

# Chromatin states define tumour-specific T cell dysfunction and reprogramming

Mary Philip<sup>1</sup>, Lauren Fairchild<sup>2,3</sup>, Liping Sun<sup>4</sup>, Ellen L. Horste<sup>1</sup>, Steven Camara<sup>1</sup>, Mojdeh Shakiba<sup>1,5</sup>, Andrew C. Scott<sup>1,5</sup>, Agnes Viale<sup>4</sup>, Peter Lauer<sup>6</sup>, Taha Merghoub<sup>5,7</sup>, Matthew D. Hellmann<sup>5,8</sup>, Jedd D. Wolchok<sup>5,7,9</sup>, Christina S. Leslie<sup>2</sup> & Andrea Schietinger<sup>1,5</sup>

**Tumour-specific CD8 T cells in solid tumours are dysfunctional, allowing tumours to progress. The epigenetic regulation of T cell dysfunction and therapeutic reprogrammability (for example, to immune checkpoint blockade) is not well understood. Here we show that T cells in mouse tumours differentiate through two discrete chromatin states: a plastic dysfunctional state from which T cells can be rescued, and a fixed dysfunctional state in which the cells are resistant to reprogramming. We identified surface markers associated with each chromatin state that distinguished reprogrammable from non-reprogrammable PD1<sup>hi</sup> dysfunctional T cells within heterogeneous T cell populations from tumours in mice; these surface markers were also expressed on human PD1<sup>hi</sup> tumour-infiltrating CD8 T cells. Our study has important implications for cancer immunotherapy as we define key transcription factors and epigenetic programs underlying T cell dysfunction and surface markers that predict therapeutic reprogrammability.**

Tumour-specific CD8 T cells (TST) are often found within solid tumours, but tumours progress despite their presence, suggesting that these TST are dysfunctional<sup>1</sup>. The clinical success of immune checkpoint blockade (for example, PD1/PDL1- and CTLA4-blocking antibodies) and adoptive T cell therapy in a subset of patients with cancer demonstrates the great potential of TST<sup>2</sup>; however, important questions remain, including how to predict which patients will respond to therapy and precisely which TST mediate clinical responses<sup>3–5</sup>. Moreover, an unmet need is the development of interventions for tumours that are refractory to checkpoint blockade despite having ample TST infiltration.

We previously demonstrated that in the early stages of tumorigenesis, TST become non-responsive, exhibiting the phenotypic, functional, and transcriptional features of tumour-reactive tumour-infiltrating lymphocytes (TIL) from late-stage human solid tumours<sup>6</sup>. TST dysfunction is initially reversible but ultimately becomes irreversible, even after removal of dysfunctional T cells from the tumour microenvironment and multiple rounds of cell division<sup>6</sup>. We hypothesized that this heritable, signal-independent dysfunctional state is epigenetically imprinted. The epigenetic programs that regulate normal differentiation of innate and adaptive lymphocytes have been described<sup>7–10</sup>. However, the epigenetic programs regulating T cell differentiation and dysfunction in tumours are not known. In this study, we used the assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq)<sup>11</sup> to assess genome-wide chromatin accessibility changes during T cell differentiation in tumours compared to acute infection.

## CD8 T cell chromatin changes during infection

We transferred congenically marked naive (N; CD44<sup>lo</sup>CD62L<sup>hi</sup>) TCR<sub>TAG</sub> cells (specific for SV40 large T antigen epitope I (TAG))<sup>12</sup> from TCR<sub>TAG</sub> transgenic mice into wild-type C57BL/6 mice, which were immunized one day later with a recombinant *Listeria monocytogenes* strain expressing TAG (LmTAG)<sup>6,13</sup>. TCR<sub>TAG</sub> cells were re-isolated,

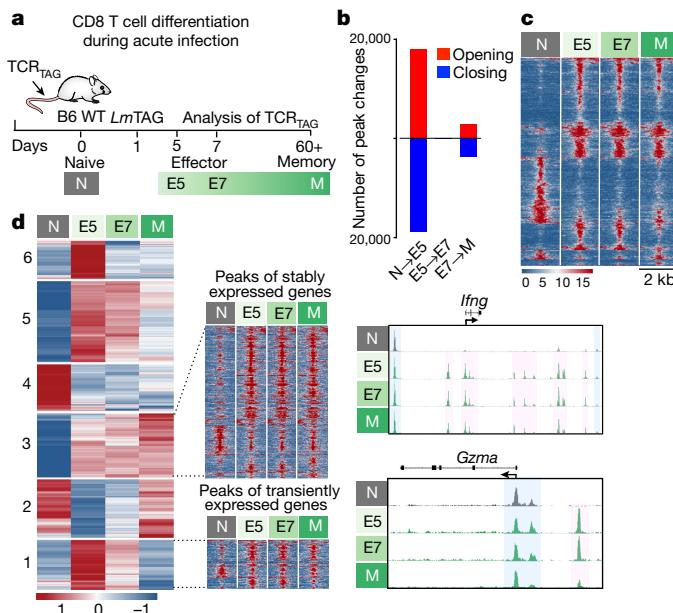
phenotypically and functionally characterized, and underwent ATAC-seq and RNA-seq at 5, 7 (effectors; E5, E7) and 60+ days (memory; M) after immunization (Fig. 1a). N, E5, E7, and M expressed characteristic activation, homing and cytokine receptors (CD44, CD62L, IL7R), transcription factors (TBET), cytotoxic molecules (GZMB, CD107), and pro-inflammatory cytokines (IFN $\gamma$ , TNF $\alpha$ ) (Extended Data Fig. 1).

ATAC-seq libraries generated from N, E5, E7, and M showed the expected distribution of fragment lengths (Extended Data Fig. 2). Using DESeq2 (ref. 14) to assess differential chromatin accessibility, we found that substantial chromatin remodelling occurred as cells differentiated from the N to the effector state (E5), with much less remodelling from E5 to E7 and E7 to M (Fig. 1b, c, Extended Data Fig. 3a). In N, effector gene loci such as *Prf1* and *Tnf* shared highly accessible chromatin and basal transcriptional activity with E5/E7 and M (Extended Data Fig. 3b), consistent with activating histone marks previously shown at these loci in naïve T cells<sup>15,16</sup>.

We analysed accessibility changes during the N to E5 transition in loci associated with early and late TCR-response genes, as defined by the Immunological Genome Project<sup>17</sup>. Early-response genes showed much fewer changes compared to late-response genes (Extended Data Fig. 3c). For example, *Ldha* (encoding LDHA, needed for the metabolic shift to aerobic glycolysis and IFN $\gamma$  production<sup>18</sup>) and *Mki67* (encoding KI67, required for chromosome segregation during mitosis<sup>19</sup>) require no change in chromatin accessibility to be rapidly induced after TCR stimulation (Extended Data Fig. 3d).

Memory T cells exhibit more rapid and robust effector function upon antigen re-encounter compared to naïve T cells<sup>20</sup>. K-means clustering of RNA expression patterns (Fig. 1d, left) revealed two trends: transient gene activation or downregulation in E5/E7 but not M (clusters 1, 2, 5, 6), and stable gene activation or downregulation in E5, E7, and M (clusters 3 and 4). In contrast, chromatin accessibility for these loci was largely similar in E5/E7 and M (Fig. 1d, middle). Thus, the ‘effector-like’

<sup>1</sup>Immunology Program, Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA. <sup>2</sup>Computational Biology Program, Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA. <sup>3</sup>Tri-Institutional Training Program in Computational Biology and Medicine, Weill Cornell Medical College, New York, New York 10065, USA. <sup>4</sup>Integrated Genomics Operation, Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA. <sup>5</sup>Weill Cornell Medical College, Cornell University, New York, New York 10065, USA. <sup>6</sup>Aduro Biotech, Inc., Berkeley, California 94720, USA. <sup>7</sup>Melanoma and Immunotherapeutics Service, Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA. <sup>8</sup>Thoracic Oncology Service, Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA. <sup>9</sup>Ludwig Center for Cancer Immunotherapy, Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA.



**Figure 1 | CD8 T cell chromatin state dynamics during acute infection.**

a, Experimental scheme. b, Number of chromatin peak accessibility changes during each transition (false discovery rate (FDR) < 0.05). c, Chromatin accessibility heat map grouped by differential accessibility patterns. Each row represents one of 8,654 selected peaks (differentially accessible between at least one sequential cell comparison; FDR < 0.05,  $\log_2(\text{fold change}) > 2$ ). d, Left, K-means clustered ( $K = 6$ , row-normalized) RNA-seq data for 1,758 differentially expressed genes ( $\log_2(\text{FC}) > 1$ , FDR < 0.05, base mean  $\log_2(\text{expression}) \geq 10$ ). Middle, heat map of differentially accessible peaks (FDR < 0.05,  $\log_2(\text{FC}) > 1$ ) presented as in c for genes in K-means clusters 1 and 3. Right, ATAC-seq signal profiles across *Ifng* and *Gzma* loci. Peaks present in all differentiation states highlighted in blue, activation-induced peaks in pink.

accessibility in M permits basal transcription of certain effector genes (cluster 3) such as *Ifng*, whereas other genes are transcriptionally silent but poised for rapid re-expression upon TCR activation (cluster 1, *Gzma*) (Fig. 1d, right).

### Chromatin state dynamics of TST dysfunction

We next assessed chromatin-state dynamics in TST over the course of tumorigenesis using the previously described tamoxifen-inducible, autochthonous liver cancer model (AST-Cre-ER<sup>T2</sup>) in which TAG is a tumour-specific antigen<sup>6</sup>. AST-Cre-ER<sup>T2</sup> mice initially develop pre-malignant lesions which eventually progress into hepatocellular carcinoma (by day 60–90)<sup>6</sup>. We transferred congenically marked naïve TCR<sub>TAG</sub> cells (N, the same as N in Fig. 1a) into AST-Cre-ER<sup>T2</sup> mice one day before tamoxifen administration and then analysed TCR<sub>TAG</sub> cells at different time points (Fig. 2a). Liver-infiltrating TCR<sub>TAG</sub> cells downregulated CD62L, uniformly expressed activation markers CD44 and inhibitory receptors PD1 and LAG3, and failed to produce IFN $\gamma$  or TNF $\alpha$  (Fig. 2b). Massive chromatin remodelling occurred by day 5, followed by a second wave of remodelling between days 7 and 14 (Fig. 2c, d, Extended Data Fig. 4a). Notably, after the second wave, few accessibility changes occurred, even after progression to established tumours at day 60+ (Fig. 2c–e). Thus, TST differentiated through two discrete chromatin states: an initial dysfunctional state 1 (L5, L7), and later dysfunctional state 2, established by day 14 and persisting thereafter. Many of the ATAC-seq peaks that were gained or lost were in intronic and intergenic regions (potential enhancer peaks), whereas peaks present in all CD8 T cells were in promoter regions (Extended Data Fig. 4b, bottom); this pattern was also seen in functional CD8 T cell differentiation (Extended Data Fig. 4b, top).

TCR<sub>TAG</sub> cells in malignant lesions followed a distinct epigenetic trajectory compared to TCR<sub>TAG</sub> cells in acute infection (L5 versus

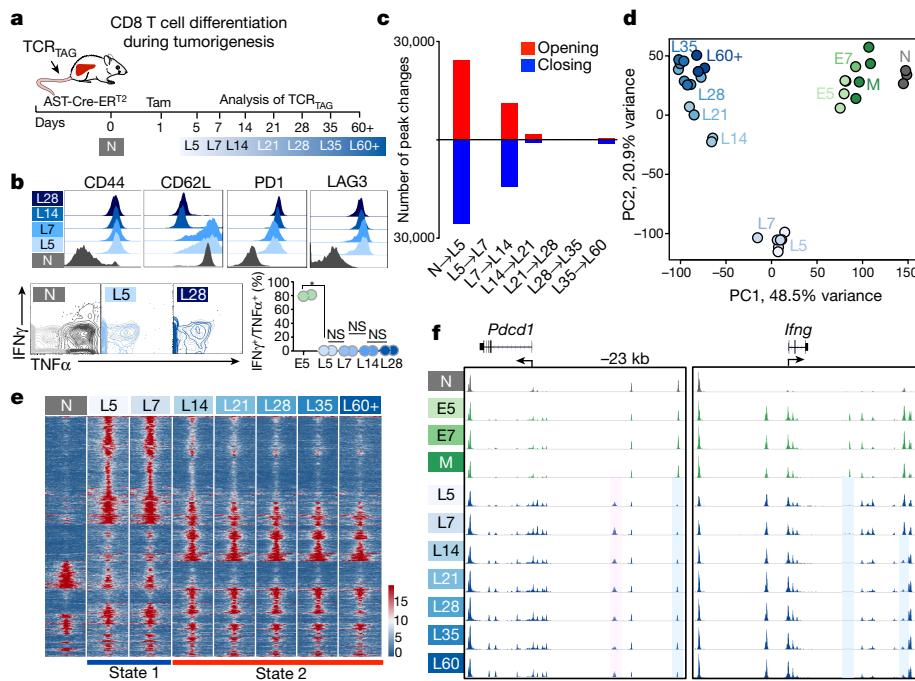
E5; Fig. 2d), and many peak changes were unique to either the early dysfunctional (L5) or functional (E5) state (Extended Data Fig. 4c) and were found in genes involved in TCR signalling and cytokine production pathways (Extended Data Fig. 4d). Enhancer peaks in the *Ifng* locus that opened during normal effector differentiation were inaccessible in dysfunctional TCR<sub>TAG</sub> cells (Fig. 2f, right). An intergenic peak near (~23.8 kb) the PD1-encoding *Pdcd1* locus was uniquely accessible in L5 to L60+, but not in E5/E7 and M (Fig. 2f, left); a similar peak was described in exhausted T cells in chronic viral infection<sup>21–23</sup>. We tested whether accessibility of potential transcription factor targets changed preferentially during differentiation from N to L5 as compared to N to E5 (Extended Data Fig. 5a). Predicted NFATC1-binding sites, including those in genes encoding inhibitory receptors and negative regulators, such as *Ctla4*, *Pdcd1*, *Tigit*, *Socs1*, and *Cblb* and transcription factors *Egr1* and *Egr2*, had increased peak accessibility in dysfunctional L5 (Extended Data Fig. 5a, b). NFAT transcription factor family members, particularly NFATC1 and NFATC2, are important regulators of T cell development and function (reviewed in ref. 24), as well as exhaustion in chronic viral infections<sup>25</sup>. Although some genes with increased NFATC1 peak accessibility in L5 showed immediate transient transcriptional activation, others were activated later (Extended Data Fig. 5c). Transcription factor footprints (nucleotide sequences protected from Tn5 transposase-mediated adapter insertion) were detected in accessible motif sites for NFATC1 as well as other transcription factors (Extended Data Fig. 6a).

### Chromatin states correlate with reprogrammability

Notably, the discrete chromatin states in dysfunctional TCR<sub>TAG</sub> cells correlated temporally with our previous observation that L8 but not L35 were capable of regaining effector function<sup>6</sup>. Indeed, when we re-isolated TCR<sub>TAG</sub> cells from liver lesions and cultured them *in vitro* with IL-15 (Fig. 3a), previously shown to induce proliferation and restore effector function in tumour-reactive CD8 T cells<sup>26,27</sup>, L5 and L7 regained the ability to produce IFN $\gamma$  and TNF $\alpha$ , but TCR<sub>TAG</sub> cells isolated at day 12 and after did not (Fig. 3a). Thus state 1 dysfunction is plastic, but with further chromatin remodelling between days 7 and 14, becomes fixed (state 2).

Chromatin peaks with TCF family motifs closed during the state 1 (L7) to state 2 (L14) transition, whereas E2F, ETS, and KLF family transcription factor motif-containing peaks opened (Fig. 3b). Indeed, TCF1 (encoded by *Tcf7*) protein levels decreased between L7 and L14 (Extended Data Fig. 6b), and analysis of closing peaks showed enrichment for WNT receptor signalling pathway genes, upstream of TCF family transcription factors, as well as cytokine response, TCR signalling, and T cell differentiation pathway genes (Extended Data Fig. 6c). Among the TCR signalling genes most upregulated during the L7–L14 transition were negative regulators such as *Cish1* and *Socs2*, whereas co-stimulatory molecule genes such as *Icos* and *Cd28* were downregulated together with closing of multiple peaks within their loci (Extended Data Fig. 6d).

We next used an *in vivo* pharmacologic strategy to test the role of NFAT and TCF in TST dysfunction. FK506 is an immunosuppressant that inhibits NFAT nuclear translocation and downstream gene activation<sup>28,29</sup>, and we used 25% of the full immunosuppression dose to partially downregulate NFAT activity without completely blocking T cell activation and/or effector function. TWS119, a GSK3 $\beta$  inhibitor, enhances differentiation of CD8 T cell to memory cells through WNT/TCF1 activation<sup>30</sup>, and we therefore treated TCR<sub>TAG</sub>-adoptively transferred AST-Cre-ER<sup>T2</sup> mice with FK506 alone or in combination with TWS119 (Extended Data Fig. 7a). Indeed, L10 TST from FK506 and FK506/TWS119-treated mice had decreased expression of the NFATC1 targets PD1 and LAG3, increased levels of TCF1 and EOMES (Extended Data Fig. 7b), and were more efficiently reprogrammable (Extended Data Fig. 7c) compared to controls or TWS119 alone (data not shown).



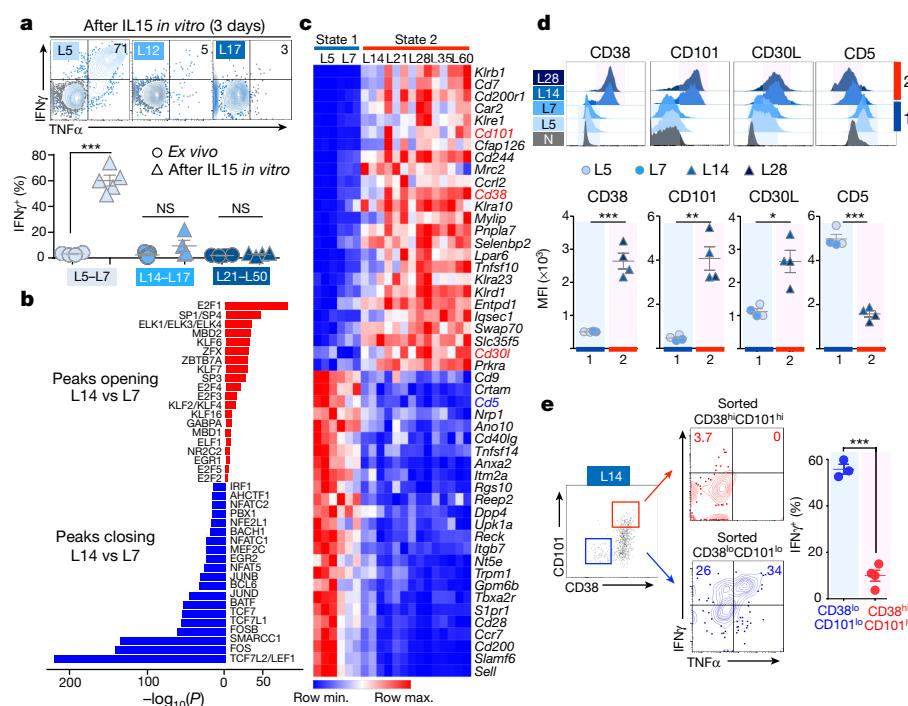
**Figure 2 | TST differentiate to a dysfunctional state in developing tumours through discrete chromatin states.** **a**, Experimental scheme. **b**, Immunophenotype and cytokine production (grey, no peptide control) ( $n = 8$  total, with  $n = 2$  per time point). Each symbol represents an individual mouse. Representative of 5 independent experiments; mean  $\pm$  s.e.m. shown. \* $P = 0.0002$  (Student's  $t$ -test); NS, not statistically significant. **c**, Number of peak changes during each transition

### Surface proteins associated with chromatin states

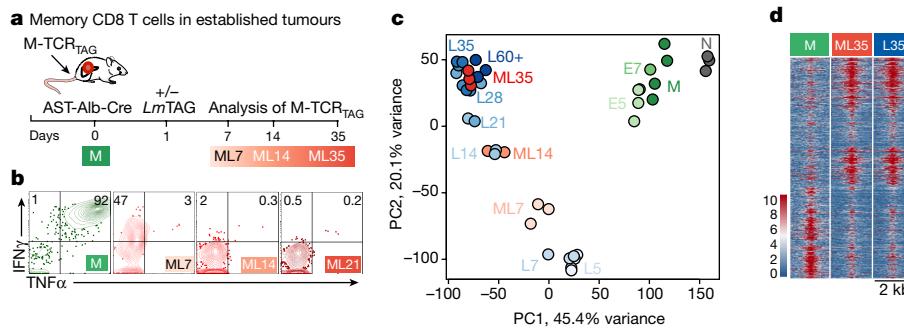
We next looked for cell surface proteins with expression that correlated with chromatin states 1 or 2 and thus might predict reprogrammability of heterogeneous TIL. PD1 and LAG3 were similarly expressed by both plastic (L5, L7) and fixed (L14+) dysfunctional TST (Fig. 2b) and thus not informative in this regard. We identified membrane protein genes differentially expressed between early (L5, L7) and late (L14 to L60+) dysfunctional TCR<sub>TAG</sub> cells (Fig. 3c) and found several markers not

(FDR < 0.05). **d**, Principal component analysis (PCA) of peak accessibility in naive TCR<sub>TAG</sub> cells (N; grey) during normal differentiation (green) and during tumorigenesis (blue). Each symbol represents a single biological replicate. **e**, Chromatin accessibility heat map (15,275 differentially accessible peaks as in Fig. 1c). **f**, ATAC-seq signal profiles across *Pdcd1* and *Ifng* loci. Peaks uniquely lost (blue) or gained (pink) in TST are highlighted.

previously associated with tumour-induced T cell dysfunction. State 1 (L5, L7) TCR<sub>TAG</sub> cells had low expression of CD38, CD101, and CD30L and high expression of CD5, whereas state 2 (L14, L28) TCR<sub>TAG</sub> cells had the opposite pattern (Fig. 3d). Consistent with its expression, the *Cd38* locus contained intergenic and intronic peaks uniquely accessible in state 2 TST (Extended Data Fig. 8a). TCF1 downregulation coincided with CD38 upregulation (Extended Data Fig. 8b), and other key regulators of CD8 T cell differentiation, such as IRF4 and BCL2, showed



**Figure 3 | Discrete chromatin states correlate with reprogrammability and surface protein expression profiles.** **a**, Top, cytokine production by L5, L12, and L17 after 3 days *in vitro* IL-15 culture (grey, no peptide control). Bottom, IFN $\gamma$  production *ex vivo* (circles) or after 3–4 days IL-15 *in vitro* culture (triangles). Pooled from three experiments. **b**, The 20 most significantly enriched transcription factor motifs in peaks opening (red) and closing (blue) between L7 and L14. **c**, RNA-seq expression (row-normalized) for the 50 most differentially expressed genes encoding membrane proteins. **d**, CD38, CD101, CD30L and CD5 expression; representative of 3 independent experiments. **e**, Cytokine production by sorted CD38 $^{hi}$ CD101 $^{lo}$  (blue) and CD38 $^{hi}$ CD101 $^{hi}$  (red) L14 after 3 days IL-15 *in vitro* culture. Similar data obtained with sorted L10 in independent experiment. Each symbol represents an individual mouse (**a**, **d**, **e**). Mean  $\pm$  s.e.m. shown; \* $P = 0.005$ , \*\* $P = 0.0005$ , \*\*\* $P \leq 0.0001$  (Student's  $t$ -test).



**Figure 4 | Memory TST rapidly enter the fixed dysfunctional chromatin state in established tumours.** **a**, Experimental scheme. **b**, Cytokine production of M cells isolated from liver tumours. **c**, PCA of peak accessibility in TCR<sub>TAG</sub> cells during acute infection (green), tumorigenesis (blue), and memory TCR<sub>TAG</sub> cells in established tumours (red). **d**, Chromatin accessibility heat map showing M cells, M cells re-isolated at day 35 from established hepatocellular carcinoma tumours (ML35), and naive TCR<sub>TAG</sub> cells isolated at day 35 (L35) from early malignant lesions (see Fig. 2). Each row represents one of 19,679 differentially accessible peaks as in Fig. 1c.

a similar binary expression in early and late TST (Extended Data Fig. 8c). Moreover, TCR<sub>TAG</sub> cells from FK506 and FK506/TWS119-treated mice expressed low CD38 and CD101 compared to controls, correlating with their improved reprogrammability (Extended Data Fig. 7d). To test whether these markers could identify reprogrammable T cells within a heterogeneous TST population, we sorted CD38<sup>lo</sup>CD101<sup>lo</sup> and CD38<sup>hi</sup>CD101<sup>hi</sup> TST from PD1<sup>hi</sup> L14 cells and assessed reprogrammability (3 days *in vitro* IL-15). CD38<sup>lo</sup>CD101<sup>lo</sup> L14 regained the ability to produce IFN $\gamma$  and TNF $\alpha$ , but CD38<sup>hi</sup>CD101<sup>hi</sup> L14 did not (Fig. 3e).

We determined whether these findings could be applied to other tumour histologies and/or T cell specificities by using mouse B16F10 (B16) melanoma cells expressing ovalbumin (B16-OVA), a model antigen recognized by OVA-specific OT1 CD8 T cells (TCR<sub>OT1</sub> cells). Naive congenically marked TCR<sub>OT1</sub> cells were adoptively transferred into B16-OVA tumour-bearing B6 mice. Tumour-infiltrating TCR<sub>OT1</sub> cells upregulated CD44, PD1, and LAG3, downregulated CD62L, and lost the ability to produce IFN $\gamma$  or TNF $\alpha$  (Extended Data Fig. 8d). At later stages, dysfunctional TCR<sub>OT1</sub> cells expressed high levels of CD38 and CD101 and downregulated CD5 compared to early dysfunctional day 5 TCR<sub>OT1</sub> cells (Extended Data Fig. 8e). Moreover, late dysfunctional TCR<sub>OT1</sub> cells at day 25 could not regain the ability to produce IFN $\gamma$  or TNF $\alpha$ , in contrast to early dysfunctional TCR<sub>OT1</sub> cells at day 5 (Extended Data Fig. 8f).

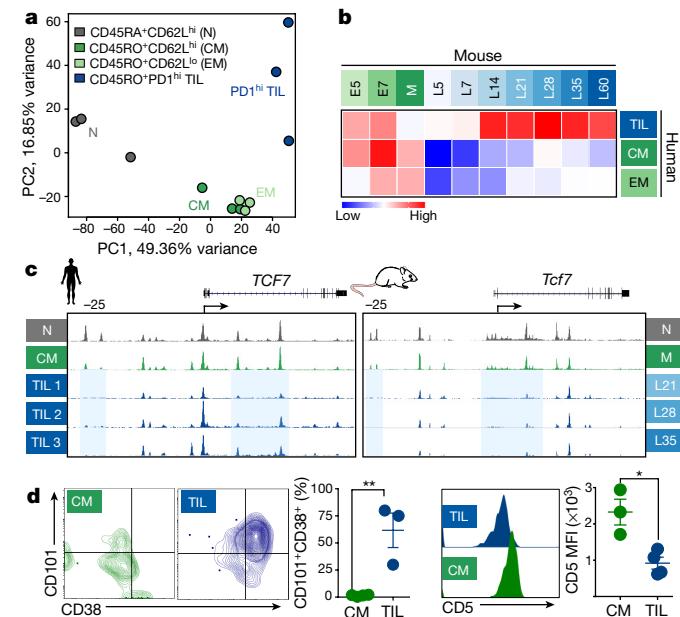
### Memory T cells enter state 2 dysfunction in tumours

We next tested whether the ‘functionally poised’ chromatin state present in memory T cells (M; Fig. 1) could prevent them from becoming dysfunctional in tumours. TCR<sub>TAG</sub> memory cells were transferred into AST-Alb-Cre mice (in which hepatocytes express TAG from birth<sup>6</sup>) bearing established hepatocellular carcinomas and, one day later, were immunized with LmTAG (Fig. 4a). By day 7, tumour-infiltrating memory T cells (ML7) rapidly upregulated PD1 and LAG3 and progressively lost effector function (Fig. 4b, Extended Data Fig. 9a). ATAC-seq revealed that M cells followed a similar epigenetic trajectory as the N cells in early malignant lesions (Fig. 4c, Extended Data Fig. 9b–d) and remarkably, by day 35, the chromatin state of transferred M cells was nearly identical to that of N at day 35 in early malignant lesions (ML35 and L35; Fig. 4d). Dysfunctional M cells displayed the same gain and loss of ATAC-seq peaks in critical gene loci including *Pdcd1*, *Ctla4*, *Cd38*, *Tcf7*, and *Ifng* (Extended Data Fig. 9e). Changes in surface protein expression (CD38, CD101, CD30L, and CD5) between ML7 and ML14 were like those seen with N (L7 and L14, respectively) (Extended Data Fig. 9f). We obtained similar results when LmTAG immunization after adoptive transfer was omitted (Extended Data Fig. 9b–d).

### Chromatin accessibility in human TIL

Finally, we examined chromatin states of human CD8 TIL and peripheral blood lymphocytes from healthy donors. We carried out

ATAC-seq on naive (N; CD45RA<sup>+</sup>CD45RO<sup>-</sup>), effector memory (EM; CD45RA<sup>-</sup>CD45RO<sup>+</sup>CD62L<sup>lo</sup>), and central memory (CM; CD45RA<sup>-</sup>CD45RO<sup>+</sup>CD62L<sup>hi</sup>) CD8 peripheral blood lymphocytes from healthy donors and PD1<sup>hi</sup> CD8 TIL isolated from human melanoma and non-small-cell lung cancer tumours (Extended Data Fig. 10a). Human N cells had a distinct chromatin state as compared to EM and CM, which were similar (Fig. 5a, Extended Data Fig. 10b), though distinct accessibility patterns in genes such as *SELL* (encoding CD62L) distinguished all three states (Extended Data Fig. 10c). PD1<sup>hi</sup> TIL uniquely gained and lost multiple peaks, for example in *IFNG*, *EGR2*, *CD5*, and *CTLA4* (Extended Data Fig. 10d). We compared the non-promoter peak changes that occurred during functional and dysfunctional mouse CD8 T cell differentiation with those observed in



**Figure 5 | Human tumour-infiltrating PD1<sup>hi</sup> CD8 T cells enter a similar chromatin accessibility state as dysfunctional state 2 TST in mice.** **a**, Peak accessibility PCA on human healthy donor peripheral blood lymphocytes and PD1<sup>hi</sup> TIL from melanoma and non-small-cell lung cancer tumours. **b**, For non-promoter peaks, normalized Spearman correlations of log<sub>2</sub>(FC) calculated between human N and EM, CM or PD1<sup>hi</sup> TIL versus log<sub>2</sub>(FC) between mouse N and E5, E7, M, and L5 to L60.  $P < 10^{-16}$  for all comparisons between human PD1<sup>hi</sup> TIL and mouse L14–L60. **c**, ATAC-seq signal profiles across human *TCF7* and mouse *Tcf7* gene loci; peaks lost in human PD1<sup>hi</sup> TIL and mouse L21, L28, L35 highlighted in blue. **d**, CD38, CD101 and CD5 expression on human CM (green) and PD1<sup>hi</sup> TIL (blue). Each symbol represents an individual healthy donor or patient. Mean  $\pm$  s.e.m. shown; \* $P = 0.01$ , \*\* $P = 0.006$  (Student’s *t*-test).

human peripheral blood lymphocytes and PD1<sup>hi</sup> TIL and found that human PD1<sup>hi</sup> TIL had the greatest overlap in peak accessibility changes with dysfunctional (state 2) mouse TST (Fig. 5b). For example, the *TCF7/Tcf7* locus showed similar intergenic and intronic peak accessibility changes in human PD1<sup>hi</sup> TIL and mouse state-2 TCR<sup>TAG</sup> cells (Fig. 5c). A subset of PD1<sup>hi</sup> TIL expressed higher levels of CD38 and CD101 and lower levels of CD5 (Fig. 5d), suggesting that these markers could potentially be used to identify T cells that are amenable to therapeutic reprogramming in human tumours.

## Discussion

In this study, we define the chromatin state dynamics underlying tumour-specific T cell dysfunction over the course of tumorigenesis. Naive TST encountering tumour antigen in pre-malignant lesions differentiated to an initially plastic, therapeutically reprogrammable chromatin state, then transitioned to a fixed dysfunctional chromatin state that did not undergo further remodelling, even with progression to large established tumours (Extended Data Fig. 10e). The rapid induction of dysfunction early during tumorigenesis without progression through an effector state resembles peripheral self-tolerance induction<sup>31,32</sup>. We identified core elements shared between mouse fixed dysfunctional TST and human PD1<sup>hi</sup> TIL. Surprisingly, memory TST differentiated to the same fixed dysfunctional chromatin state in tumours, suggesting that antigen exposure in tumours can overwrite pre-existing epigenetic programs regardless of the initial differentiation state.

We identified surface markers, including CD101 and CD38, which were associated with discrete dysfunctional chromatin states and demarcated reprogrammable from non-reprogrammable PD1<sup>hi</sup> T cells within heterogeneous TIL populations, a finding of important potential clinical relevance, and human PD1<sup>hi</sup> TIL showed heterogeneous expression of these markers. In patients who do not respond to immune checkpoint blockade (non-responders), PD1<sup>hi</sup> TIL may be in a fixed dysfunctional state, in contrast to responders whose PD1<sup>hi</sup> TIL are in a plastic state, amenable to reprogramming. Our studies on the epigenetic and transcriptional programs underlying TST dysfunctional states and therapeutic reprogrammability point to new targets and strategies to transform TST into potent anti-tumour agents.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 1 November 2016; accepted 12 April 2017.

Published online 17 May 2017.

- Hellström, I., Hellström, K. E., Pierce, G. E. & Yang, J. P. Cellular and humoral immunity to different types of human neoplasms. *Nature* **220**, 1352–1354 (1968).
- Khalil, D. N., Smith, E. L., Brentjens, R. J. & Wolchok, J. D. The future of cancer treatment: immunomodulation, CARs and combination immunotherapy. *Nat. Rev. Clin. Oncol.* **13**, 273–290 (2016).
- Snyder, A. et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* **371**, 2189–2199 (2014).
- Kelderman, S., Schumacher, T. N. & Haanen, J. B. Acquired and intrinsic resistance in cancer immunotherapy. *Mol. Oncol.* **8**, 1132–1139 (2014).
- Rizvi, N. A. et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348**, 124–128 (2015).
- Schietinger, A. et al. Tumor-specific T cell dysfunction is a dynamic antigen-driven differentiation program initiated early during tumorigenesis. *Immunity* **45**, 389–401 (2016).
- Zhang, J. A., Mortazavi, A., Williams, B. A., Wold, B. J. & Rothenberg, E. V. Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity. *Cell* **149**, 467–482 (2012).
- Scharer, C. D., Barwick, B. G., Youngblood, B. A., Ahmed, R. & Boss, J. M. Global DNA methylation remodeling accompanies CD8 T cell effector function. *J. Immunol.* **191**, 3419–3429 (2013).
- Russ, B. E. et al. Distinct epigenetic signatures delineate transcriptional programs during virus-specific CD8<sup>+</sup> T cell differentiation. *Immunity* **41**, 853–865 (2014).
- Shih, H. Y. et al. Developmental acquisition of regulomes underlies innate lymphoid cell functionality. *Cell* **165**, 1120–1133 (2016).

- Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
- Staveley-O'Carroll, K. et al. *In vivo* ligation of CD40 enhances priming against the endogenous tumor antigen and promotes CD8<sup>+</sup> T cell effector function in SV40 T antigen transgenic mice. *J. Immunol.* **171**, 697–707 (2003).
- Brockstedt, D. G. et al. *Listeria*-based cancer vaccines that segregate immunogenicity from toxicity. *Proc. Natl. Acad. Sci. USA* **101**, 13832–13837 (2004).
- Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
- Araki, Y., Fann, M., Wersto, R. & Weng, N. P. Histone acetylation facilitates rapid and robust memory CD8 T cell response through differential expression of effector molecules (eomesodermin and its targets: perforin and granzyme B). *J. Immunol.* **180**, 8102–8108 (2008).
- Denton, A. E., Russ, B. E., Doherty, P. C., Rao, S. & Turner, S. J. Differentiation-dependent functional and epigenetic landscapes for cytokine genes in virus-specific CD8<sup>+</sup> T cells. *Proc. Natl. Acad. Sci. USA* **108**, 15306–15311 (2011).
- Best, J. A. et al. Transcriptional insights into the CD8<sup>+</sup> T cell response to infection and memory T cell formation. *Nat. Immunol.* **14**, 404–412 (2013).
- Peng, M. et al. Aerobic glycolysis promotes T helper 1 cell differentiation through an epigenetic mechanism. *Science* **354**, 481–484 (2016).
- Cuylen, S. et al. Ki-67 acts as a biological surfactant to disperse mitotic chromosomes. *Nature* **535**, 308–312 (2016).
- Kaech, S. M. & Cui, W. Transcriptional control of effector and memory CD8<sup>+</sup> T cell differentiation. *Nat. Rev. Immunol.* **12**, 749–761 (2012).
- Sen, D. R. et al. The epigenetic landscape of T cell exhaustion. *Science* **354**, 1165–1169 (2016).
- Pauken, K. E. et al. Epigenetic stability of exhausted T cells limits durability of reinvigoration by PD-1 blockade. *Science* **354**, 1160–1165 (2016).
- Scott-Browne, J. P. et al. Dynamic changes in chromatin accessibility occur in CD8<sup>+</sup> T cells responding to viral infection. *Immunity* **45**, 1327–1340 (2016).
- Macian, F. NFAT proteins: key regulators of T-cell development and function. *Nat. Rev. Immunol.* **5**, 472–484 (2005).
- Martinez, G. J. et al. The transcription factor NFAT promotes exhaustion of activated CD8<sup>+</sup> T cells. *Immunity* **42**, 265–278 (2015).
- Teague, R. M. et al. Interleukin-15 rescues tolerant CD8<sup>+</sup> T cells for use in adoptive immunotherapy of established tumors. *Nat. Med.* **12**, 335–341 (2006).
- Li, Y. et al. MART-1-specific melanoma tumor-infiltrating lymphocytes maintaining CD28 expression have improved survival and expansion capability following antigenic restimulation *in vitro*. *J. Immunol.* **184**, 452–465 (2010).
- Flanagan, W. M., Corthésy, B., Bram, R. J. & Crabtree, G. R. Nuclear association of a T-cell transcription factor blocked by FK-506 and cyclosporin A. *Nature* **352**, 803–807 (1991).
- Jain, J. et al. The T-cell transcription factor NFATp is a substrate for calcineurin and interacts with Fos and Jun. *Nature* **365**, 352–355 (1993).
- Gattinoni, L. et al. Wnt signaling arrests effector T cell differentiation and generates CD8<sup>+</sup> memory stem cells. *Nat. Med.* **15**, 808–813 (2009).
- Schietinger, A., Delrow, J. J., Basom, R. S., Blattman, J. N. & Greenberg, P. D. Rescued tolerant CD8 T cells are preprogrammed to reestablish the tolerant state. *Science* **335**, 723–727 (2012).
- Waugh, K. A. et al. Molecular profile of tumor-specific CD8<sup>+</sup> T cell hypofunction in a transplantable murine cancer model. *J. Immunol.* **197**, 1477–1488 (2016).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank J. van der Veeken, V. Krisnawan, Y. Pritykin, and B. Gasmi for discussions and technical support; S. Reiner for discussions; and the MSKCC Flow Cytometry Core and Animal Facility, T.M., M.H., J.D.W., and A.S. are members of the Parker Institute for Cancer Immunotherapy. This work was supported by NIH-NCI grants R00 CA172371 (to A.S.), K08 CA158069 (to M.P.), and U54 CA209975 (to C.S.L. and A.S.), NHGRI grant U01 HG007893 (to C.S.L.), V Foundation for Cancer Research (to A.S.), the William and Ella Owens Medical Research Foundation (to A.S.), the Josie Robertson Young Investigator Award (to A.S.), and the MSKCC Core Grant P30 CA008748. The Integrated Genomics Operation Core was supported by Cycle for Survival and the Marie-Josée and Henry R. Kravis Center for Molecular Oncology.

**Author Contributions** M.P. and A.S. conceived and designed the study, carried out experiments, and analysed and interpreted data. L.F. designed and performed all high-throughput computational analyses; C.S.L. designed and supervised computational analyses; E.H., S.C., M.S., and A.C.S. assisted with experiments; L.S. and A.V. performed ATAC-seq; P.L. generated *Listeria* strains; and T.M., M.H., and J.D.W. provided human samples. M.P. and A.S. wrote the manuscript, with all authors contributing to writing and providing feedback.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to A.S. (schietia@mskcc.org).

**Reviewer Information** *Nature* thanks J. Wherry and the other anonymous reviewer(s) for their contribution to the peer review of this work.

## METHODS

**Mice.** TCR<sub>TAG</sub> transgenic mice (B6.Cg-Tg(TcraY1,TcrbY1)416Tev/J)<sup>12</sup>, Cre-ER<sup>T2</sup> (B6.129-Gt(ROSA)26Sor<sup>tm1(cre/ERT2)Ty/J</sup>/J), Alb-Cre (B6.Cg-Tg(Alb-cre)21Mgn/J), TCR-OT1 (C57BL/6-Tg(TcraTcrb)1100Mjb/J), Ly5.1 (B6.SJL-Ptprc<sup>a</sup> Pepc<sup>b</sup>/BoyJ), and C57BL/6J Thy1.1 mice were purchased from The Jackson Laboratory. TCR<sub>TAG</sub> mice were crossed to Thy1.1 mice to generate TCR<sub>TAG</sub> Thy1.1 mice. TCR-OT1 were crossed to Ly5.1 mice to generate TCR-OT1 Ly5.1 mice. AST (Albumin-floxStop-SV40 large T antigen (TAG))<sup>33</sup> were crossed to Cre-ER<sup>T2</sup> or Alb-Cre mice to obtain AST-Cre-ER<sup>T2</sup> and AST-Alb-Cre mice, respectively<sup>6</sup>. Both female and male mice were used for studies. Mice were age- and sex-matched and between 1.5–3 months old when used for experiments. Animals were assigned randomly to experimental groups. All mice were bred and maintained in the animal facility at Memorial Sloan Kettering Cancer Center (MSKCC). Experiments were performed in compliance with the MSKCC Institutional Animal Care and Use Committee (IACUC) regulations.

**Antibodies and reagents.** Fluorochrome-conjugated antibodies were purchased from BD Biosciences, eBioscience, Biolegend, and Cell Signaling Technology. Tamoxifen (Sigma) stock solution was prepared by warming tamoxifen in 1 ml sterile corn oil at 50 °C for 15 min, then further diluted in corn oil to obtain the stock concentration (5 mg ml<sup>-1</sup> in corn oil). A single dose of tamoxifen (1 mg) was administered intraperitoneally (i.p.) into AST-Cre-ER<sup>T2</sup> mice.

**Intracellular cytokine staining.** Intracellular cytokine staining was performed using the Cytofix/Cytoperm Plus kit (BD Biosciences) per the manufacturer's instructions. In brief, T cells were mixed with 2 × 10<sup>6</sup> congenically marked splenocytes and incubated with Tag-I peptide (0.5 µg ml<sup>-1</sup>) or OVA peptide (0.1 µg ml<sup>-1</sup>) for 4–5 h at 37 °C in the presence of GolgiPlug (brefeldin A). After staining for cell-surface molecules, the cells were fixed, permeabilized, and stained with antibodies against IFN $\gamma$  (XMG1.2) and TNF $\alpha$  (MP6-XT22).

**Flow cytometric analysis.** Flow cytometric analysis was performed using Fortessa and LSR FACS analysers (BD Biosciences); cells were sorted using BD FACS Aria (BD Biosciences) at the MSKCC Flow Core Facility. Flow data were analysed with FlowJo v. 10 software (Tree Star Inc.).

**Listeria infection.** The *Listeria monocytogenes* (*Lm*)  $\Delta$ actA  $\Delta$ inlB strain<sup>13</sup> expressing the Tag-I epitope (SAINNYAQKL, SV40 large T antigen<sup>206–215</sup>) was generated by Aduro Biotech as previously described<sup>34</sup>. Experimental vaccination stocks were prepared by growing bacteria to early stationary phase, washing in phosphate buffered saline, formulated at approximately 1 × 10<sup>10</sup> colony-forming units (c.f.u.) ml<sup>-1</sup>, and stored at –80 °C. Mice were infected i.p. with 5 × 10<sup>6</sup> c.f.u. of *Lm*TAG.

**Adoptive T cell transfer.** For the generation of effector and memory TCR<sub>TAG</sub> CD8<sup>+</sup> T cells, 10<sup>5</sup> CD8<sup>+</sup> splenocytes from TCR<sub>TAG</sub> Thy1.1 transgenic mice were adoptively transferred into B6 (Thy1.2) mice; one day later, mice were infected with 5 × 10<sup>6</sup> c.f.u. *Lm*TAG. Effector TCR<sub>TAG</sub> CD8<sup>+</sup> T cells were isolated from the spleens of B6 host mice and analysed 5 or 7 days after *Lm*TAG immunization; memory TCR<sub>TAG</sub> CD8<sup>+</sup> T cells were isolated from spleens of B6 host mice and analysed at least 2–3 months after *Lm*TAG immunization. For the transfer of naive TCR<sub>TAG</sub> T cells into AST-Cre-ER<sup>T2</sup> mice, 1 × 10<sup>5</sup> to 2.5 × 10<sup>6</sup> CD8<sup>+</sup> splenocytes from TCR<sub>TAG</sub> Thy1.1 transgenic mice were adoptively transferred into AST-Cre-ER<sup>T2</sup> mice; 1 day later, mice were treated with 1 mg tamoxifen and donor T cells isolated for subsequent analyses. For memory TCR<sub>TAG</sub> transfer experiments (3–4) × 10<sup>4</sup> TCR<sub>TAG</sub> Thy1.1<sup>+</sup>CD44<sup>hi</sup>CD62L<sup>hi</sup> sorted central memory CD8 T cells were adoptively transferred into AST-Alb-Cre mice; one day later, mice were infected with 5 × 10<sup>6</sup> c.f.u. *Lm*TAG (10<sup>5</sup> central memory T cells were sorted and transferred for experiments without subsequent listeria immunization).

**B16-OVA tumour model.** 5 × 10<sup>5</sup> to 1 × 10<sup>6</sup> B16 tumour cells expressing OVA (full-length or cytosolic as previously described<sup>35</sup>) were injected into C57BL/6J wild-type mice. Once tumours were established (1–2 weeks later) naive Ly5.1 congenically marked TCR<sub>OT1</sub> CD8 T cells were adoptively transferred and isolated from tumours at indicated time points. Tumour volumes did not exceed the permitted volumes specified by the MSKCC IACUC protocol. The B16 cell line was obtained from ATCC. It was tested negative for all rodent pathogens including *Mycoplasma pulmonis*.

**Cell isolation for subsequent analyses.** Spleens were mechanically disrupted with the back of a 3-ml syringe, filtered through a 70-µm strainer, and red blood cells were lysed with ammonium chloride potassium buffer. Cells were washed twice with cold RPMI 1640 media supplemented with 2 µM glutamine, 100 U ml<sup>-1</sup> penicillin/streptomycin, and 5–10% FCS (cRPMI). Liver tissue was mechanically disrupted to a single-cell suspension using a 150 µ metal mesh and glass pestle in ice-cold 3% FCS/HBSS and passed through a 70-µm strainer. The liver homogenate was spun down at 400g for 5 min at 4 °C, and the pellet was resuspended in 30 ml 3% FCS/HBSS, 500 µl (500 U) heparin, and 17 ml Percoll (GE), mixed by inversion, and spun at 500g for 10 min at 4 °C. Pellet was lysed with ammonium chloride potassium buffer and cells were further processed for downstream applications.

**IL-15 *in vitro* culture.** TCR<sub>TAG</sub> or TCR<sub>OT1</sub> cells were isolated from tumours at various time points after transfer and cultured *in vitro* in the presence of IL-15 (100 ng ml<sup>-1</sup>) in cRPMI for 3–4 days.

**Pharmacologic rescue studies.** Naive TCR<sub>TAG</sub> (Thy1.1<sup>+</sup>) cells were transferred into AST-Cre-ER<sup>T2</sup> (Thy1.2<sup>+</sup>) mice which were treated with tamoxifen one day later. On days 2–9, mice were treated with the calcineurin inhibitor FK506 (Prograf, 5 mg ml<sup>-1</sup>) (2.5 mg per kg per mouse i.p. once daily) alone, or in combination with the GSK3β inhibitor TWS119 (Sigma; 0.75 mg per mouse i.p. once daily; days 5–8). Control mice were treated with PBS and/or DMSO.

**Human samples.** Human tumour samples and healthy donor peripheral blood lymphocytes were obtained as per protocols approved by the MSKCC Institutional Review Board (IRB), and all patient and healthy donors provided informed consent. Peripheral blood lymphocytes were flow-sorted for naïve, effector memory-like and central memory-like phenotypes as described in Extended Data Fig. 10a. Human melanoma and lung tumours were mechanically disrupted as described for solid tumours in mice, and CD45RO<sup>+</sup>PD<sup>hi</sup>CD8<sup>+</sup> T cells were flow-sorted for subsequent ATAC-seq analysis.

**Statistical analyses.** Statistical analyses on flow cytometric data were performed using unpaired two-tailed Student's *t* tests (Prism 6.0, GraphPad Software). A *P* value of <0.05 was considered statistically significant.

**Sample preparation for ATAC-seq and RNA-seq.** Mouse samples: replicate samples were isolated from spleens or livers and sorted as follows. (i) Naive TCR<sub>TAG</sub> Thy1.1<sup>+</sup> T cells were sorted by flow cytometry (CD8<sup>+</sup>CD44<sup>lo</sup>) from spleens of TCR<sub>TAG</sub> Thy1.1 transgenic mice. (ii) Day 5 and day 7 effector, and memory TCR<sub>TAG</sub> Thy1.1<sup>+</sup> T cells were sorted by flow cytometry (CD8<sup>+</sup>Thy1.1<sup>+</sup>) from spleens of infected B6 (Thy1.2) host mice (see above) 5 and 7 days or 2–3 months after listeria infection. (iii) TCR<sub>TAG</sub> Thy1.1<sup>+</sup> T cells from pre/early malignant liver lesions: naïve TCR<sub>TAG</sub> Thy1.1<sup>+</sup> T cells were adoptively transferred into AST-Cre-ER<sup>T2</sup> mice. 1 day later, mice were given 1 mg tamoxifen i.p. At given time points after tamoxifen treatment, T cells were isolated and sorted (CD8<sup>+</sup>Thy1.1<sup>+</sup>) from livers as described above. (iv) TCR<sub>TAG</sub> Thy1.1<sup>+</sup> memory T cells from established hepatocellular carcinomas in AST-Alb-Cre mice: TCR<sub>TAG</sub> memory T cells were isolated from tumours and flow sorted (CD8<sup>+</sup>Thy1.1<sup>+</sup>) as described above.

Human samples: samples were flow-sorted as described in Extended Data Fig. 10a.

After flow-sorting, all samples for downstream ATAC-seq analysis were frozen in 10% DMSO/FCS and stored at –80 °C; samples for RNA-seq were directly sorted into Trizol and frozen and stored at –80 °C.

**Transcriptome sequencing.** RNA from sorted cells was extracted using RNeasy mini kit (Qiagen) as per instructions provided by the manufacturer. After ribogreen quantification and quality control of Agilent BioAnalyzer, 6–15 ng of total RNA was amplified (12 cycles) using the SMART-seq V4 (Clontech) ultralow input RNA kit for sequencing. 10 ng of amplified cDNA was used to prepare Illumina hiseq libraries with the Kapa DNA library preparation chemistry (Kapa Biosystems) using 8 cycles of PCR. Samples were barcoded and run on a Hiseq 2500 1T in a 50 bp/50 bp Paired end run, using the TruSeq SBS Kit v3 (Illumina). An average of 51 million paired reads were generated per sample and the percent of mRNA bases was 62.5% on average.

**ATAC sequencing.** Chromatin profiling was performed by ATAC-seq as described previously<sup>11</sup>. In brief, 12,000 to 50,000 cells were washed in cold PBS and lysed. Transposition was performed at 42 °C for 45 min. After purification of the DNA with the MinElute PCR purification kit (Qiagen), material was amplified for 5 cycles. Additional PCR cycles were evaluated by real time PCR. Final product was cleaned by Ampure Beads at a 1.5× ratio. Libraries were sequenced on a Hiseq 2500 1T in a 50 bp/50 bp Paired end run, using the TruSeq SBS Kit v3 (Illumina). An average of 47 × 10<sup>6</sup> paired reads was generated per sample.

**ATAC data and preprocessing.** Raw ATAC-seq reads were trimmed and filtered for quality using Trim Galore! v0.4.0 ([http://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore/](http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/)), powered by CutAdapt v1.8.1 (<http://dx.doi.org/10.14806/ej.17.1.200>) and FastQC v0.11.3 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Paired-end reads were aligned using Bowtie2 v2.2.5 (ref. 36) against either mm10 or hg38 and non-uniquely mapping reads were removed. To correct for the fact that the Tn5 transposase binds as a dimer and inserts two adaptors in the Tn5 tagmentation step<sup>37</sup>, all positive-strand reads were shifted 4 bp downstream and all negative-strand reads were shifted 5 bp upstream to centre the reads on the transposon binding event<sup>11</sup>. We then pooled the shifted reads by sample type and identified peaks using MACS2 (ref. 38) with a threshold of FDR-corrected *P* < 1 × 10<sup>-2</sup> using the Benjamini–Hochberg procedure for multiple hypothesis correction. As called peaks may be caused by noise in the assay and not reflect true chromatin accessibility, we calculated an irreproducible discovery rate (IDR)<sup>39</sup> for all pairs of replicates across a cell type. The IDR is an estimate of the threshold where two ranked lists of results, in this case peak calls ranked by *P* value, no longer represent reproducible events. Using this measure, we excluded peaks

that were not reproducible ( $IDR < 5 \times 10^{-3}$ ) across at least one pair of replicates in each mouse or human cell type.

**ATAC-seq atlas creation.** Peaks found reproducibly in each mouse cell type were combined to create a genome-wide atlas of accessible chromatin regions. Reproducible peaks from different samples were merged if they overlapped by more than 75%. To create the atlas of accessible peaks for the human samples, reproducible peaks from the normal human cell types (HN, HCM, and HEM) and the tumour-derived cells (PD1<sup>hi</sup>) were combined. There was greater variation between the human TIL samples than between T cell samples from healthy donors; this led to fewer reproducible peaks being called in the TIL samples. Like the mouse atlas, peaks overlapping by more than 75% were merged in the human atlas. Numbers of called peaks and reproducible peaks for each sample type are listed in Supplementary Data.

**Assignment of ATAC-seq peaks to genes.** The RefSeq transcript annotations of the hg38 version of the human genome and the mm10 version of the mouse genome were used to define the genomic location of transcription units. For genes with multiple gene models, the longest transcription unit was used for the gene locus definition. ATAC peaks located in the body of the transcription unit, together with the 2-kb regions upstream of the TSS and downstream of the 3' end, were assigned to the gene. If a peak was found in the overlap of the transcription units of two genes, one of the genes was chosen arbitrarily. Intergenic peaks were assigned to the gene with a TSS or 3' end that was closest to the peak. In this way, each peak was unambiguously assigned to one gene. Peaks were annotated as promoter peaks if they were within 2 kb of a transcription start site. Non-promoter peaks were annotated as intergenic, intronic or exonic according to the relevant RefSeq transcript annotation.

**ATAC-seq peak atlas summary.** We found a total of 75,689 reproducible ATAC-seq peaks in the mouse samples. Examining genomic locations, 39.6% of the peaks were found in introns, 36.3% were found in intergenic regions, 22.1% were found in promoters and 2.1% were found in exons. In the human samples, we found a total of 42,104 reproducible ATAC-seq peaks. Among these peaks, 34.0% were found in introns, 29.9% were found in intergenic regions, 34.0% were found in promoters, and 2.0% were found in exons. Chromosome-wide genomic coverage for all (autosomal) chromosomes and all samples was examined and no systemic bias was observed.

**Principal component analysis.** PCA plots were generated using read counts against all mouse or human atlas peaks. These read counts were processed using the variance-stabilizing transformation built into the DESeq2 package<sup>40</sup>.

**Differential peak accessibility.** Reads aligning to atlas peak regions were counted using the summarizeOverlaps function of the R packages GenomicAlignments v1.2.2 and GenomicRanges v1.18.4 (ref. 41). Differential accessibility of these peaks was then calculated for all pairwise comparisons of cell types using DESeq2 v1.6.3 (ref. 40).

**Peak heat maps and genome coverage plots.** The ATAC-seq peak heat maps were created by pooling the DESeq size-factor normalized read counts per atlas peak across replicates of ATAC-seq data and binning the region  $\pm 1$  kb around the peak summit in 20 bp bins. To improve visibility, bins with read counts greater than the 75th percentile + 1.5  $\times$  IQR were capped at that value. All analysis was performed using the original uncapped read counts. Genome coverage plots were generated for each replicate of ATAC-seq and RNA-seq by calculating genome-wide coverage of aligned reads using the bedtools function genomecov<sup>42</sup>. For ATAC-seq samples, this coverage was calculated after shifting the reads to account for the Tn5-induced bias. The coverage values were then normalized using DESeq2-derived size factors and replicates were combined to create one signal track for each sample type. ATAC-seq and RNA-seq coverage plots were generated using the Integrated Genomics Viewer (Broad)<sup>43</sup>.

**Transcription factor peak assignment.** Using the MEME<sup>44</sup>-curated CisBP<sup>45</sup> transcription factor binding motif (TFBM) reference, we scanned the mouse ATAC-seq peak atlas with FIMO<sup>46</sup> to find peaks likely to contain each TFBM ( $P < 10^{-4}$ ). The MEME cisBP reference for direct and inferred motifs for *Mus musculus* was curated by the MEME suite developers as follows: to reduce redundancy, for each transcription factor a single motif was selected according to the following precedence rules. The direct motif was chosen if there was one, otherwise the inferred motif with the highest DNA binding domain (DBD) similarity (according to CisBP) to a transcription factor in another species with a direct motif was chosen. If there was more than one direct motif or inferred motif with the highest DBD similarity, a motif was chosen according to its provenance (CisBP 'Motif\_Type' attribute) in the following order: ChIP-seq, HocoMoco, DeBoer11, PBM, SELEX, B1H, High-throughput Selex CAGE, PBM:CSA:DIP-chip, ChIP-chip, COMPILED, DNaseI footprinting. Each motif thus determined was linked to a single transcription factor in the CisBP database, following the same precedence rules. The final reference contained 718 motifs between 6 and 30 bp in width (average width, 10.7 bp). Transcription factors with similar FIMO-predicted target

peaks were combined into transcription factor families. Similarity of predicted target peak sets was measured using the Jaccard index (size of intersection/size of union). Transcription factors with Jaccard indices greater than 0.7 were combined for further analyses. Relative transcription factor accessibility was calculated using two one-sided Wilcoxon rank-sign tests comparing the distributions of peak heights for peaks containing FIMO-predicted transcription factor binding sites. Peak height was defined as the maximum observed number of reads overlapping at any point in the defined peak region.

**Footprinting analysis.** ATAC-seq footprints containing FIMO-predicted transcription factor binding sites ( $P < 1 \times 10^{-4}$ ) were selected. Positive- and negative-strand ATAC-seq cut sites were counted 100 bp up- and down-stream of the centre of the motif site in each of the selected peaks. The mean number of ATAC-seq cut sites across matching atlas peaks was then plotted to generate the footprint figures.

**Diamond plots.** In these plots, each gene is represented by a stack of diamonds corresponding accessible chromatin regions of the same gene. The bottom-most peak in this stack corresponds to the  $\log_2$  fold change in expression of the gene. The diamonds are coloured according to the accessibility change of their ATAC-seq peak with blue indicating closing and red indicating opening. The colour scale was based on the rank-order of the peak accessibility changes. In Extended Data Fig. 6d, the colour scale ranges from a  $\log_2$  fold change of -3.92 to 4.96 (L14/L7).

**Comparison of human and mouse ATAC-seq atlases.** The UCSC liftOver tool<sup>47</sup> was used to convert the mouse ATAC-seq peak atlas from mm10 coordinates to hg38 coordinates. The converted mouse atlas was then compared to the human atlas and 20,642 mouse peaks were within 100 bp of a human peak. We compared the results from the UCSC liftover tool and an alternative method, bnMapper<sup>48</sup>, and confirmed that the set of peaks mapped by bnMapper and by the UCSC liftOver tool was nearly identical (57,383 out of 75,689 by liftOver and 58,299 out of 75,689 by bnMapper). Additionally, all 57,223 peaks mapped to hg38 by both tools were mapped to the same chromosomal positions. The majority of these conserved peaks were found in promoter regions (56.4%), whereas relatively fewer were found in intergenic (22.4%), intronic (19.6%), and exonic (1.5%) regions. For non-promoter peaks conserved between human and mouse, Spearman correlations of  $\log_2(FC)$  were calculated between human N and human EM, CM or PD1<sup>hi</sup> TIL versus  $\log_2(FC)$  between mouse N and functional E5, E7, M and dysfunctional L5 to L60.

**RNA-seq.** Raw ATAC-seq reads were trimmed and filtered for quality using Trim Galore! v0.4.0 ([http://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore/](http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/)), powered by CutAdapt v1.8.1 (<http://dx.doi.org/10.14806/ej.17.1.200>) and FastQC v0.11.3 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Paired-end reads were aligned using STAR<sup>49</sup> against either mm10 or hg38. The RefSeq transcript annotations of the hg38 version of the human genome and the mm10 version of the mouse genome were used for the genomic location of transcription units. Reads aligning to annotated exon regions were counted using the summarizeOverlaps function of the R packages GenomicAlignments v1.2.2 and GenomicRanges v1.18.4 (ref. 41). Differential expression of genes across cell types was calculated using DESeq2 v1.6.3 (ref. 40). FDR correction of 0.05 was imposed unless otherwise stated. A  $\log_2$  fold change cutoff of 1 was used in some analyses as indicated.

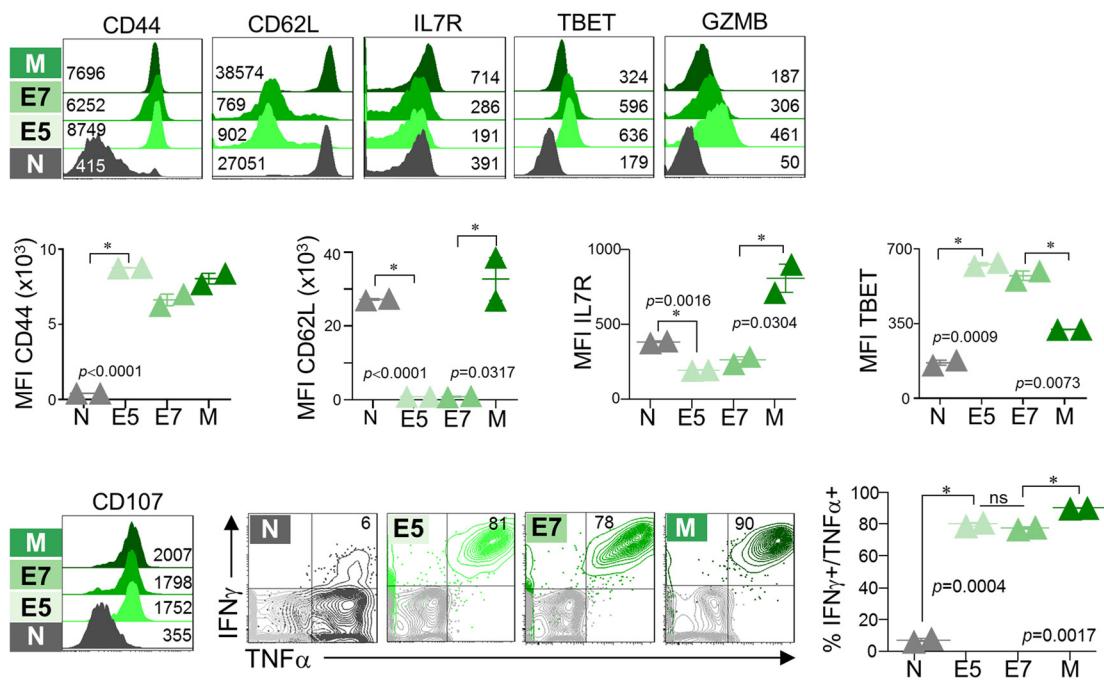
**Pathway analysis.** Enrichment of gene ontology terms in sets of ATAC-seq peaks was calculated using GREAT (Genomic Regions Enrichment of Annotations Tool) using default parameters<sup>50</sup>. The full ATAC-seq atlas was used as the background set.

**Membrane protein analysis.** To identify membrane proteins that distinguished early (L5–L7) from late (L14–L60) dysfunctional TST, RNA-seq data was analysed for genes contained within the gene ontology category 0016020 (membrane proteins). The top 50 most up- and downregulated genes (size-factor normalized RPKM) when compared between L5–L7 and L14–L60 were plotted in a heat map (row-normalized). Protein expression was assessed by flow cytometry for those membrane proteins for which monoclonal antibodies were available. Mouse targets (clone; supplier): CD5 (53-7.3; eBioscience), CD30L (RM153; eBioscience), CD38 (90; Biolegend), and CD101 (Moushi101; eBioscience). Human targets: CD5 (L17F12; Biolegend), CD38 (HB7; eBioscience), CD101 (BB27; Biolegend).

**Data reporting.** No statistical methods were used to predetermine sample size. The investigators were not blinded to allocation during experiments and outcome assessment. Mice or human samples were excluded if donor or tumour-infiltrating CD8 T cells could not be found.

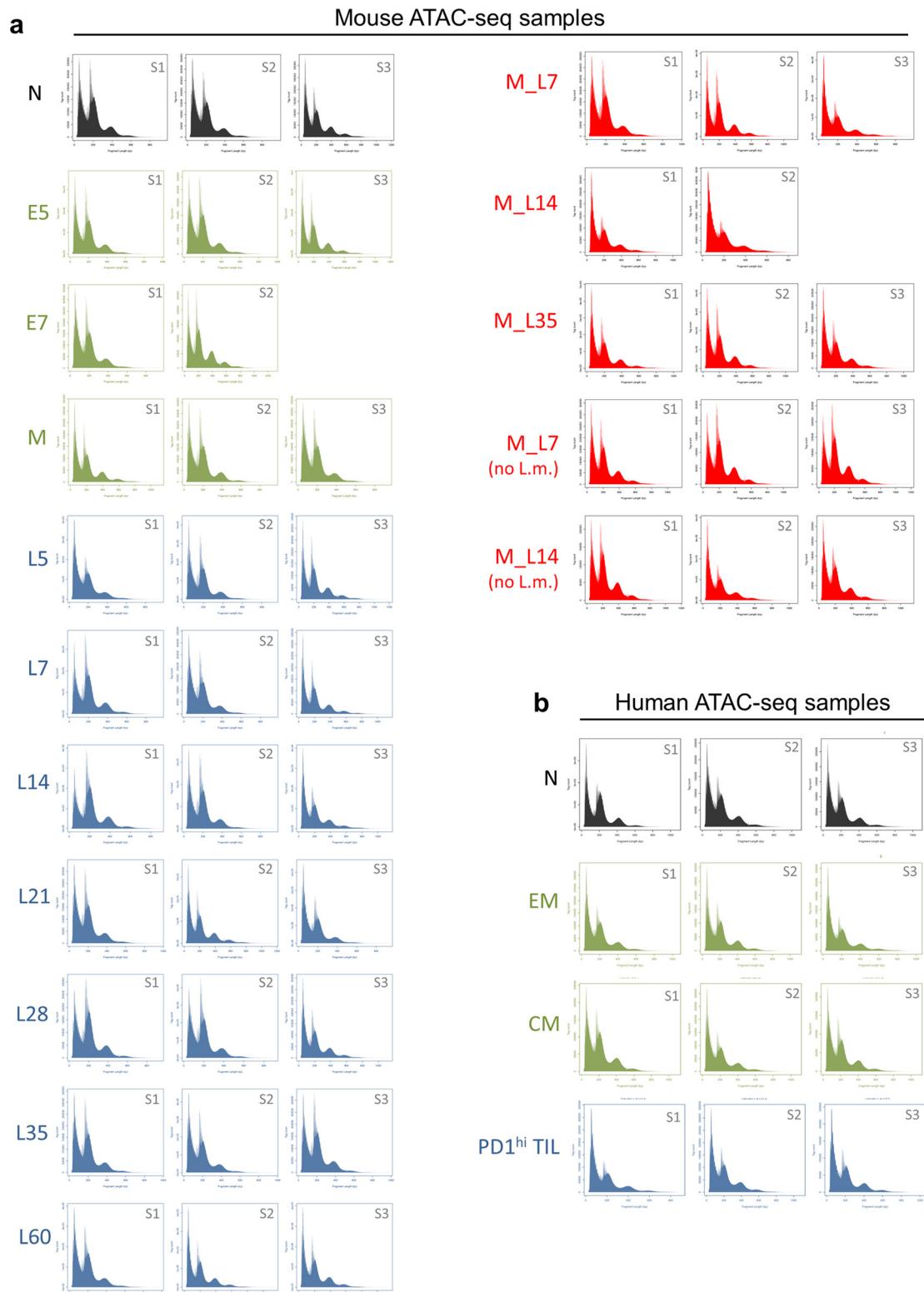
**Data availability.** All data generated and supporting the findings of this study are available within the paper. The RNA-seq and ATAC-seq data have been deposited in the Gene Expression Omnibus (GEO Super-Series accession number GSE89309 (GSE89307 for RNA-seq, GSE89308 for ATAC-seq)). Source Data for Figs 1–5 and Extended Data Figs 1, 3, and 7 are provided with the online version of the paper. Additional information and materials will be made available upon request.

33. Stahl, S. *et al.* Tumor agonist peptides break tolerance and elicit effective CTL responses in an inducible mouse model of hepatocellular carcinoma. *Immunol. Lett.* **123**, 31–37 (2009).
34. Sinnathamby, G. *et al.* Priming and activation of human ovarian and breast cancer-specific CD8<sup>+</sup> T cells by polyvalent *Listeria monocytogenes*-based vaccines. *J. Immunother.* **32**, 856–869 (2009).
35. Engels, B. *et al.* Relapse or eradication of cancer is predicted by peptide-major histocompatibility complex affinity. *Cancer Cell* **23**, 516–526 (2013).
36. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie2. *Nat. Methods* **9**, 357–359 (2012).
37. Adey, A. *et al.* Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome Biol.* **11**, R119 (2010).
38. Zhang, Y. *et al.* Model-based analysis of ChIP-seq (MACS). *Genome Biol.* **9**, R137 (2008).
39. Zigler, C. M. & Belin, T. R. The potential for bias in principal causal effect estimation when treatment received depends on a key covariate. *Ann. Appl. Stat.* **5**, 1876–1892 (2011).
40. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
41. Lawrence, M. *et al.* Software for computing and annotating genomic ranges. *PLOS Comput. Biol.* **9**, e1003118 (2013).
42. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
43. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
44. Bailey, T. L. *et al.* MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202–W208 (2009).
45. Weirauch, M. T. *et al.* Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* **158**, 1431–1443 (2014).
46. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017–1018 (2011).
47. Hinrichs, A. S. *et al.* The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res.* **34**, D590–D598 (2006).
48. Denas, O. *et al.* Genome-wide comparative analysis reveals human-mouse regulatory landscape and evolution. *BMC Genomics* **16**, 87 (2015).
49. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
50. McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).

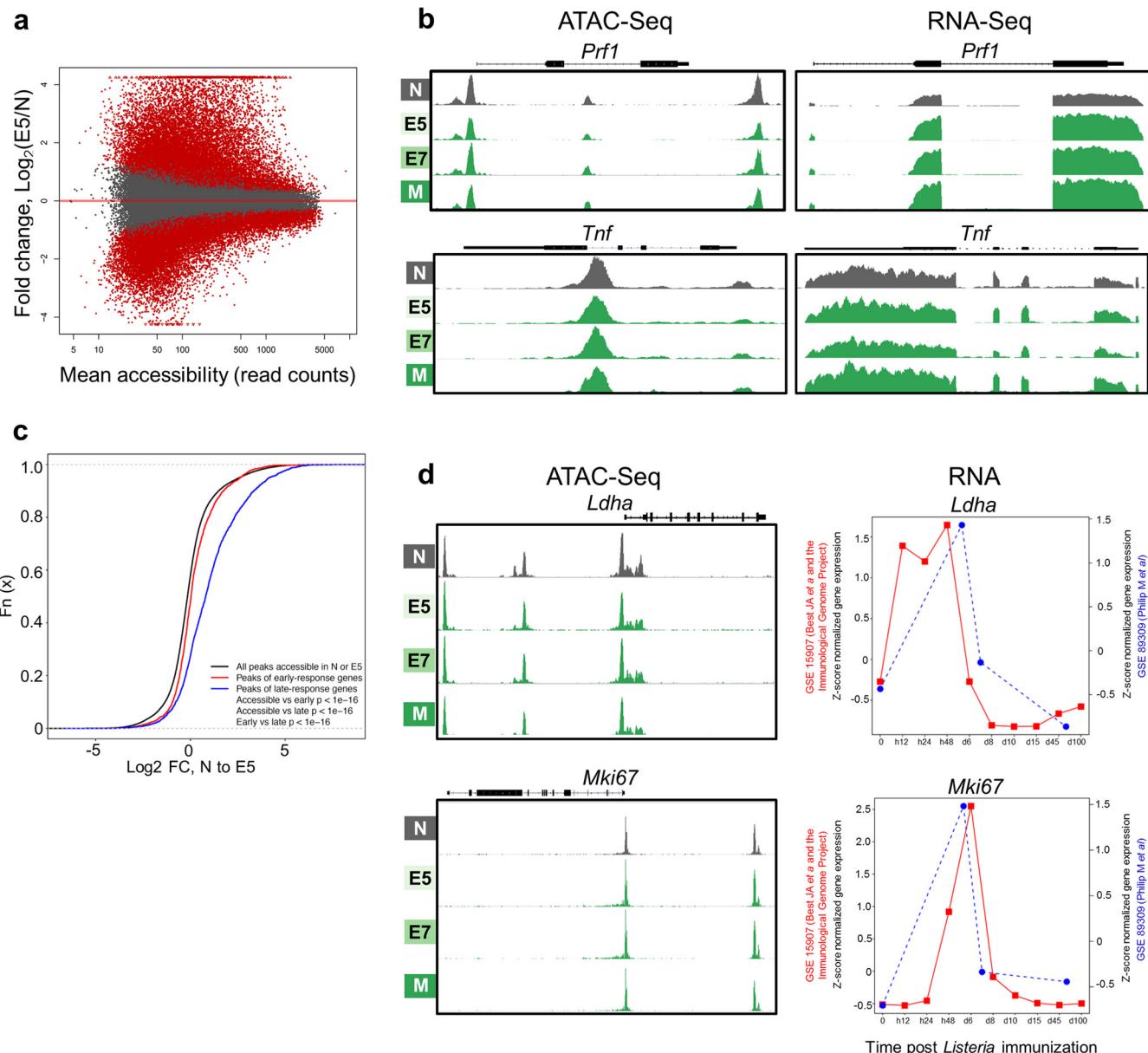


**Extended Data Figure 1 | Phenotypic and functional characteristics of naive TCR<sub>TAG</sub> CD8 T cells differentiating to effector and memory T cells during acute *Listeria* infection.** Naive TCR<sub>TAG</sub> cells (N; Thy1.1<sup>+</sup>) were transferred into B6 (Thy1.2<sup>+</sup>) mice, which were immunized with *LmTAG* one day later. At days 5, 7, and 60+ after *LmTAG*, effector (E5 and E7), and memory (M) T cells were isolated from spleens and assessed for phenotype and function. Flow cytometric analysis of CD44, CD62L, IL7R $\alpha$ , TBET, and GZMB expression directly *ex vivo* (upper panel; inset

numbers show MFI), and intracellular IFN $\gamma$  and TNF $\alpha$  production and CD107 expression after 4 h of *ex vivo* TAG peptide stimulation (lower panel). Flow plots are gated on CD8<sup>+</sup>Thy1.1<sup>+</sup> cells. For cytokine production, in grey are shown no-peptide control cells. ( $n = 8$  total, with  $n = 2$  per cell state). Each symbol represents an individual mouse. Data show mean  $\pm$  s.e.m.;  $P$  values calculated using unpaired, two-tailed Student's *t*-test. Data are representative of more than four independent experiments.

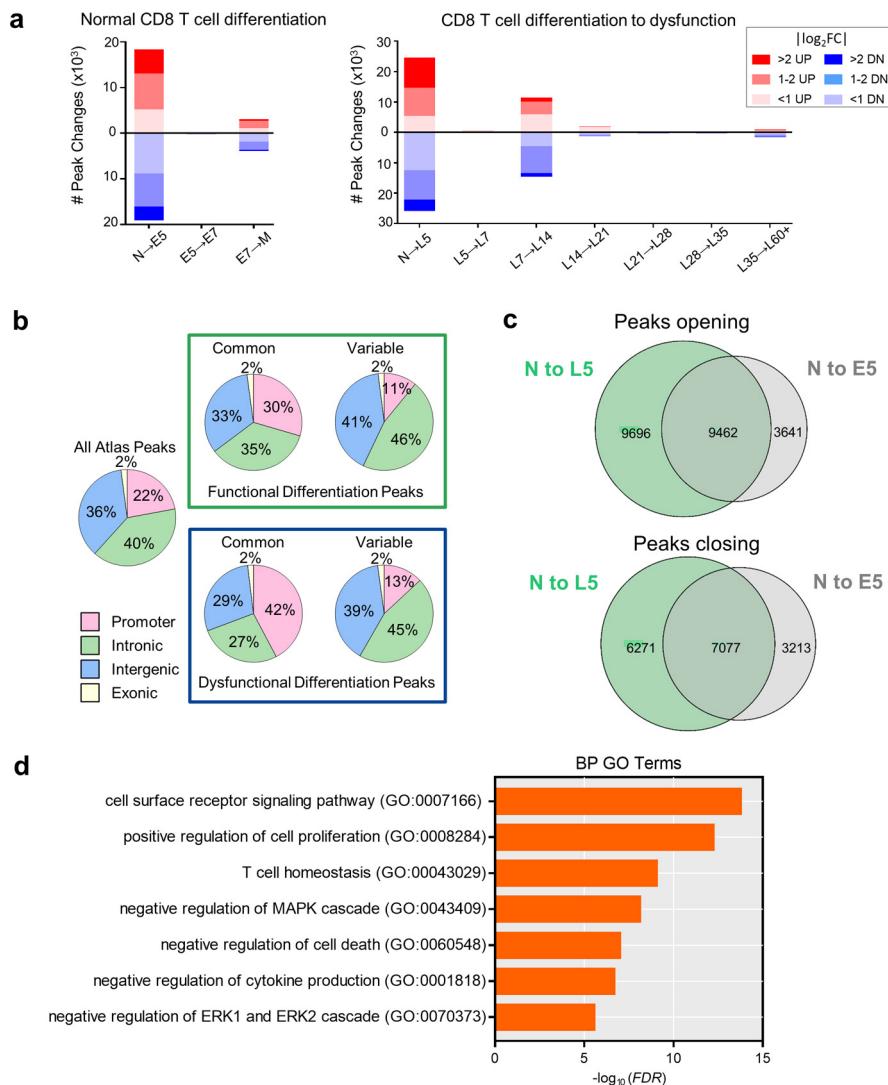


**Extended Data Figure 2 | Fragment length distribution plots of ATAC-seq samples.** **a, b,** Plots are shown for all mouse (**a**) and human (**b**) CD8 T cell ATAC-seq samples displaying fragment length (bp; x axis) and read counts (y axis). (S1, S2, S3 represent replicates per sample group.)



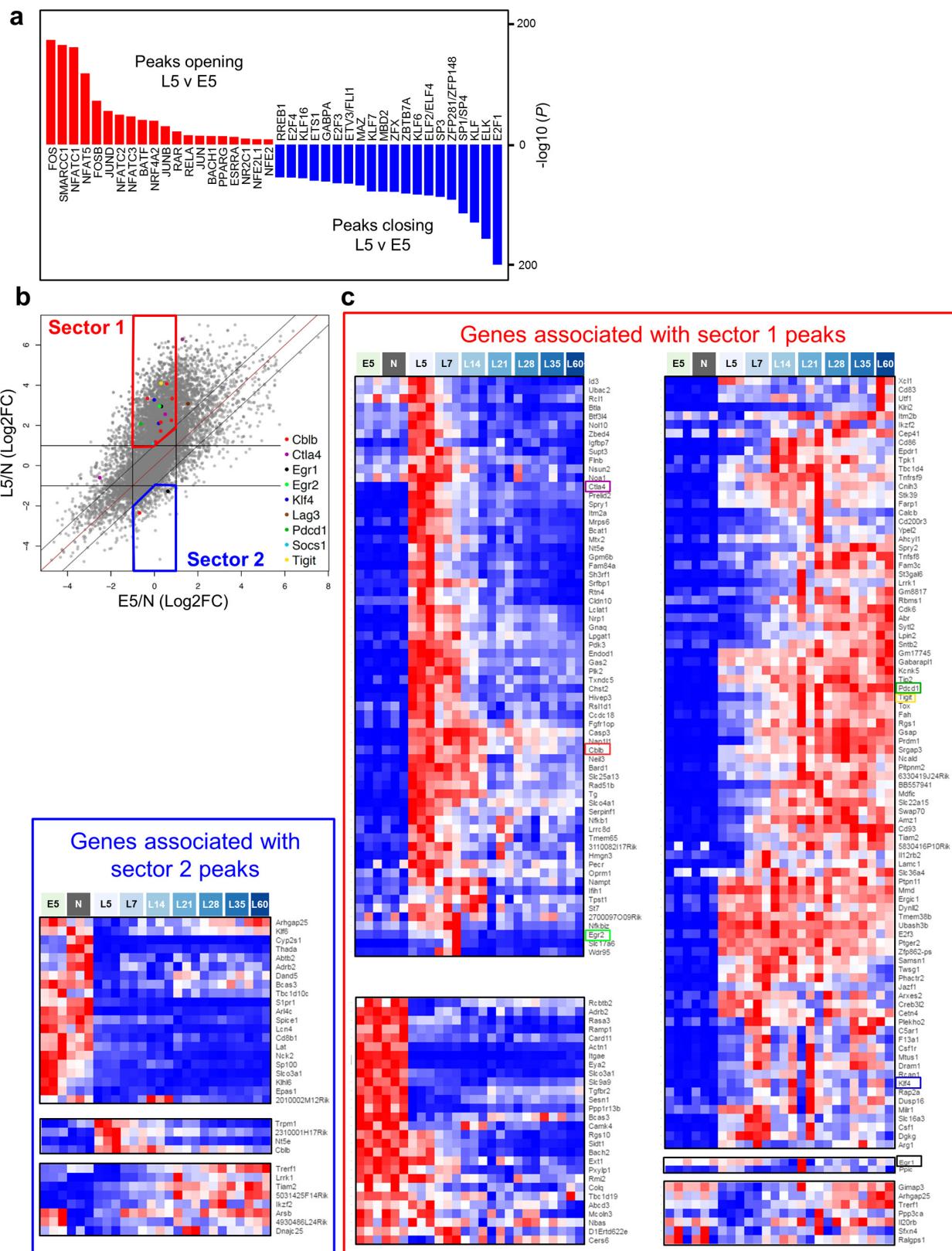
**Extended Data Figure 3 | Epigenetic and transcriptional regulation of normal CD8 differentiation.** **a**, ATAC-seq data reveals massive chromatin remodelling during normal CD8 T cell differentiation. MA plot of naive (N) and day 5 effectors (E5) showing  $\log_2$  ratios of peak accessibility (E5/N) versus mean read counts for all atlas peaks. Significantly differentially accessible peaks are shown in red (FDR < 0.05). **b**, Epigenetic and transcriptional regulation of CD8 effector genes. ATAC-seq (left) and RNA-seq (right) signal profiles of *Prf1* and *Tnf* in naive, effectors (E5 and E7), and memory (M) TCR<sub>TAG</sub> cells during acute *LmTAG* infection. **c, d**, Epigenetic and transcriptional regulation of early CD8 response genes in TCR<sub>TAG</sub> cells during acute listeria infection. Published expression data from the Immunological Genome Project (ref. 17; GSE15907) were used; early-response genes increase in expression within the first 12–24 h

and late-response genes increase expression 24–48 h after naive T cells encounter *LmOVA* as determined in ref. 17. **c**, Cumulative distribution function of peak accessibility changes between N and E5. Peaks associated with early-response genes show fewer changes in accessibility as compared to peaks associated with late-response genes. The black line shows all peaks accessible in N or E5, the red line shows peaks associated with early-response genes and the blue line shows peaks associated with late-response genes. **d**, ATAC-seq signal profiles (left) and RNA expression (right) of the early response genes *Ldha* (top) and *Mki67* (bottom) in N, E5/E7, and M TCR<sub>TAG</sub> cells during acute *LmTAG* infection (blue line; GSE89309, current data set) overlaid with expression data from ref. 17/Immunological Genome Project (red line).



**Extended Data Figure 4 | Chromatin peak accessibility changes during normal and dysfunctional CD8 T cell differentiation.** **a**, Number of DESeq-determined chromatin peak accessibility changes during each transition during normal CD8 T cell differentiation (*Listeria* infection) (right) and CD8 T cell differentiation to dysfunction during tumorigenesis (left) broken down by  $\log_2(\text{FC}) > 2$ ,  $\log_2(\text{FC}) = 1-2$ , and  $\log_2(\text{FC}) < 1$ . **b**, Chromatin accessibility peaks gained or lost during normal and dysfunctional CD8 T cell differentiation were mainly found in intergenic and intronic regions. Pie charts showing the proportions of reproducible ATAC-seq peaks in exonic, intronic, intergenic, and promoter regions (left, distribution for all peaks in the atlas). Green box: normal CD8 T cell differentiation during *LmTAG* immunization; distribution for common and variably accessible peaks in N, E5, E7, and M functional CD8

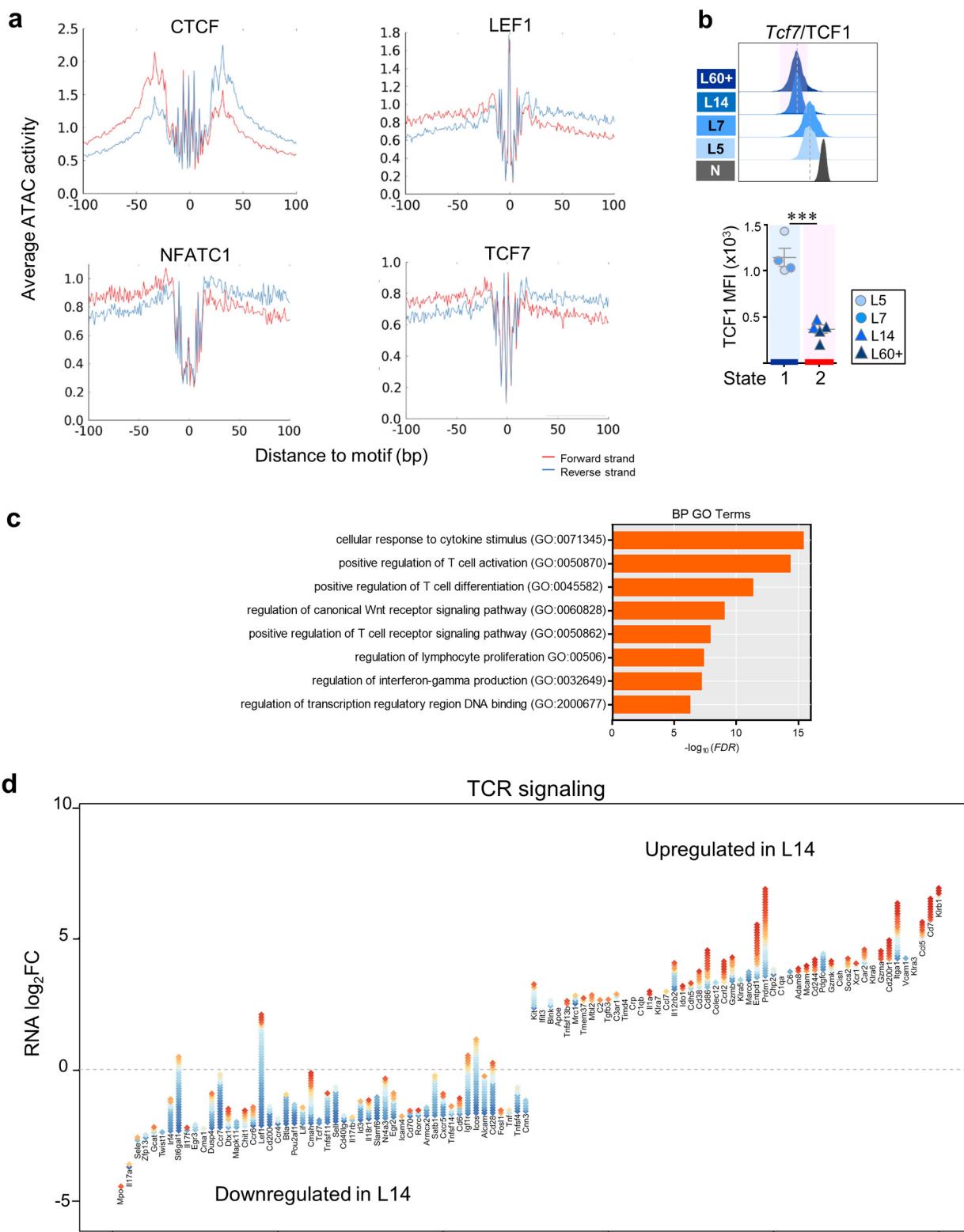
T cells. Blue box: differentiation to dysfunction in progressing tumours; distribution for common and variably accessible peaks in N, L5, L7, L14, L21, L28, L35, and L60+. Variable: significant change in at least one cell type comparison (FDR  $< 0.05$ ,  $\log_2(\text{FC}) > 1$ ). Common: no change in any cell type comparison. **c**, Venn diagrams show the number of significantly changed peaks during the transition from naive (N) to day 5 effectors (E5) TCR<sub>TAG</sub> cells during acute listeria *LmTAG* infection versus N to L5 early malignant lesion-infiltrating TCR<sub>TAG</sub> cells (FDR  $< 0.05$ ,  $\log_2(\text{FC}) > 2$ ). Upper, Venn diagram shows opening peaks; lower, Venn diagram shows closing peaks. **d**, Selected biological process (BP) Gene Ontology (GO) terms enriched in peaks open in L5 relative to E5 as determined through GREAT analysis.



Extended Data Figure 5 | See next page for caption.

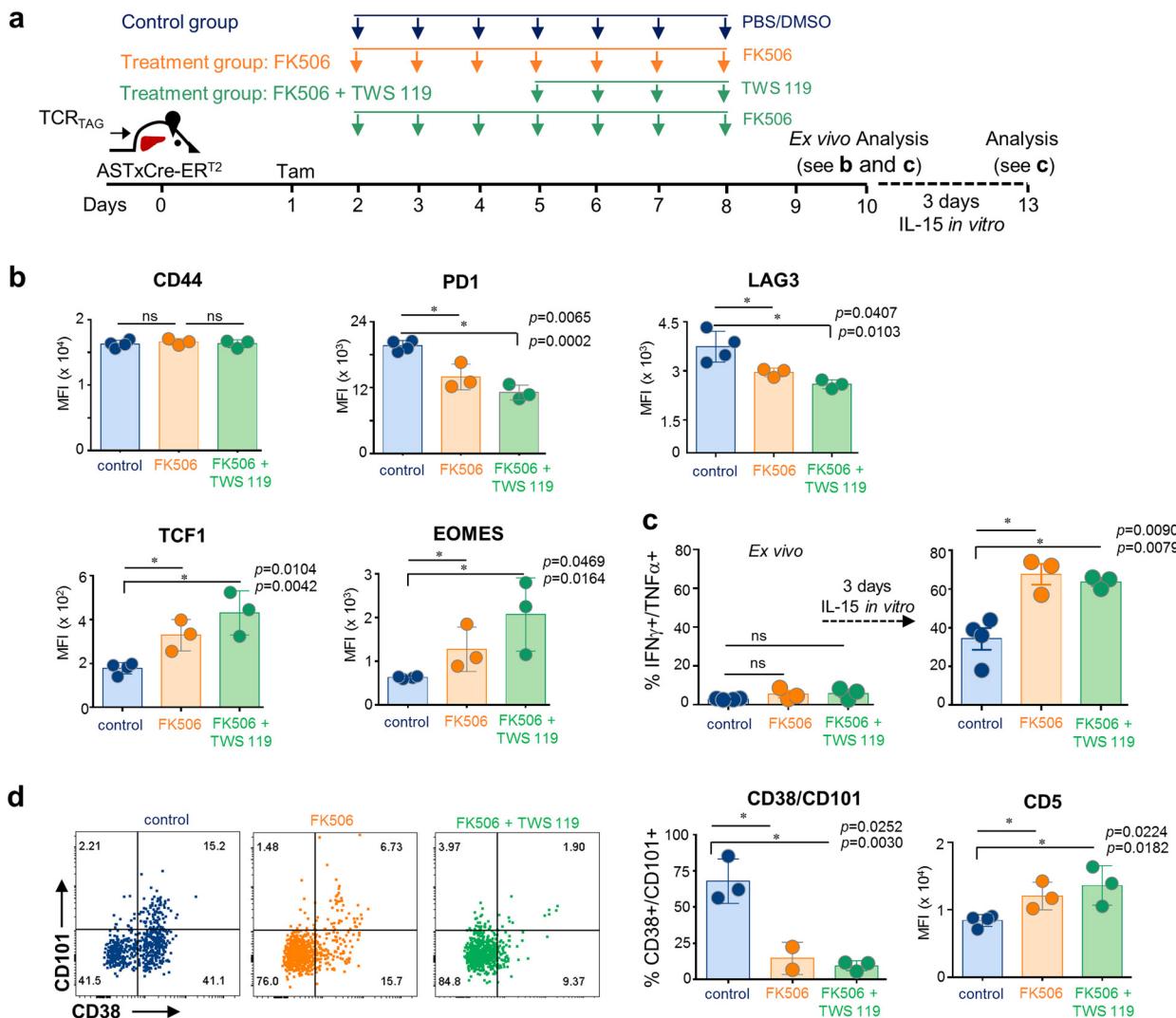
**Extended Data Figure 5 | NFATC1 targets become significantly more accessible during differentiation to dysfunction in early malignant lesions as compared to normal effector differentiation.** **a**, The 20 most significantly enriched transcription factor motifs in peaks opening (red) and closing (blue) between L5 and E5. **b**, Scatterplot comparing the changes in peak accessibility for all differentially accessible peaks containing the NFATC1 motif during the transition from naive (N) to day 5 effectors (E5) TCR<sup>+</sup> cells during acute listeria *Lm*TAG infection versus N to L5 in pre-malignant lesions (FDR < 0.05, log<sub>2</sub>(FC) > 1). Highlighted are NFATC1 target peaks associated with genes encoding negative regulatory transcription factors and inhibitory receptors. Some genes, for example, *Cblb* and *Klf4*, had multiple NFATC1 target peaks, including peaks that decreased in accessibility. **c**, Genes with more accessible NFATC1 target peaks during differentiation to dysfunction in malignant lesions show increased expression levels. Gene expression for

genes with peaks in sector 1 and sector 2, with increased and decreased accessibility in L5 versus E5, respectively. Heat maps show RNA-seq expression data (row-normalized) for differentially expressed ( $P < 0.01$ ,  $\log_2(\text{FC}) > 1$ ) genes with NFATC1 target peaks contained in sector 1 (red box) or sector 2 (blue box) of scatterplot presented **b**. The majority of sector 1 genes (195 out of 223, 87%) revealed increased expression in dysfunctional TST as compared to E5, whereas the majority of sector 2 genes (21 out of 33, 63%) had decreased expression. Genes are clustered by row according to expression across the samples. Interestingly, although many genes in sector 1 had transiently increased expression in L5 and L7 (red box, upper left), many genes increased in expression at later stages of tumorigenesis at L14 and beyond (red box, upper right). This suggests that NFATC1 activation of downstream targets (negative regulators of T cell function) may not only induce early dysfunction, but may cause or contribute to the transition from plastic to fixed dysfunction.



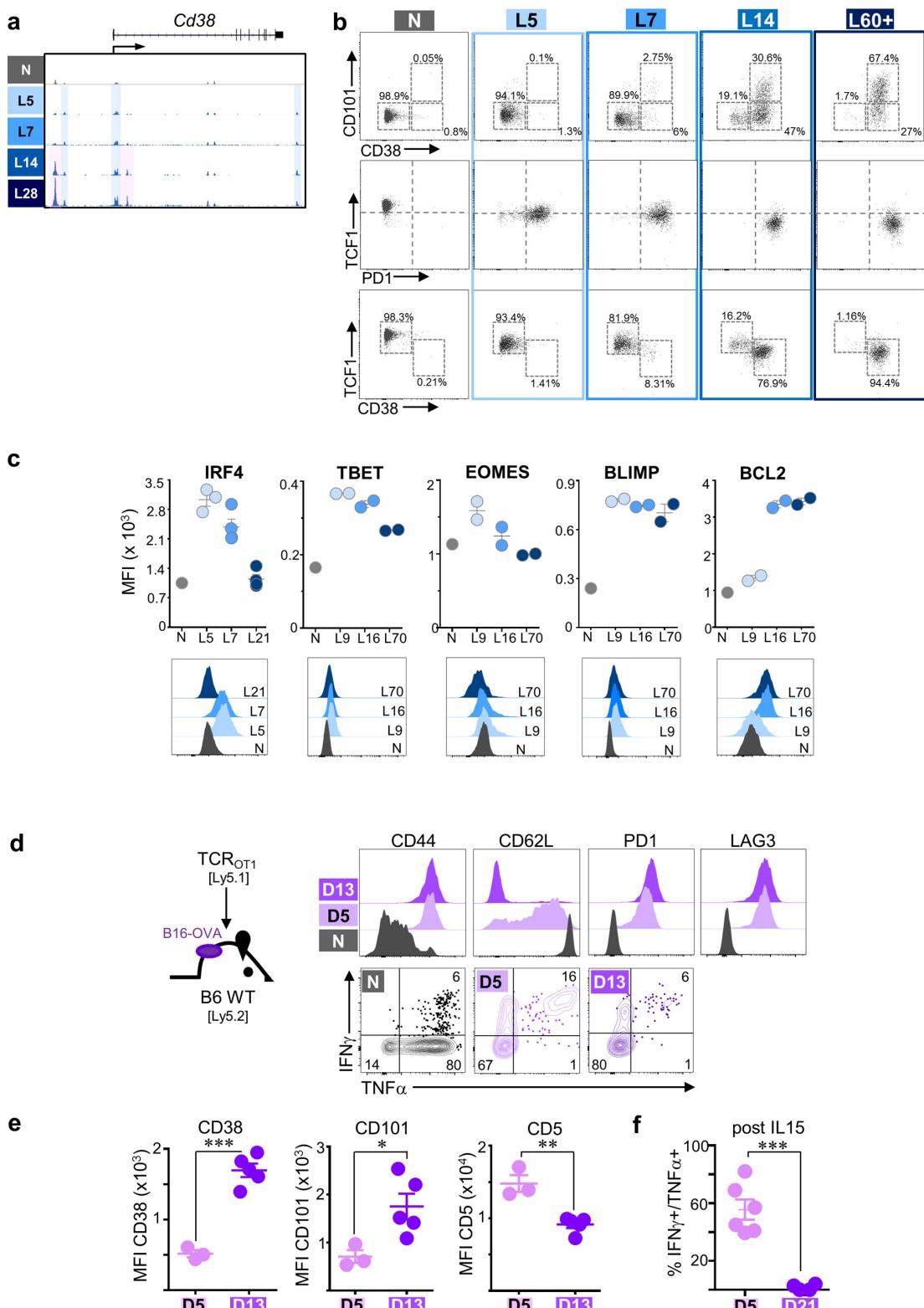
**Extended Data Figure 6 | Epigenetic and transcriptional changes during the L7 to L14 transition.** **a**, Transcription factor footprinting in chromatin accessible regions. ATAC cut site distributions show footprints for CTCF, LEF1, NFATC1, and TCF7 in naïve CD8 T cells. Shown is the mean number of ATAC cut sites on the forward (red) or reverse (blue) strand 100 bp up and downstream of the transcription factor motif site, calculated for atlas peaks predicted by FIMO to be bound by the respective transcription factor ( $P < 10^{-4}$ ). **b**, TCF1 expression (MFI; mean fluorescence intensity). Each symbol represents individual mouse. Mean  $\pm$  s.e.m. shown; \*\*\* $P \leq 0.0001$  (Student's *t*-test). **c**, Selected

biological processes (BP) (gene ontology (GO) terms) enriched in genes which significantly lost chromatin accessibility during the L7 to L14 transition as determined through GREAT analysis. **d**, Gain and losses of regulatory elements for top 50 most differentially expressed genes associated with TCR signalling during the L7 to L14 transition. Top 25 genes associated with TCR signalling with highest and lowest logFC gene expression changes are shown. Each gene is illustrated by a stack of diamonds, where each diamond represents a chromatin peak associated with the gene. Red diamonds denote peaks gained in the transition, blue diamonds denote peaks that were lost.



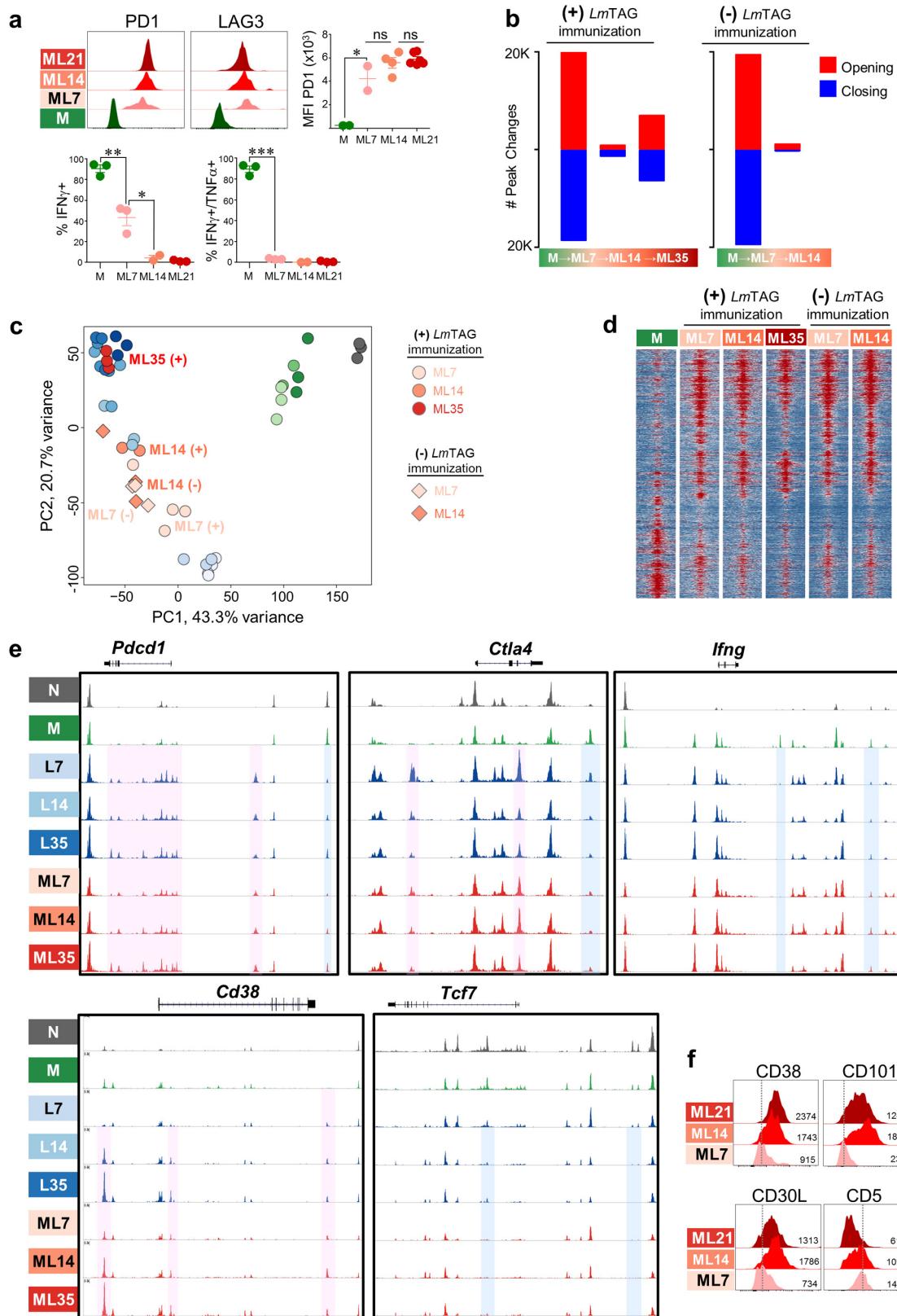
**Extended Data Figure 7 | Pharmacological targeting of NFAT and Wnt/β-catenin signalling prevents TST differentiation to the fixed dysfunctional state *in vivo*.** **a**, Experimental scheme. Naive TCR<sub>TAG</sub> cells (Thy1.1<sup>+</sup>) were transferred into AST-Cre-ER<sup>T2</sup> (Thy1.2<sup>+</sup>) mice which were treated with tamoxifen (tam) one day later. At days 2–9 mice were treated with the calcineurin inhibitor FK506 (2.5 mg kg<sup>-1</sup> per mouse) alone (FK506 treatment group; orange), or in combination with the GSK3β inhibitor TWS119 (0.75 mg per mouse; days 5–8) (FK506 + TWS119 treatment group; green), or PBS/DMSO (control group; blue) as indicated. At day 10, TCR<sub>TAG</sub> cells were isolated from livers and assessed for phenotype and function. **b**, Flow cytometric analysis of CD44, PD1, LAG3, TCF1, and EOMES expression of TCR<sub>TAG</sub> cells.

TCF1, and EOMES expression of TCR<sub>TAG</sub> cells. **c**, Production of IFN $\gamma$  and TNF $\alpha$  by TCR<sub>TAG</sub> cells isolated at day 10 (left panel; *ex vivo*), and after 3 days IL-15 *in vitro* culture (right panel). Each symbol represents an individual mouse. Data show mean  $\pm$  s.e.m.; *P* values calculated using unpaired two-tailed *t*-test. **d**, Representative flow cytometric analysis of CD38 and CD101 expression of TCR<sub>TAG</sub> cells (numbers indicate %); CD38, CD101 and CD5 expression. Each symbol represents an individual mouse. Data show mean  $\pm$  s.e.m.; *P* values calculated using unpaired two-tailed *t*-test. These data are representative of 2 independent experiments (with total *n* = 10 for experiment 1; *n* = 9, experiment 2).



**Extended Data Figure 8 | Epigenetic and expression dynamics of membrane proteins and transcription factors associated with T cell dysfunction.** **a**, ATAC-seq signal profile across the *Cd38* loci with 'state 2' uniquely accessible peaks highlighted in pink; activation-associated peaks highlighted in blue. **b**, Expression profiles of N, L5, L7, L14, and L60+ TCR<sub>OT1</sub> cells for CD101 versus CD38, TCF versus PD1, and TCF1 versus CD38 by flow cytometric analysis. **c**, Expression of transcription factors and other proteins on tumour-specific TCR<sub>OT1</sub> T cells over the course of tumorigenesis (MFI; mean fluorescence intensity). Each symbol represents an individual mouse. Data shows mean  $\pm$  s.e.m. (bottom panel).

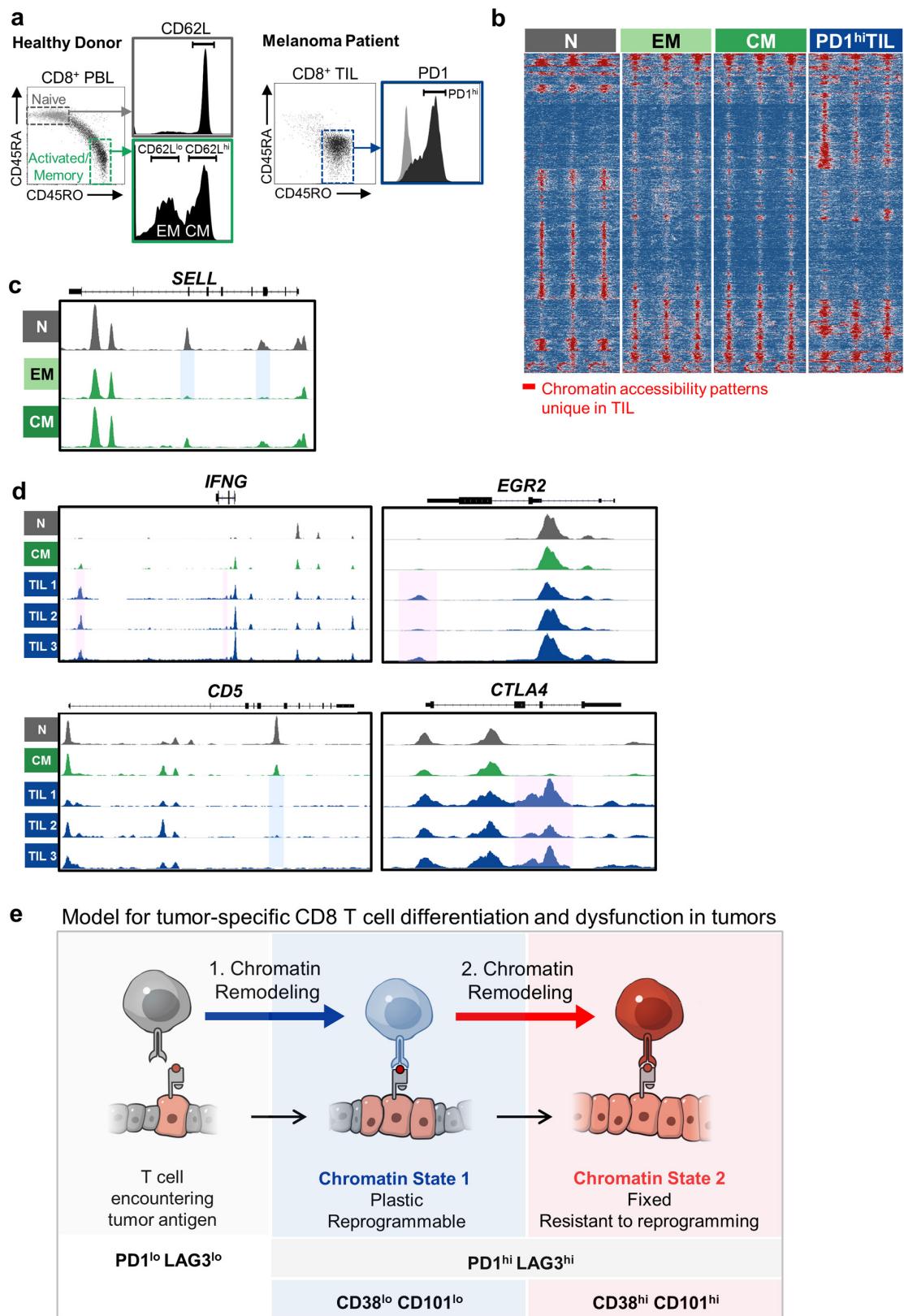
Representative flow histogram overlays are shown. **d–f**, TCR<sub>OT1</sub> TST in established B16-OVA tumours enter plastic and fixed dysfunctional states. **d**, Immunophenotype of and cytokine production by TCR<sub>OT1</sub> cells re-isolated from established B16-OVA tumours 5 (D5) and 13 (D13) days after transfer. **e**, CD38, CD101 and CD5 expression on day 5 and day 13 TCR<sub>OT1</sub> cells. **f**, Cytokine production by day 5 and day 21 TCR<sub>OT1</sub> cells after 3 days of IL-15 *in vitro* culture. Each symbol represents individual mouse. Mean  $\pm$  s.e.m. shown; \* $P = 0.03$ , \*\* $P = 0.002$ , \*\*\* $P \leq 0.0003$  (Student's *t*-test).



Extended Data Figure 9 | See next page for caption.

**Extended Data Figure 9 | Chromatin state dynamics of memory TCR<sub>TAG</sub> cells differentiating to the dysfunctional state in solid tumours.** **a**, PD1 and LAG3 expression and cytokine production of memory TCR<sub>TAG</sub> cells in liver tumours. Each symbol represents individual mouse. Mean  $\pm$  s.e.m. shown; \* $P=0.03$ , \*\* $P=0.006$ , \*\*\* $P<0.0001$  (Student's *t*-test); representative of four independent experiments. **b**, Numbers of ATAC-seq peaks significantly opening or closing ( $FDR < 0.05$ ) during each transition as memory TCR<sub>TAG</sub> cells differentiate to the dysfunctional state 7, 14, and 35 days after transfer into hepatocellular-carcinoma-tumour bearing AST-Alb-Cre mice with ((+); left) and without ((-); right) listeria *Lm*TAG immunization; peaks opening (red), peaks closing (blue). **c**, Principal component analysis of peak accessibility during naive TCR<sub>TAG</sub> cells differentiation in acute infection (green), early tumorigenesis

(blue), and memory TCR<sub>TAG</sub> cells in established hepatocellular carcinomas (red). Circles, with *Lm*TAG immunization; diamonds, no *Lm*TAg immunization. **d**, Chromatin accessibility heat map. Each row represents 1 of 11,698 selected peaks (differentially accessible between any sequential cell comparison;  $FDR < 0.05$ ,  $\log_2(\text{FC}) > 2$ ). Shown are  $\pm 1$  kb from the peak summit (2 kb total per region). **e**, ATAC-seq signal profiles of *Pdcld1*, *Ctla4*, *Cd38*, *Tcf7*, and *Ifng* genes of naive (N; grey), memory (M; green), L7, L14, L35 (blue series), and ML7, ML14, and ML35 (red series) TCR<sub>TAG</sub> cells. Pink boxes highlight peaks that become accessible in dysfunctional T cells compared to naive and memory; blue boxes highlight peaks that become inaccessible in dysfunctional TCR<sub>TAG</sub> cells compared to naive and memory TCR<sub>TAG</sub> cells. **f**, CD38, CD101, CD30L, and CD5 expression on ML7, ML14, ML21. Inset numbers show MFI.



Extended Data Figure 10 | See next page for caption.

**Extended Data Figure 10 | Chromatin states of human PD1<sup>hi</sup> tumour-infiltrating CD8<sup>+</sup> T cells and model for CD8 TST differentiation and dysfunction in tumours.** **a**, Sorting scheme of peripheral blood lymphocytes for naive (N), effector memory (EM), central memory (CM) CD8 T cell populations (left), and PD1<sup>hi</sup> CD8 TIL from patients with melanoma or non-small-cell lung cancer. **b**, Differentially accessible ATAC-seq peaks grouped by DESeq-defined differential accessibility pattern. Each column represents one biological replicate. Samples shown include CD45RA<sup>+</sup>CD45RO<sup>-</sup> (naive; grey), CD45RA<sup>-</sup>CD45RO<sup>+</sup>CD62L<sup>-</sup> (effector memory; light green) and CD45RA<sup>-</sup>CD45RO<sup>+</sup>CD62L<sup>+</sup> (central memory; dark green) peripheral blood CD8<sup>+</sup> T cells from healthy donors,

and CD45RA<sup>-</sup>CD45RO<sup>+</sup>PD1<sup>hi</sup>CD8<sup>+</sup> T cells isolated and flow-sorted from human melanoma and lung tumours (PD1<sup>hi</sup> TIL; blue). Open, accessible chromatin regions are presented in red; inaccessible chromatin regions are presented in blue. **c**, ATAC-seq signal profiles of *SELL* in naive, effector memory, and central memory. Blue boxes highlight peaks that remain accessible in central memory or become inaccessible in effector member compared to naive respectively. **d**, ATAC-seq signal profiles of *IFNG*, *EGR2*, *CD5*, and *CTLA4*. Pink and blue boxes highlight peaks that become accessible or inaccessible in PD1<sup>hi</sup> TIL compared to naive or central memory, respectively. **e**, Model for tumour-specific CD8 T cell differentiation and dysfunction in tumours.