of value (Kalhor et al., 2018, Alemany et al., 2018, Spanjaard et al., 2018, McKenna et al., 2016).

MATERIALS AND METHODS

Mice

C57BL/6J-Ly5.1, C57BL/6J, OT-I, mTmG and UCB-GFP mice were obtained from Jackson Laboratories, and strains were maintained in the animal department of The Netherlands Cancer Institute (NKI). The mTmG and UCB-GFP mice were crossed with OT-I mice to obtain the mTmG-OT-I and GFP-OT-I strains, respectively. All animal experiments were approved by the Animal Welfare Committee of the NKI, in accordance with national guidelines.

Generation of the BC2.0 high diversity retroviral barcode library

The BC1DS_lib oligo (**Supplementary Table 1**) containing a 21nt random barcode sequence was PCR amplified (10 cycles: 10sec 98 °C, 30 sec 55 °C, 1 min 72 °C) with Phusion polymerase (New England Biolabs). The resulting PCR amplified product was column purified (MinElute PCR cleanup kit, Qiagen) and subsequently digested with XhoI and EcoRI, followed by ligation into the 3' UTR of the GFP cDNA sequence within the pMX retroviral vector, using the Electroligase kit (New England Biolabs). Electrocompetent DH10b bacteria

(Invitrogen) were then electroporated with 16 ng ligation product and a small fraction of the transformed bacteria was plated on Luria-Bertani agar plates to determine transformation efficiency, while the remaining bacteria were grown overnight in 400 ml Luria-Bertani medium (VWR Life Science) supplemented with ampicillin (Sigma). DNA was extracted from the bacterial culture using the Maxiprep kit (Invitrogen).

Establishment of the barcode reference list

To be able to match barcode sequences observed in biological samples to a reference list of barcodes present in the BC2.0 library, barcode sequences in the library were PCR amplified in duplicate (referred to as repA and repB) and subsequently sequenced as independent samples. In brief, barcodes were amplified from 10 ng retroviral library DNA using a combination of native Taq DNA polymerase (Invitrogen) and Deep Vent polymerase (New England Biolabs) at a 2:1 ratio, in three consecutive rounds of PCR. First round PCR was performed using the Top_lib and Bot_lib primers (15 cycles: 5 sec 94 °C; 5 sec 57.2 °C; 10 sec 72 °C), second round PCR was performed using the BC1v2DS_For and BC1v2DS_Rev primers (15 cycles: 5 sec 95 °C; 5 sec 58 °C; 10 sec 72 °C), Third round PCR was performed using the P5_For and P7_Index_Rev primers (7 cycles: 5 sec 94 °C; 10 sec 58 °C; 10 sec 72 °C). Resulting PCR products were sequenced on an Illumina hiSeq2500 lane. For primer sequences, see **Supplementary Table 1**.

In the sequencing data of repA and repB, 349,439 and 333,422 unique barcode sequences were detected, respectively, with 64.32% of all detected sequences being shared between the two replicates. Many of these sequences are likely to be spurious, resulting from PCR and sequencing errors. Such spurious sequences derive from true 'mother barcodes' that have a much higher abundance than the 'child sequences'; with child sequences differing up to several nucleotides from the mother sequence and having a reproducible frequency of occurrence of up to ~5% of the abundance of the mother barcode (Beltman et al., 2016). To remove those spurious barcode variants, we removed all sequences that had a Levenshtein distance of ≤4 nucleotides (Levenshtein, 1966) from a potential mother barcode and that also had a read count of ≤5% of that potential mother barcode. Additional spurious barcodes that occur at a very low abundance are likely to escape from this cleaning procedure, for instance because they contain >4 nucleotides differences from their mother. For this reason, only barcodes that were detected at least 3 times in the two replicates combined were retained in the barcode reference list. After this filtering, a list of 263,582 unique barcodes was obtained, of which only 1.27% was not shared between technical replicates.

Generation of barcode labeled T cells

Retrovirus of the barcode library was produced by transfection of Phoenix-E packaging cells using FuGene™6 (Roche). Retroviral supernatant was harvested 48 h after transfection and stored at -80 °C. To generate naïve barcode labeled OT-I T cells, thymocytes were harvested

from 5-7 week old OT-I mice and transduced with the barcode library virus by spin-infection (90 min, 400g), in IMDM (Gibco Life Technologies) supplemented with 8% FCS, 100 U/ml penicillin, 100 µg/ml streptomycin and 10 ng/ml recombinant murine IL-7 (PeproTech). To limit the fraction of T cells with multiple barcode integrations, barcode library virus was diluted prior to transduction to obtain a transduction efficiency of 8-10%. After 24 h of culture, cells were harvested and viable thymocytes were enriched using Lympholite-M Cell Separation Medium (Cedarlane) followed by purification of GFP^+ cells by FACS (FACSAria II (BD Biosciences) and MoFLo Astrios (Beckman Coulter). Subsequently, ~1 million sorted GFP^+ thymocytes were intra-thymically injected into 5-7 week old C57Bl/6 or C57Bl/6-Ly5.1 primary recipient mice, as described previously (Gerlach et al., 2010, Gerlach et al., 2013). After a maturation period of 2-4 weeks, whole blood, spleen and LN (cervical, axillary, brachial, mesenteric, inguinal and lumbar) were harvested and pooled, followed by enrichment of $CD8^+$ T cells using the Mouse CD8 T Lymphocyte Enrichment Set (BD Biosciences). The fraction of GFP^+ cells in the $CD8^+$ T cell pool was determined by flow cytometry (Fortessa (BD Biosciences)) and 500-1,000 GFP^+ cells were adoptively transferred into 8-14 week old secondary C57Bl/6 or C57Bl/6-Ly5.1 recipient mice.

Immunization by DNA vaccination and Herpes Simplex-1 (HSV)-infection

One day prior to vaccination with the 'HELP-OVA' vector that encodes the $OVA_{257–264}$ epitope (SIINFEKL), the shuffled HPV E7 sequence, and MHC-II class restricted helper epitopes (Oosterhuis et al., 2012, Ahrends et al., 2016), fur was removed from hind legs with Veet depilation cream (Reckitt Benckiser). Primary DNA vaccination was subsequently performed on day 0, 3 and 6 by tattooing (Bins et al., 2005) a droplet of 15 µl of a 2 µg/µl DNA solution in 10 mmol/L Tris pH 8.0 and 1 mmol/L EDTA pH 8.0 per leg, by means of a sterile disposable 9-needle bar mounted on a rotary tattoo device oscillating at a frequency of 100 Hz for 1 min with a needle depth of 1 mm (MT.DERM). For secondary vaccinations, mice received a single DNA tattoo with 20 µl of the 2 µg/µl plasmid solution on the inside and on the outside of the leg, at >60 days after start of primary vaccination.

The $HSV_{TOM-OVA}$ virus, containing a CMV immediate-early promoter tomato-$OVA_{257–264}$ gene cassette in the intergenic region between the UL26 and UL27 genes of the HSV-1 strain KOS (Halford et al., 2004) was grown in Vero cells, as described previously (Weeks et al., 2000). One day prior to infection, fur was removed from hind legs with Veet depilation cream (Reckitt Benckiser). On day 0, a droplet of 7 µl containing ~3.125×10^5 PFU $HSV_{TOM-OVA}$ in DMEM (Gibco Life Technologies) per area was given once to both legs of anesthetized mice by means of a tattoo, using a sterile disposable 9-needle bar mounted on a rotary tattoo device oscillating at a frequency of 100 Hz for 1 min with a needle depth of 0.5 mm (MT.DERM). First macroscopic skin lesions became visible on treated areas around day 3 post infection (data not shown).

Recovery of barcode-labeled cells from vaccinated and HSV-infected recipient mice

To sample the effector T cell pool without sacrificing the animal, a blood sample of 300 µL was drawn from the tail vein. Erythrocytes were lysed using NH_4Cl, and samples were stored as cell pellets at -80 °C. To recover GFP$^+$ T cells from skin and secondary lymphoid organs, either in the effector or memory phase, mice were sacrificed, whole blood was collected by heart puncture and spleen and LNs (cervical, axillary, brachial, mesenteric, inguinal and lumbar) were harvested. The blood, spleen and LNs derived from one mouse were processed as one sample, unless indicated otherwise. In addition, skin tissue from the hind legs was collected and processed separately. For isolation of barcode labeled cells from skin tissue, Veet-depilated (Reckitt Benckiser) full-thickness skin was collected using scissors and forceps and minced into small pieces. Subsequently, skin fragments were taken up in DMEM (Gibco Life Technologies) supplemented with collagenase IV (Gibco) and II (Worthington Biochemical Corporation) (both 1.25 mg/ml final), deoxyribonuclease type I (DNAse I, 0.25 mg/ml final, Sigma Aldrich), 4% fetal calf serum (FCS, Sigma), 0.25% bovine serum albumin fraction IV (BSA, Fisher Scientific), HBSS (Gibco Life Technologies), and rotated at 37 °C overnight. After digestion, skin preparations were diluted with DMEM containing 8% FCS, filtered over 100 µm and 70 µm strainers (Falcon), washed twice, and taken up in HBSS supplemented with 0.5% BSA, pulmozyme (40 µg/ml final, Roche) and the indicated antibodies (**Supplementary Table 2**). After staining for 30 min at 4°C, samples were washed and filtered over a 30 µm strainer (Miltenyi Biotec). To exclude dead cells, samples were stained with 4',6-Diamidino-2-Phenylindole Dihydrochloride (DAPI, Sigma-Aldrich). Barcode-labeled skin-resident CD8$^+$ memory T cells were sorted on a FACSAria II (BD Biosciences) or FACSAria Fusion (BD Biosciences). Typical yields were 1,000-10,000 GFP$^+$ CD8$^+$ cells per leg.

Harvested spleen and LN tissue of individual mice was mashed through a 70 µm strainer into a single cell suspension and pooled with matched blood sample. This pooled cell pool, referred to as the circulatory compartment, was treated with NH_4Cl to remove erythrocytes, stained with the indicated antibodies (**Supplementary Table 2**), and GFP$^+$CD8$^+$ cells were then isolated by cell sorting on a MoFLo Astrios (Beckman Coulter), with typical yields of 1,000-10,000 GFP$^+$CD8$^+$ cells per mouse. Following isolation, sorted cells derived from either the skin or circulatory compartment were lysed in DirectPCR Lysis Reagent (Viagen Biotech), supplemented with 0.4 mg/ml Proteinase K (Sigma), and resulting samples were stored at -20 °C.

Analysis of the presence of blood borne T cells in the skin T_{EFF} pool

To determine the fraction of blood borne T cells in skin preparations of the vaccination site obtained during the effector phase, splenocytes of GFP-OT-I transgenic mice were first negatively enriched with the Mouse CD8 T Lymphocyte Enrichment Set (BD Biosciences).

Subsequently, C57BL/6J-Ly5.1 animals received ~700 naïve GFP-OT-I splenocytes intravenously (i.v.), followed by primary DNA vaccination on Veet-depilated (Reckitt Benckiser) hind legs as described above. At day 10 post vaccination, mice received a one-time injection of $1.5x10^6$ $CD8^+$ negatively enriched mTmG-OT-I splenocytes as a reference for blood borne T cells, 5 minutes prior to sacrificing the animals. Subsequently, blood and skin tissue was harvested and cells were isolated from the two compartments, as described above. Single cells suspensions were then stained with IR-dye (Thermo Fisher Scientific) and analyzed on an LSR II SORP (BD Biosciences).

Barcode DNA amplification and Next Generation Sequencing

Genomic DNA was isolated from frozen pellets of effector blood samples using DNeasy Blood and Tissue (Qiagen) for downstream PCR. Sorted samples of lymphoid tissues and from skin were lysed in DirectPCR Lysis Reagent (Viagen Biotech). Products of samples in experiments in which all samples contained more than approximately 3,000 barcode-labeled T cells were used for PCR amplification without intermediate steps. To enhance barcode recovery in experiments with samples with a lower GFP^+ cell count, barcode sequences were first captured from the obtained genomic DNA (gDNA) preparations, utilizing biotinylated DNA capture oligos that anneal either 5' or 3' of the barcode sequence in the GFP gDNA (for oligo sequences, see **Supplementary Table 1**). If at least one sample in an experiment contained <3000 GFP^+ cells, all samples in that experiment (independent of their GFP^+ cell count) were subjected to the barcode gDNA capture protocol, to avoid the possible generation of bias by this procedure. In brief, gDNA was sheared on the ME220 Focused-ultrasonicator (Covaris) under the following conditions: time: 20 sec; peak power:70; duty% 20; cycles/burst:1000. Next, sheared gDNA was denatured and mixed 1:1 with hybridization buffer (1 ml composition: 667.6 µl 20x SSPE (Gibco); 267.6 µl 50x Denhardt's solution (Sigma-Aldrich); 13.2 µl 20% SDS (Sigma-Aldrich); 26.8 µl 0.5M EDTA pH 8.0; 26.8 µl water supplemented with the biotinylated Capt_For_BClibv2 (50 fmol) and Capt_Rev_BClibv2 (50 fmol) oligos). Hybridization with biotinylated capture oligos was performed overnight at 65 °C. The following day, Streptavidin beads (Dynabeads^{tm} MyOne^{tm} streptavidin T1, Invitrogen) were washed with 2x B&W buffer (2M NaCl in TE buffer, pH 8.0) in low retention microtubes (Axygen) that were pre-rinsed with 400 ml 10 mM Tris, pH 8.0 solution, and the hybridized gDNA was mixed with the streptavidin beads for 30 min at room temperature. Subsequently, bead-bound gDNA was isolated by magnetic pull down using the Dynamag-2 magnet (Invitrogen). The isolated biotinylated gDNA beads were sequentially washed once with 500 µl 1x B&W buffer (diluted in TE buffer pH 8.0), 200 µl 0.5x B&W buffer (diluted in Tris buffer pH 8.0), 200 µl 0.25x B&W buffer (diluted in Tris buffer pH 8.0), and twice with 200 µl 10 mM Tris buffer (pH 8.0). The bead-bound gDNA was directly used for downstream PCR amplification.

All samples were split into two separate technical replicates prior to the first PCR amplification. Genomic barcodes were amplified by nested PCRs using Taq polymerase

(Invitrogen). First, the barcode sequence was amplified using the Top_Lib and Bot_Lib primers (30 cycles: 15 sec 95 °C; 30 sec 57.2 °C; 30 sec 72 °C). Subsequently, PCR products were subjected to a second amplification (30 cycles: 15 sec 95 °C; 30 sec 57.2 °C; 30 sec 72 °C) using the BC1v2_DS_For and BC1v2_DS_Rev primers that share the annealing sites of the Top_lib and Bot_lib primer respectively, but are tailed with sequences representing the Illumina primer annealing sites. Finally, the resulting PCR products were subjected to a third amplification (15 cycles: 15 sec 95 °C; 30 sec 57.2 °C; 30 sec 72 °C) using the P5_For and P7_index_Rev primers that are tailed with the P5 or P7 adaptors, respectively. The P7_index_Rev primer harbors a unique 7 bp index sequence that allows multiplexed analysis of up to 144 samples on one sequencing lane. Used 7 bp indexes had a Levenshtein distance of at least 3 bp from each other to avoid incorrect assignment of reads, due to PCR or sequence errors (Faircloth and Glenn, 2012). The final PCR products of individual samples were pooled, 322 bp fragments were purified using E-gel extraction (Invitrogen), and PCR products were sequenced on a HiSeq2500 Illumina platform with a read length of 65 bp. For primer sequences, see **Supplementary Table 1**.

Filtering of sequencing data

The reads obtained after sequencing were mapped to the barcode reference library, and reads that showed a 100% match to the barcode constant region, an index sequence that corresponded to 1 of the indices used during the PCR amplification, and a full match to one of the 21 bp barcode sequences listed in the reference library were retained. Using these filtering steps, approximately 150 – 190 million reads (75% - 95% of total reads) were considered of appropriate quality for downstream analysis.

To determine barcode sampling efficiency in biological samples, reproducibility between technical replicates was analyzed and biological samples were excluded from further analysis when the spearman correlation coefficient between technical replicates was <0.7. Next, barcodes that were not detected in both technical replicates were excluded, removing on average 0.66% of the total reads (and hence inferred cell fraction) per biological sample. After removal of non-reproducibly detected barcodes, the normalized read counts of the barcodes detected in the two technical replicates were averaged. As an additional noise filtering step, all barcodes that represented less than 0.01% of reads per sample were excluded. Finally, read counts were renormalized to 10,000, yielding values that represent relative T cell clone sizes in the biological samples. Data filtering and downstream analysis were performed in the software package R version 3.6.0 (Planting of a Tree, https://www.r-project.org/).

Deep-sequencing data analysis after filtering

To allow the visualization of clones with a read count of 0 on a log scale, read counts of all clones were plotted as read count+1, but original read count values were used for all calculations. Correlations between samples were calculated over the barcodes that were

shared between the two compared samples, using the spearman rank correlation. For data visualization, software package R (ggplot2) and Graphpad Prism 7.03 were used.

All ratios were calculated as: Readcount $_{SampleA}$/ Readcount $_{SampleB}$, taking the inverse of this ratio in case "Readcount $_{SampleA}$" was lower than "Readcount $_{SampleB}$", ensuring all outcomes are ≥1. Non-shared barcodes were excluded from the ratio calculations.

To determine the clonal bias threshold described in **Fig. 2 D**, technical replicate samples of all biological samples used in **Fig. 2** were used, with barcodes with a normalized read count of <0.5 excluded from the analysis. For all remaining barcodes, the ratio in read counts between technical replicate A and B was calculated, and a threshold was established such that 98% of barcodes detected in all technical replicates would have a ratio lower than this threshold (plotted in **Supplementary Fig. 2 C**). This resulted in a clonal bias threshold of 4.8, indicating that a clone had to contribute at least 4.8 times more to one of the normalized cell compartments than to the other cell compartment to be considered biased. Biased clones that were only detected in either the T_{CIRCM} or T_{RM} compartment cannot be ascribed a read count ratio. To allow for the visualization of these clones in **Fig. 2 E**, we applied the formula: (clone size T_{RM} – clone size T_{CIRCM})/(clone size T_{RM} + clone size T_{CIRCM}), resulting in values that range from -1 to 1, with -1 being completely biased toward T_{CIRCM} formation and 1 being completely biased toward T_{RM} formation.

To allow statistical analysis of the magnitude of clonal disparity between different combinations of cell compartments, an additional measurement of disparity was established (applied in **Fig. 3 C** and **5 C**). Specifically, to compare the magnitude of the differences between sample A and two other samples (i.e. A – B versus A – C), all barcodes observed in samples A, B and C were ranked in descending order based on the normalized read counts observed in sample A (reference sample), taking along shared and non-shared barcodes detected in the biological samples. Next, the cumulative read count of the ordered barcodes in sample A was plotted against the cumulative read counts in sample A (providing a reference curve) and against the cumulative read counts in samples B and C (**Supplementary Fig. 3 A**). The level of disparity was then determined by calculating the area between the reference curve and the curves obtained for samples B and C. In this analysis, a value of 0 signifies that samples are fully identical with respect to clonal composition, and a value of 0.5 signifies a complete lack of overlap between samples.

Modeling stochastic survival of memory T cells

To model the composition of a memory T cell pool that is purely formed by the stochastic survival of T_{EFF} cells, random *in silico* sampling of barcodes detected in the effector cell pool present in peripheral blood was conducted (**Fig. 3 D, E**). Specifically, to mimic stochastic memory formation, the probability of a clone surviving was considered to be directly proportional to its relative contribution to the effector pool (i.e. if a clone represented 50% of the total T_{EFF} pool, the probability of its offspring to be sampled per draw would be 0.5). *In*

silico modeling of the memory pool of 4 mice was performed using the following conditions: 1). By drawing a number of cells that was equal to the number of experimentally observed T_{RM} and T_{CIRCM} cells; 2). By drawing a number of cells that was equal to a fraction 0.1 of the number of experimentally observed T_{RM} and T_{CIRCM} cells; 3). By drawing a number of cells that was equal to the number of experimentally observed barcodes in the T_{RM} and T_{CIRCM} pool. The first setting models a situation in which the memory compartment is derived from the effector compartment without any further proliferation. The second setting models a situation in which the memory compartment is formed by a combination of cell death and expansion. The third scenario represents the most extreme bottleneck scenario in which each barcode observed in a memory compartment would be derived from a single cell that survived following the effector phase. Notably, for the second and third setting we assumed that the final T_{RM} pool is formed by proliferation of the drawn founder pool, and that during this expansion the hierarchy between founder clones does not alter. For the three settings, sampling was performed 1,000 times with replacement. To measure the resemblance of the modeled memory pool with the experimentally observed effector pool, Spearman correlations were calculated over the relative sizes of all clones, and were compared to the correlation between the experimentally observed effector pool and experimentally observed memory pool.

Statistics

Statistical analyses were performed using the two-tailed Mann-Whitney U-test and Spearman correlation test, using R (freely available at www.r-project.org) and Graphpad (Prism 7.03). Results were regarded as statistically significant at a P-value of <0.05, with * $P<0.05$, ** $P<0.005$ and ***$P<0.0005$.

Data availability

Data supporting the findings of this study are available from the corresponding author upon request.

Code availability

Codes were written in R and are available from the corresponding author upon request.

REFERENCES

Agarwal, P., Raghavan, A., Nandiwada, S. L., Curtsinger, J. M., Bohjanen, P. R., Mueller, D. L. & Mescher, M. F. 2009. Gene regulation and chromatin remodeling by IL-12 and type I IFN in programming for CD8 T cell effector function and memory. *J Immunol,* 183**,** 1695-704.

Ahrends, T., Babala, N., Xiao, Y., Yagita, H., Van Eenennaam, H. & Borst, J. 2016. CD27 Agonism Plus PD-1 Blockade Recapitulates CD4+ T-cell Help in Therapeutic Anticancer Vaccination. *Cancer Res,* 76**,** 2921-31.

Ahrends, T., Spanjaard, A., Pilzecker, B., Babala, N., Bovens, A., Xiao, Y., Jacobs, H. & Borst, J. 2017. CD4(+) T Cell Help Confers a Cytotoxic T Cell Effector Program Including Coinhibitory Receptor Downregulation and Increased Tissue Invasiveness. *Immunity,* 47**,** 848-861 e5.

Alemany, A., Florescu, M., Baron, C. S., Peterson-Maduro, J. & Van Oudenaarden, A. 2018. Whole-organism clone tracing using single-cell sequencing. *Nature,* 556**,** 108-112.

Araki, K., Turner, A. P., Shaffer, V. O., Gangappa, S., Keller, S. A., Bachmann, M. F., Larsen, C. P. & Ahmed, R. 2009. mTOR regulates memory CD8 T-cell differentiation. *Nature,* 460**,** 108-12.

Ariotti, S., Beltman, J. B., Chodaczek, G., Hoekstra, M. E., Van Beek, A. E., Gomez-Eerland, R., Ritsma, L., Van Rheenen, J., Maree, A. F., Zal, T., De Boer, R. J., Haanen, J. B. & Schumacher, T. N. 2012. Tissue-resident memory CD8+ T cells continuously patrol skin epithelia to quickly recognize local antigen. *Proc Natl Acad Sci U S A,* 109**,** 19739-44.

Beltman, J. B., Urbanus, J., Velds, A., Van Rooij, N., Rohr, J. C., Naik, S. H. & Schumacher, T. N. 2016. Reproducibility of Illumina platform deep sequencing errors allows accurate determination of DNA barcodes in cells. *BMC Bioinformatics,* 17**,** 151.

Bins, A. D., Jorritsma, A., Wolkers, M. C., Hung, C. F., Wu, T. C., Schumacher, T. N. & Haanen, J. B. 2005. A rapid and potent DNA vaccination strategy defined by in vivo monitoring of antigen expression. *Nat Med,* 11**,** 899-904.

Buchholz, V. R., Flossdorf, M., Hensel, I., Kretschmer, L., Weissbrich, B., Graf, P., Verschoor, A., Schiemann, M., Hofer, T. & Busch, D. H. 2013. Disparate individual fates compose robust CD8+ T cell immunity. *Science,* 340**,** 630-5.

Casey, K. A., Fraser, K. A., Schenkel, J. M., Moran, A., Abt, M. C., Beura, L. K., Lucas, P. J., Artis, D., Wherry, E. J., Hogquist, K., Vezys, V. & Masopust, D. 2012. Antigen-independent differentiation and maintenance of effector-like resident memory T cells in tissues. *J Immunol,* 188**,** 4866-75.

Cui, W. & Kaech, S. M. 2010. Generation of effector CD8+ T cells and their conversion to memory T cells. *Immunol Rev,* 236**,** 151-66.

Faircloth, B. C. & Glenn, T. C. 2012. Not all sequence tags are created equal: designing and validating sequence identification tags robust to indels. *PLoS One,* 7**,** e42543.

Feinerman, O., Veiga, J., Dorfman, J. R., Germain, R. N. & Altan-Bonnet, G. 2008. Variability and robustness in T cell activation from regulated heterogeneity in protein levels. *Science,* 321**,** 1081-4.

Frost, E. L., Kersh, A. E., Evavold, B. D. & Lukacher, A. E. 2015. Cutting Edge: Resident Memory CD8 T Cells Express High-Affinity TCRs. *J Immunol,* 195**,** 3520-4.

Gaide, O., Emerson, R. O., Jiang, X., Gulati, N., Nizza, S., Desmarais, C., Robins, H., Krueger, J. G., Clark, R. A. & Kupper, T. S. 2015. Common clonal origin of central and resident memory T cells following skin immunization. *Nat Med,* 21**,** 647-53.

Gebhardt, T., Wakim, L. M., Eidsmo, L., Reading, P. C., Heath, W. R. & Carbone, F. R. 2009. Memory T cells in nonlymphoid tissue that provide enhanced local immunity during infection with herpes simplex virus. *Nat Immunol,* 10**,** 524-30.

Gerlach, C., Rohr, J. C., Perie, L., Van Rooij, N., Van Heijst, J. W., Velds, A., Urbanus, J., Naik, S. H., Jacobs, H., Beltman, J. B., De Boer, R. J. & Schumacher, T. N. 2013. Heterogeneous differentiation patterns of individual CD8+ T cells. *Science,* 340**,** 635-9.

Gerlach, C., Van Heijst, J. W., Swart, E., Sie, D., Armstrong, N., Kerkhoven, R. M., Zehn, D., Bevan, M. J., Schepers, K. & Schumacher, T. N. 2010. One naive T cell, multiple fates in CD8+ T cell differentiation. *J Exp Med,* 207**,** 1235-46.

Halford, W. P., Balliet, J. W. & Gebhardt, B. M. 2004. Re-evaluating natural resistance to herpes simplex virus type 1. *J Virol,* 78**,** 10086-95.

Hendriks, J., Xiao, Y., Rossen, J. W., Van Der Sluijs, K. F., Sugamura, K., Ishii, N. & Borst, J. 2005. During viral infection of the respiratory tract, CD27, 4-1BB, and OX40 collectively determine formation of CD8+ memory T cells and their capacity for secondary expansion. *J Immunol,* 175**,** 1665-76.

Herndler-Brandstetter, D., Ishigame, H., Shinnakasu, R., Plajer, V., Stecher, C., Zhao, J., Lietzenmayer, M., Kroehling, L., Takumi, A., Kometani, K., Inoue, T., Kluger, Y., Kaech, S. M., Kurosaki, T., Okada, T. & Flavell, R. A. 2018. KLRG1(+) Effector CD8(+) T Cells Lose KLRG1, Differentiate into All Memory T Cell Lineages, and Convey Enhanced Protective Immunity. *Immunity,* 48**,** 716-729 e8.

Iborra, S., Martinez-Lopez, M., Khouili, S. C., Enamorado, M., Cueto, F. J., Conde-Garrosa, R., Del Fresno, C. & Sancho, D. 2016. Optimal Generation of Tissue-Resident but Not Circulating Memory T Cells during Viral Infection Requires Crosspriming by DNGR-1(+) Dendritic Cells. *Immunity,* 45**,** 847-860.

Jiang, X., Clark, R. A., Liu, L., Wagers, A. J., Fuhlbrigge, R. C. & Kupper, T. S. 2012. Skin infection generates non-migratory memory CD8+ T(RM) cells providing global skin immunity. *Nature,* 483**,** 227-31.

Joshi, N. S., Cui, W., Chandele, A., Lee, H. K., Urso, D. R., Hagman, J., Gapin, L. & Kaech, S. M. 2007. Inflammation directs memory precursor and short-lived effector CD8(+) T cell fates via the graded expression of T-bet transcription factor. *Immunity,* 27**,** 281-95.

Kaech, S. M. & Wherry, E. J. 2007. Heterogeneity and cell-fate decisions in effector and memory CD8+ T cell differentiation during viral infection. *Immunity,* 27**,** 393-405.

Kalhor, R., Kalhor, K., Mejia, L., Leeper, K., Graveline, A., Mali, P. & Church, G. M. 2018. Developmental barcoding of whole mouse via homing CRISPR. *Science,* 361.

Levenshtein, V. 1966. Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doclady* 10**,** 707–10.

Mackay, L. K., Braun, A., Macleod, B. L., Collins, N., Tebartz, C., Bedoui, S., Carbone, F. R. & Gebhardt, T. 2015a. Cutting edge: CD69 interference with sphingosine-1-phosphate receptor function regulates peripheral T cell retention. *J Immunol,* 194**,** 2059-63.

Mackay, L. K., Rahimpour, A., Ma, J. Z., Collins, N., Stock, A. T., Hafon, M. L., Vega-Ramos, J., Lauzurica, P., Mueller, S. N., Stefanovic, T., Tscharke, D. C., Heath, W. R., Inouye, M., Carbone, F. R. & Gebhardt, T. 2013. The developmental pathway for CD103(+)CD8+ tissue-resident memory T cells of skin. *Nat Immunol,* 14**,** 1294-301.

Mackay, L. K., Wynne-Jones, E., Freestone, D., Pellicci, D. G., Mielke, L. A., Newman, D. M., Braun, A., Masson, F., Kallies, A., Belz, G. T. & Carbone, F. R. 2015b. T-box Transcription Factors Combine with the Cytokines TGF-beta and IL-15 to Control Tissue-Resident Memory T Cell Fate. *Immunity,* 43**,** 1101-11.

Marchingo, J. M., Prevedello, G., Kan, A., Heinzel, S., Hodgkin, P. D. & Duffy, K. R. 2016. T-cell stimuli independently sum to regulate an inherited clonal division fate. *Nat Commun,* 7**,** 13540.

Masopust, D., Choo, D., Vezys, V., Wherry, E. J., Duraiswamy, J., Akondy, R., Wang, J., Casey, K. A., Barber, D. L., Kawamura, K. S., Fraser, K. A., Webby, R. J., Brinkmann, V., Butcher, E. C., Newell, K. A. & Ahmed, R. 2010. Dynamic T cell migration program provides resident memory within intestinal epithelium. *J Exp Med,* 207**,** 553-64.

Mckenna, A., Findlay, G. M., Gagnon, J. A., Horwitz, M. S., Schier, A. F. & Shendure, J. 2016. Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science,* 353**,** aaf7907.

Milner, J. J., Toma, C., Yu, B., Zhang, K., Omilusik, K., Phan, A. T., Wang, D., Getzler, A. J., Nguyen, T., Crotty, S., Wang, W., Pipkin, M. E. & Goldrath, A. W. 2017. Runx3 programs CD8(+) T cell residency in non-lymphoid tissues and tumours. *Nature,* 552**,** 253-257.

Mousavi, S. F., Soroosh, P., Takahashi, T., Yoshikai, Y., Shen, H., Lefrancois, L., Borst, J., Sugamura, K. & Ishii, N. 2008. OX40 costimulatory signals potentiate the memory commitment of effector CD8+ T cells. *J Immunol,* 181**,** 5990-6001.

Mueller, S. N. & Mackay, L. K. 2016. Tissue-resident memory T cells: local specialists in immune defence. *Nat Rev Immunol,* 16**,** 79-89.

Muschaweckh, A., Buchholz, V. R., Fellenzer, A., Hessel, C., Konig, P. A., Tao, S., Tao, R., Heikenwalder, M., Busch, D. H., Korn, T., Kastenmuller, W., Drexler, I. & Gasteiger, G. 2016. Antigen-dependent competition shapes the local repertoire of tissue-resident memory CD8+ T cells. *J Exp Med,* 213**,** 3075-3086.

Obar, J. J. & Lefrancois, L. 2010. Early events governing memory CD8+ T-cell differentiation. *Int Immunol,* 22**,** 619-25.

Oosterhuis, K., Aleyd, E., Vrijland, K., Schumacher, T. N. & Haanen, J. B. 2012. Rational design of DNA vaccines for the induction of human papillomavirus type 16 E6- and E7-specific cytotoxic T-cell responses. *Hum Gene Ther,* 23**,** 1301-12.

Parameswaran, N., Suresh, R., Bal, V., Rath, S. & George, A. 2005. Lack of ICAM-1 on APCs during T cell priming leads to poor generation of central memory cells. *J Immunol,* 175**,** 2201-11.

Sarkar, S., Kalia, V., Haining, W. N., Konieczny, B. T., Subramaniam, S. & Ahmed, R. 2008. Functional and genomic profiling of effector CD8 T cell subsets with distinct memory fates. *J Exp Med,* 205**,** 625-40.

Scholer, A., Hugues, S., Boissonnas, A., Fetler, L. & Amigorena, S. 2008. Intercellular adhesion molecule-1-dependent stable interactions between T cells and dendritic cells determine CD8+ T cell memory. *Immunity,* 28**,** 258-70.

Sheridan, B. S., Pham, Q. M., Lee, Y. T., Cauley, L. S., Puddington, L. & Lefrancois, L. 2014. Oral infection drives a distinct population of intestinal resident memory CD8(+) T cells with enhanced protective function. *Immunity,* 40**,** 747-57.

Smith, N. L., Patel, R. K., Reynaldi, A., Grenier, J. K., Wang, J., Watson, N. B., Nzingha, K., Yee Mon, K. J., Peng, S. A., Grimson, A., Davenport, M. P. & Rudd, B. D. 2018. Developmental Origin Governs CD8(+) T Cell Fate Decisions during Infection. *Cell,* 174**,** 117-130 e14.

Sowell, R. T., Rogozinska, M., Nelson, C. E., Vezys, V. & Marzo, A. L. 2014. Cutting edge: generation of effector cells that localize to mucosal tissues and form resident memory CD8 T cells is controlled by mTOR. *J Immunol,* 193**,** 2067-71.

Spanjaard, B., Hu, B., Mitic, N., Olivares-Chauvet, P., Janjuha, S., Ninov, N. & Junker, J. P. 2018. Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. *Nat Biotechnol,* 36**,** 469-473.

Steinert, E. M., Schenkel, J. M., Fraser, K. A., Beura, L. K., Manlove, L. S., Igyarto, B. Z., Southern, P. J. & Masopust, D. 2015. Quantifying Memory CD8 T Cells Reveals Regionalization of Immunosurveillance. *Cell,* 161**,** 737-49.

Turner, D. L., Bickham, K. L., Thome, J. J., Kim, C. Y., D'ovidio, F., Wherry, E. J. & Farber, D. L. 2014. Lung niches for the generation and maintenance of tissue-resident memory T cells. *Mucosal Immunol,* 7**,** 501-10.

Weeks, B. S., Ramchandran, R. S., Hopkins, J. J. & Friedman, H. M. 2000. Herpes simplex virus type-1 and -2 pathogenesis is restricted by the epidermal basement membrane. *Arch Virol,* 145**,** 385-96.

Zehn, D., Lee, S. Y. & Bevan, M. J. 2009. Complete but curtailed T-cell response to very low-affinity antigen. *Nature,* 458**,** 211-4.

Acknowledgements

We would like to thank the Genomics Core facility, in particular Arno Velds and Marja Nieuwland, the Flow cytometry facility, and Animal research facility of the Netherlands Cancer Institute for technical support, M. Hoekstra for illustrations and J. Borst for providing the OVA-vaccination plasmid. We are grateful to all members of the Schumacher group for scientific input. This work was supported by ERC AdG grant Life-His-T (to T.N.S.).

Author information

Lianne Kok and Feline E Dijkgraaf contributed equally to this work.

Contributions

L.K. and F.E.D. designed and performed experiments. L.K. analyzed lineage-tracing data. J.U. designed and produced the barcode library and performed capture experiments. K.B. assisted in the design of analysis methods. R.F.C. contributed to design and execution of $HSV_{TOM-OVA}$ experiments. D.W.V. and F.E.D. developed the T_{RM} isolation protocol. L.P. wrote the DNA barcode filtering script. J.B.B generated the barcode reference list. L.K., F.E.D. and T.N.S. contributed to experimental design and prepared the manuscript with input of all co-authors.

Competing interests

The authors declare no competing financial interests.

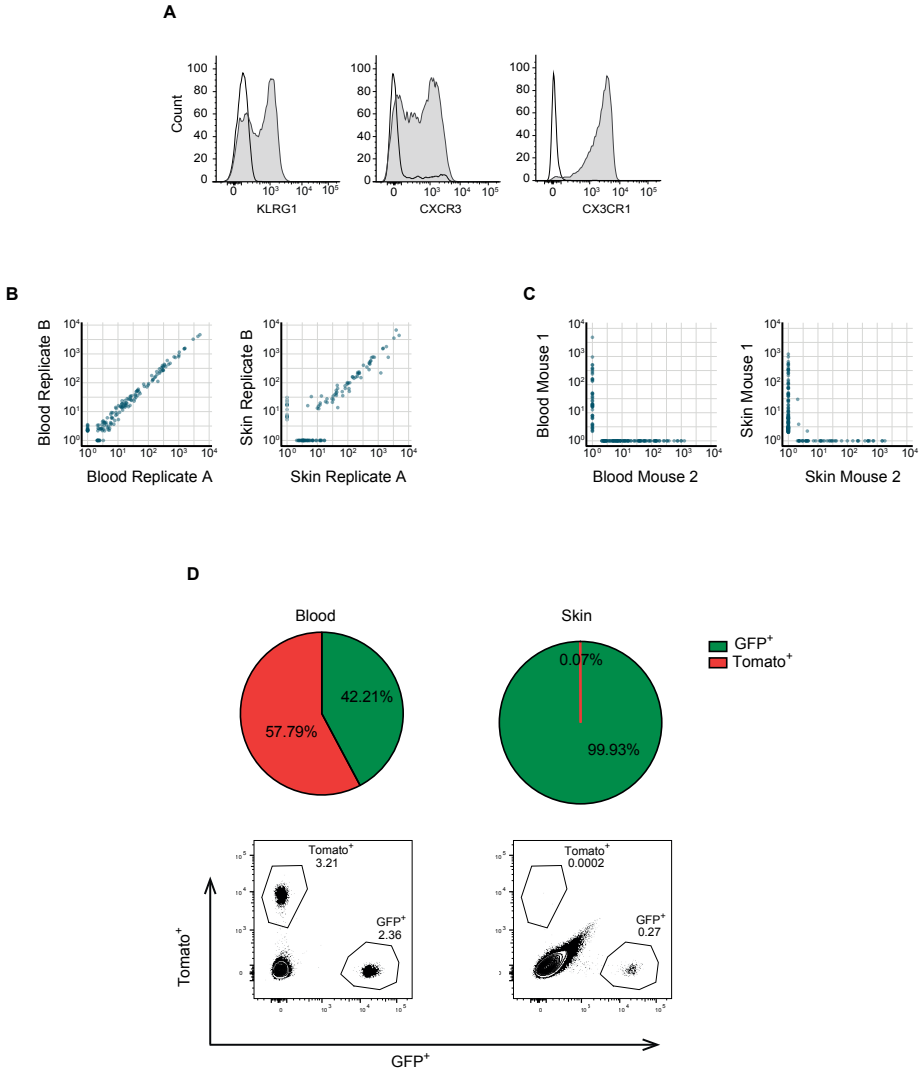
SUPPLEMENTARY INFORMATION

Supplementary Table 1 | DNA sequences. Oligo-DNA and primer sequences used to generate the barcode library and to PCR-amplify barcode sequences from biological samples.

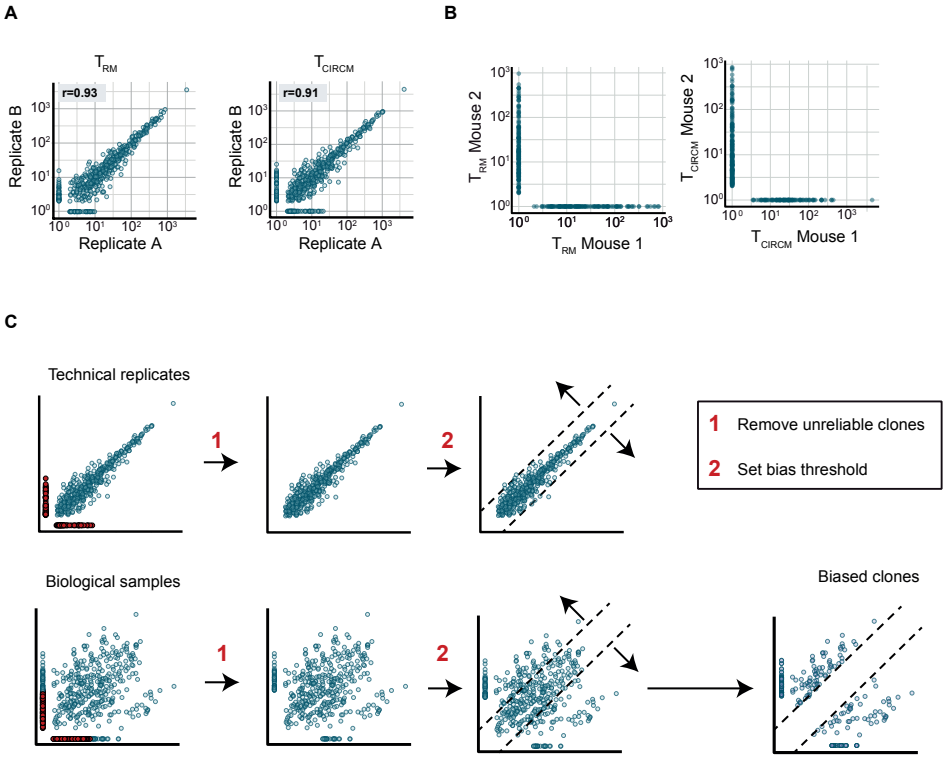
Name	Sequence (5' – 3')
BC1DS_lib_oligo	aagcttttgctgccgtcaactagaacactcgagatcagnnnnnnnnnnnnnnnnnnnnnatgtggtatgatgtatc atctgagtagctagttccgcaactaccttcagggacctggtgctgttcatcctactcttgatcttgaattcatagatatc-gtctgtactcagttgtcactttcat
BC1DS_lib_rev	atgaaagtgacaactgagtacagacgatat
Capt_For_BClibv2	ccactacctgagcacccagtccgccctgagcaaagaccccaacgagaagcgcgatcacatggtcctgctg-gagttcgtgaccgccgccgggatcactctc
Capt_Rev_BClibv2	ctagcttgccaaacctacaggtggggtctttcattccccccttttttctggagactaaataaaatcttttattttatcgtcgac-cactgtgctggcggccg
Top_lib	tgctgccgtcaactagaaca
Bot_lib	gatctcgaatcaggcgctta
BC1v2_DS_For	acactctttccctacacgacgctcttccgatctnnnnctagaacactcgagatcag
BC1v2_DS_Rev	gtgactggagttcagacgtgtgctcttccgatcgatctcgaatcaggcgctta
P5_For	aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatct
P7_index_Rev	caagcagaagacggcatacgagatnnnnnnngtgactggagttcagacgtgtgctcttccgatc

Supplementary Table 2 | Antibodies used for flow cytometry. Top: single cell suspensions of skin samples. Bottom: single cell suspensions of circulation samples.

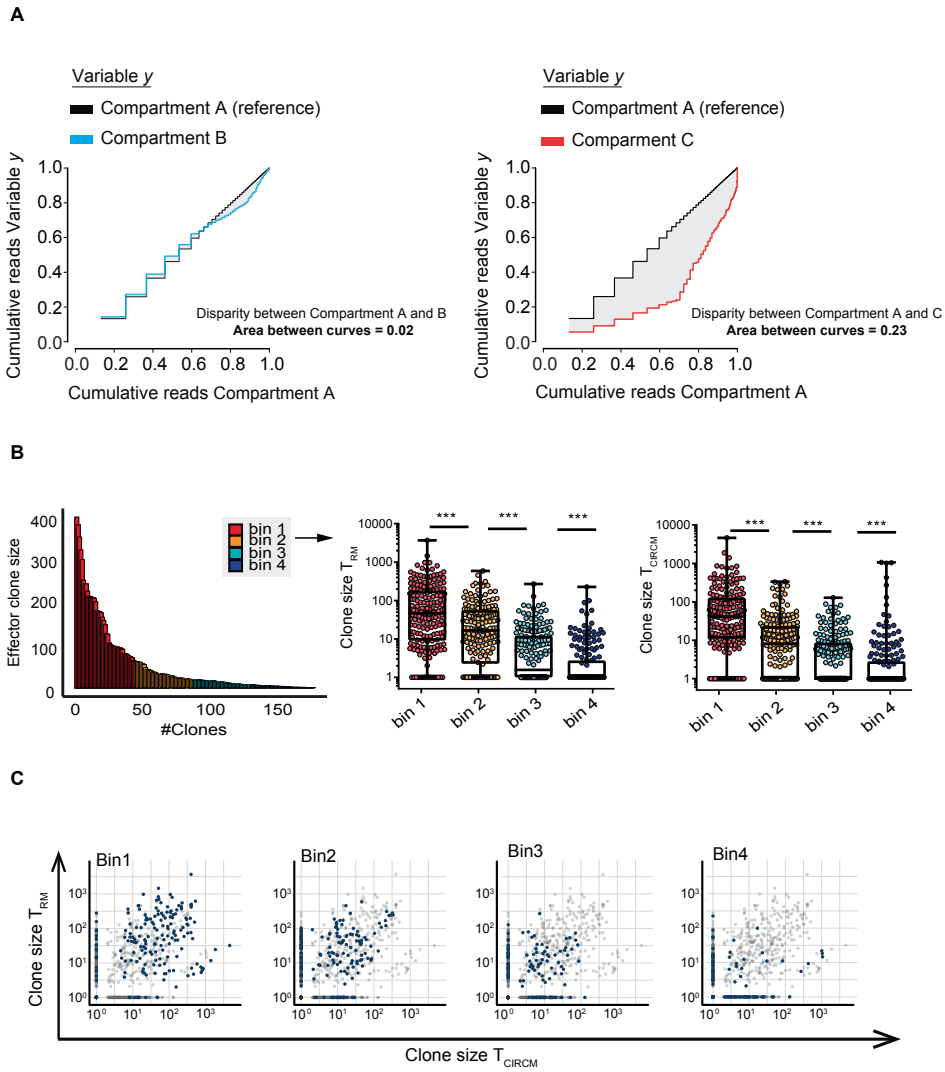
Antibody	Clone	Catalogue #	Company
anti-CD8beta-PeCy7	eBioH35-17.2	25-0083-82	Thermo Fisher Scientific
anti-CD69-PE	H1.2F3	12-0691-81	Thermo Fisher Scientific BD Biosciences
anti-CD103-PerCP-Cy5.5	2E7	121415	BioLegend
anti-CD103-BV711	M290	564320	BD Biosciences
anti-CD8alpha-Percp-Cy5.5	53-6.7	55162	BD Biosciences
anti-CD62L-PE	Mel-14	12-0621-83	eBioscience
anti-CD27-APC	LG.7F9	17-0271-81	eBioscience
anti-KLRG1-APC	2F1/KLRG1	138411	BioLegend
anti-CX3CR1-BV421	SA011F11	149023	BioLegend
anti-CXCR3-PE	CXCR3-173	126505	BioLegend



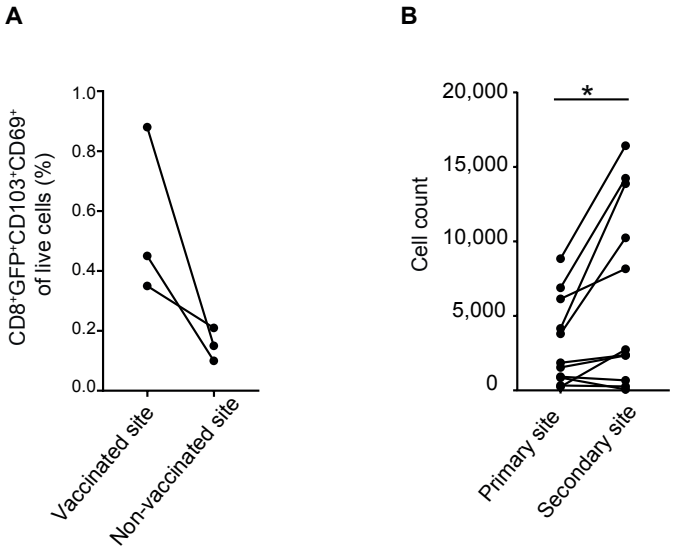
Supplementary Figure 1 | Quality of barcode quantification and analysis of blood borne T cell contamination in effector phase skin samples. A, **B**, **C**, Recipients of barcode labeled T cells were vaccinated and whole blood and organs were harvested at d12 post first vaccination. **A**, Analysis of effector marker expression on barcode labeled OT-I T cells (gray) and endogenous $CD8^{hi}$ T cells (transparent) present in blood. Data from one mouse, representative of eleven mice, are depicted. **B**, Measured clone sizes detected in representative technical replicates of blood (left) and skin (right) samples. **C**, Measured clone sizes detected in blood (left) and skin (right) of independent mice. Dots represent individual clones. **D**, Analysis of the presence of blood borne T cells in skin preparations. Recipients of GFP^+ OT-I T cells were DNA vaccinated and then received $1.5x10^6$ $Tomato^+$ cells 5 min prior to sacrifice, at day 10 post vaccination. (Top) Pie charts depicting the relative percentage of GFP^+ and $Tomato^+$ cells in blood (left) and skin (right) preparations. (Bottom) representative flow cytometry plots. Cells are gated on live lymphocytes. Data are representative of four mice.



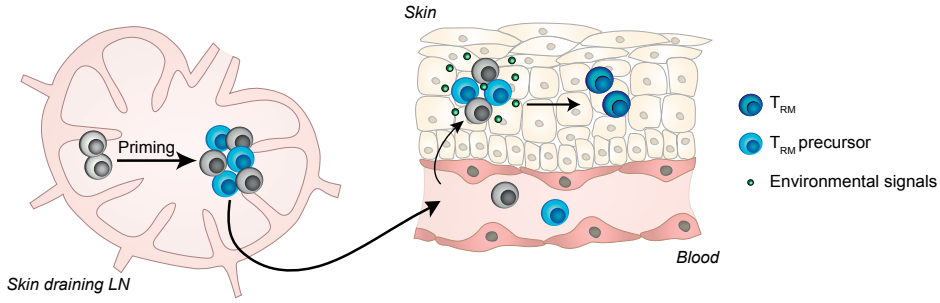
Supplementary Figure 2 | Quality control and analyses of the barcode-labeled T_{RM} and T_{CIRCM} compartment. A, Measured clone sizes detected in technical replicates of T_{RM} (left) and T_{CIRCM} (right) samples derived from the mice described in Fig. 2B. Spearman correlation r was calculated over clones that were detected in both technical replicates A (left): P<0.0005, A (right): P<0.0005. **B**, Measured clone sizes detected in T_{RM} (left) and T_{CIRCM} (right) of independent mice described in Fig. 2B. **C**, Step by step description of the strategy used to filter biological data and to define biased clones, as depicted in Fig. 2D. First, unreliably detected clones (indicated in red) are removed. Second, a bias threshold (dashed lines) is set, such that 98% of the clones in technical replicates fall below this threshold. This threshold is subsequently applied to the biological data to identify clones with a bias in output that goes beyond the variation that occurs because of technical noise. Clones that contribute >4.8 times to one sample than to the other are considered biased. A-C, Dots represent individual clones.



Supplementary Figure 3 | Remodeling of the skin-resident and the circulating memory compartment. A, Example plots depicting the strategy to determine the disparity between two cellular compartments, as applied in Fig. 3C and Fig. 5C. Disparity of compartment B and C to compartment A can be assessed by plotting the fraction of cumulative reads of clones in compartment B and C, which are ordered based on their size (largest to smallest) in compartment A (y-axis), to the cumulative reads of the ordered clones in compartment A (x-axis). Area between the compartment A reference curve and compartment B (left) and C (right) curves is calculated to generate a measure of disparity. **B** (left), Illustration of the subdivision of ordered effector-stage T cell clones (large to small) into 4 bins, with each bin containing 25% of all observed clones. B (middle, right), Quantitative contribution of binned clones detected in effector blood to the T_{RM} and T_{CIRCM} compartment. Median with whiskers representing mix/max, ***P< 0.0005, Mann-Whitney U test. **C**, Relative contribution of T_{EFF} clones in bin 1-4 to the T_{RM} and T_{CIRCM} compartment. B, C, Representative data from two independent experiments, dots represent individual clones.



Supplementary Figure 4 | *De novo* T_{RM} generation upon generation upon secondary vaccination in previously unperturbed sites. Mice received GFP+ OT-I T cells (**A**) or barcode-labeled OT-I T cells (**B**) and were subjected to DNA vaccination on the right hind leg, while the other hind leg remained unperturbed. **A**, Analysis of the number of T_{RM} detected in the vaccinated (right leg) and non-vaccinated (left leg) skin site >60 days after vaccination. **B**, >60 days post primary vaccination, the non-vaccinated (left leg) skin site was subjected to DNA vaccination and >60 days after secondary vaccination, the primary and secondary vaccinated skin sites were harvested and GFP+ T_{RM} at the two sites were quantified. B, Number of barcode-labeled T_{RM} detected at the primary and secondary skin vaccination site of nine mice. *P<0.05, Wilcoxon signed-rank test. A, B, Dots represent individual mice.



Supplementary Figure 5 | Model of tissue-resident memory CD8+ T cell differentiation. During priming in the skin-draining lymph node, naïve CD8+ T cells become activated, undergo clonal expansion and a selection of responding T cells develops a heightened capacity to form tissue-resident memory T cells. In the effector phase, these putative T_{RM} precursors migrate alongside non-T_{RM} precursors from the blood into the inflamed skin. At this site, T_{RM} precursors display a heightened capacity to mature into long-term persisting T_{RM} cells in response to tissue-derived external factors such as TGFβ, IL-15 and antigen. Note that formation of T_{RM} precursor cells may either occur early during clonal expansion, as depicted here, or may reflect heterogeneity in T cell potential that already exists prior to antigen stimulation.