Chromatin states define tumour-specific T cell dysfunction and reprogramming

Mary Philip1, Lauren Fairchild2,3, Liping Sun4, Ellen L. Horste1, Steven Camara1, Mojdeh Shakiba1,5, Andrew C. Scott1,5, Agnes Viale6, Peter Laufer6, Taha Merghoub5,7, Matthew D. Hellmann5,8, Jedd D. Wolchok5,7,9, Christina S. Leslie2 & Andrea Schietinger1,5

Tumour-specific CD8 T cells (TST) are often found within solid tumours, but tumours progress despite their presence, suggesting that these TST are dysfunctional. 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; however, important questions remain, including how to predict which patients will respond to therapy and precisely which TST mediate clinical responses3–5. 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 tumours8. 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 division6. 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 described7–10. 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)11 to assess genome-wide chromatin accessibility changes during T cell differentiation in tumours compared to acute infection.

C8 D T cell chromatin changes during infection

We transferred congenically marked naive (N; CD44loCD62Lhi) TCRγδ cells (specific for SV40 large T antigen epitope I (TAG))12 from TCRγδ transgenic mice into wild-type C57BL/6 mice, which were immunized one day later with a recombinant Listeria monocytogenes strain expressing TAG (LmTAG)6,13. TCRγδ 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 transcription factors (TBET), cytotoxic molecules (GZMB, CD107), and pro-inflammatory cytokines (IFNγ, TNFα) (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 naive T cells15,16.

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 Project17. Early-response genes showed much fewer changes compared to late-response genes (Extended Data Fig. 3c). For example, Ldhα (encoding LDHA, needed for the metabolic shift to aerobic glycolysis and IFNγ production18) and Mki67 (encoding Ki67, required for chromosome segregation during mitosis19) 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 naive T cells20. 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’
Article research

is a tumour-specific antigen. AST-Cre-ERT2 mice initially develop T cell differentiation (Extended Data Fig. 4b, top). This pattern was also seen in functional CD8 peaks present in all CD8 T cells were in promoter regions (Extended intronic and intergenic regions (potential enhancer peaks), whereas later dysfunctional state 2, established by day 14 and persisting thereafter. Changes occurred, even after progression to established tumours at day 60+. 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 TAG cells in malignant lesions followed a distinct epigenetic trajectory compared to TCR TAG 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 TAG 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. 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 Cila4, 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. 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 TAG cells correlated temporally with our previous observation that L8 but not L35 were capable of regaining effector function. Indeed, when we re-isolated TCR TAG 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, L5 and L7 regained the ability to produce IFNγ and TNFα, but TCR TAG 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 targets changed preferentially during differentiation from N to E5 (Extended Data Fig. 5a). Predicted NFATC1-binding sites, including those in genes encoding inhibitory receptors and negative regulators, such as Cila4, 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. 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).

We next assessed chromatin-state dynamics in TST in the course of tumorigenesis using the previously described tamoxifen- inducible, autochthonous liver cancer model (AST-Cre-ER22) in which TAG is a tumour-specific antigen. AST-Cre-ER22 mice initially develop pre-malignant lesions which eventually progress into hepatocellular carcinoma (by day 60–90). We transferred congenically marked naive TCR TAG cells (N, the same as N in Fig. 1a) into AST-Cre-ER22 mice one day before tamoxifen administration and then analysed TCR TAG cells at different time points (Fig. 2a). Liver-infiltrating TCR TAG cells downregulated CD62L, uniformly expressed activation markers CD44 and inhibitory receptors PD1 and LAG3, and failed to produce IFNγ or TNFα (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 TAG cells in malignant lesions followed a distinct epigenetic trajectory compared to TCR TAG 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 TAG 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. 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 Cila4, 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. 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 TAG cells correlated temporally with our previous observation that L8 but not L35 were capable of regaining effector function. Indeed, when we re-isolated TCR TAG 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, L5 and L7 regained the ability to produce IFNγ and TNFα, but TCR TAG 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 Itgα and Cd28 were downregulated together with closing of multiple peaks within their loci (Extended Data Fig. 6d).

We next used an in vitro 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, 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 inhibitor, enhances differentiation of CD8 T cell to memory cells through WNT/ TCF1 activation, and we therefore treated TCR TAG-adoptively trans-

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.

25 MAY 2017 | VOL 545 | NATURE | 453
We identified membrane protein genes with chromatin states 1 or 2 and thus might predict reprogrammability. We next looked for cell surface proteins with expression that correlated with chromatin states through discrete chromatin states. Figure 2 highlights.

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 TCRTAG cells (Fig. 3c) and found several markers not previously associated with tumour-induced T cell dysfunction. State 1 (L5, L7) TCRTAG cells had low expression of CD38, CD101, and CD30L and high expression of CD5, whereas state 2 (L14, L28) TCRTAG 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 significant enrichment.

Figure 2 and Figure 3 captions are not transcribed here as they are not essential for understanding the content of the text.
a similar binary expression in early and late TST (Extended Data Fig. 8c). Moreover, TCR\textsubscript{TAG} 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\textsuperscript{lo}CD101\textsuperscript{lo} and CD38\textsuperscript{hi}CD101\textsuperscript{hi} TST from PD1\textsuperscript{hi} L14 cells and assessed reprogrammability (3 days \textit{in vitro} IL-15). CD38\textsuperscript{lo}CD101\textsuperscript{lo} L14 regained the ability to produce IFN-\gamma and TNF-\alpha, but CD38\textsuperscript{hi}CD101\textsuperscript{hi} 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 OTI CD8 T cells (TCROT\textsubscript{1} cells). Naive congenerically marked TCR\textsubscript{OT1} cells were adoptively transferred into B16-OVA tumour-bearing B6 mice. Tumour-infiltrating TCR\textsubscript{OT1} 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\textsubscript{OT1} cells expressed high levels of CD38 and CD101 and downregulated CD5 compared to early dysfunctional day 5 TCR\textsubscript{OT1} cells (Extended Data Fig. 8e). Moreover, late dysfunctional TCR\textsubscript{OT1} cells at day 25 could not regain the ability to produce IFN-\gamma or TNF-\alpha, in contrast to early dysfunctional TCR\textsubscript{OT1} cells at day 5 (Extended Data Fig. 8f).

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

We next tested whether the ‘functionally poised’ state present in memory T cells (M; Fig. 1) could prevent them from becoming dysfunctional in tumours. TCR\textsubscript{TAG} memory cells were transferred into AST-Alb-Cre mice (in which hepatocytes express TAG from birth) bearing established hepatocellular carcinomas and, one day later, were immunized with Lm\textsubscript{TAG} (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, Cldn4, Cd38, Tcf7, and Ifng (Extended Data Fig. 9e). Changes in surface protein expression (CD38, CD301, CD101, 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 Lm\textsubscript{TAG} 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\textsuperscript{+}CD45RO\textsuperscript{−}), effector memory (EM; CD45RA\textsuperscript{−}CD45RO\textsuperscript{−}CD62L\textsuperscript{lo}), and central memory (CM; CD45RA\textsuperscript{−}CD45RO\textsuperscript{−}CD62L\textsuperscript{hi}) CD8 peripheral blood lymphocytes from healthy donors and PD1\textsuperscript{hi} 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\textsuperscript{hi} TIL uniquely gained and lost multiple peaks, for example in Ifng, Egr2, Cdt5, 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 human melanoma and non-small-cell lung cancer tumours.
human peripheral blood lymphocytes and PD1hi TIL and found that human PD1hi TIL had the greatest overlap in peak accessibility changes with dysfunctional (state 2) mouse TST (Fig. 5b). For example, the TCF7/TcF2 locus showed similar intergenic and intronic peak accessibility changes in human PD1hi TIL and mouse state-2 TCR-tag Cells (Fig. 5c). A subset of PD1hi 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 tumors.

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 induction1,12. We identified core elements shared between mouse fixed dysfunctional TST and human PD1hi 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 PD1hi T cells within heterogeneous TIL populations, a finding of important potential clinical relevance, and human PD1hi TIL showed heterogeneous expression of these markers. In patients who do not respond to immune checkpoint blockade (non-responders), PD1hi TIL may be in a fixed dysfunctional state, in contrast to responders whose PD1hi TIL are in a plastic state, amenable to reprogramming. Our studies on the epigenetic and transcriptional programs underlying TST dysfunctional states and therapeutic reprogrammablility 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.

1. 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).
2. Khalil, D. N., Smith, E. L., Brentjens, R. J. & Wolchok, J. D. The future of cancer immunotherapy. Nat. Methods 10, 1213–1218 (2013).
3. Stavely-O’Carroll, K. et al. In vivo ligation of CD40 enhances priming against the endogenous tumor antigen and promotes CD8+ T cell effector function in SV40 T antigen transgenic mice. J. Immunol. 171, 697–707 (2003).
4. Brockstedt, D. G. et al. Listeria-based cancer vaccines that integrate innate immunogenicity from toxicity. Proc. Natl Acad. Sci. USA 101, 13832–13837 (2004).
5. Anders, S. & Huber, W. Differential expression analysis for sequence count data. Genome Biol. 11, R106 (2010).
6. Aradhya, S., Fann, M., Wiersto, K., 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).
7. Denton, A. E., Russ, B. E., Dull, T. P., Cui, W. & Greenberg, P. D. Differentiation-dependent functional and epigenetic landscapes for cytokine genes in virus-specific CD8+ T cells. Proc. Natl Acad. Sci. USA 108, 15306–15311 (2011).
8. Best, J. A. et al. Transcriptional insights into the CD8+ T cell response to infection and memory T cell formation. Nat. Immunol. 14, 404–412 (2013).
9. Peng, M. et al. Aerobic glycolysis promotes T helper 1 cell differentiation through an epigenetic mechanism. Science 354, 481–484 (2016).
10. Cuylen, S. et al. Ki-67 acts as a biological surfactant to disperse mitotic chromosomes. Nature 535, 308–312 (2016).
11. Keach, S. M. & Cui, W. Transcriptional control of effector and memory CD8+ T cell differentiation. Nat. Rev. Immunol. 12, 749–761 (2012).
12. Sen, D. R. et al. The epigenetic landscape of T cell exhaustion. Science 354, 1159–1169 (2016).
13. Paakko, K. E. et al. Epigenetic stability of exhausted T cells limits durability of reinvigoration by PD-1 blockade. Science 354, 1160–1165 (2016).
14. Scott-Browne, J. P. et al. Dynamic changes in chromatin accessibility occur in CD8+ T cells responding to viral infection. Immunity 45, 1327–1340 (2016).
15. Macian, F. NFAT proteins: key regulators of T-cell development and function. Nat. Rev. Immunol. 5, 472–484 (2005).
16. Martínez, G. J. et al. The transcription factor NFAT promotes exhaustion of activated CD8+ T cells. Immunity 42, 265–278 (2015).
17. Teague, R. M. et al. Interleukin-15 rescues tolerant CD8+ T cells for use in adoptive immunotherapy of established tumors. Nat. Med. 12, 335–341 (2006).
18. Li, Y. et al. MART-1-specific melanoma tumor-infiltrating lymphocytes maintaining CD28 expression have improved survival and expansion capability following antigen restimulation in vitro. J. Immunol. 184, 452–465 (2010).
19. Blattman, J. N., (to A.S.), and the MSKCC Core Grant P30 CA008748. The Integrated Genomics Program is supported by NIH-NCI grants R00 CA172371 (to A.S.), K08 CA158069 (to M.P.), 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 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. performed all high-throughput 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 gene expression data and performed all high-throughput computational analyses; 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. 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. TCRαβ-negative mice (B6.Cg-Tg[TcraY1, TcrbY1]A16Tev/J), Cre-ERT2 (B6.129-Gt(ROSA)26SorGnat(tetr/ERT2))f1/J, Alb-Cre (B6.Cg-Tg(Alb-cre)21Mgn/J), TCR-OT1 (C57BL/6-Tg[TcraTcrb]100Mjb/J), Ly-5.1 (B6.SJL-Ptprc8 Pepe8/JBoy), and C57BL/6 Thy1.1 mice were purchased from The Jackson Laboratory. TCRαβ mice were crossed to Thy.1 mice to generate TCRαβ Thy.1.1 mice. TCR-OT1 were crossed to Ly5.1 mice to generate TCR-OT1 Ly5.1 mice. Alb-ER (Albumin-foxStop-SV40 large T antigen (TAG)) f1 were crossed to Cre-ERf2 or Alb-Cre mice to obtain Alb-ERf2 and Alb-ERf2 mice, respectively. 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. Tamanoxifen (Sigma) stock solution was prepared by warming tamanoxifen 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−1 in corn oil). A single dose of tamanoxifen (1 mg) was administered intraperitoneally (i.p.) into Alb-Cre−/− 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 × 10^6 congenerically marked splenocytes and incubated with Tag-1 peptide (0.5 µg ml−1) or OVA peptide (0.1 µg ml−1) 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 IFNγ (XM16.2) and TNFα (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) Sacta Janib strain expressing the Tag-1 epitope (SAINNFAQKL, SV40 large T antigen[26–21]) was generated by Aduro Biotech as previously described[1]. Experimental vaccination stocks were prepared by growing bacteria to early stationary phase, washing in phosphate buffered saline, formulated at approximately 1 × 10^9 colony-forming units (cfu) ml−1, and stored at −80 °C. Mice were infected i.p. with 5 × 10^5 c.f.u. of LmTAG.

Adoptive T cell transfer. For the generation of effector and memory TCRαβ CD8+ T cells, 10^6 CD8+ splenocytes from TCRαβ Thy1.1 transgenic mice were adoptively transferred into B6 (Thy1.2) mice; one day later, mice were infected with 5 × 10^5 c.f.u. of LmTAG. Effector TCRαβ CD8+ T cells were isolated from the spleens of B6 host mice and analysed 5 or 7 days after LmTAG immunization; memory TCRαβ CD8+ T cells were isolated from spleens of B6 host mice and analysed at least 2–3 months after LmTAG immunization. For the transfer of naive TCRαβ T cells into AST-Cre−/− mice, 1 × 10^6 to 2.5 × 10^6 CD8+ splenocytes from TCRαβ Thy1.1 transgenic mice were adoptively transferred into AST-Cre−/− mice; 1 day later, mice were treated with 1 mg tamanoxifen. 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.

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 naive, 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+ PD-CD8+ T cells were flow-sorted for subsequent ATAC-seq analysis.

Statistical analyses. Statistical analyses on flow cytometric data were performed using unpaired 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 TCRαβ Thy1.1+ T cells were sorted by flow cytometry (CD8+ CD44hi) from spleens of TCRαβ Thy1.1 transgenic mice. (ii) Day 5 and day 7 effector, and memory TCRαβ Thy1.1+ T cells were sorted by flow cytometry (CD8+ Thy1.1+) from spleens of infected B6 (Thy1.2) host mice (see above) 5 and 7 days or 2–3 months after listeria infection. (iii) TCRαβ Thy1.1+ T cells from pre/early malignant liver lesions: naive TCRαβ Thy1.1+ T cells were adoptively transferred into AST-Cre−/− mice. 1 day later, mice were given 1 mg tamanoxifen i.p. At given time points after tamanoxifen treatment, T cells were isolated and sorted (CD8+ Thy1.1+) from livers as described above. (iv) TCRαβ Thy1.1+ memory T cells from established hepatocellular carcinomas in AST-Alb-Cre mice: TCRαβ Thy1.1 memory T cells were isolated from tumours and flow sorted (CD6 CD8 Thy1.1+) 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 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[11]. 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 × 10^6 paired reads was generated per sample.

ATAC data and preprocessing. Raw ATAC-seq reads were trimmed and filtered for quality using Trim Galore! v.0.40 (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Read data was adapted with adapterRemoval v1.8.1 (http://dx.doi.org/10.14866/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) among 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 adapters in the Tn5 translocation step[12], 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[13]. 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^{-5} using the Benjamini–Hochberg procedure for multiple hypothesis correction. A single peak may be caused by noise in the assay and not reflect true chromatin accessibility; we calculated an irreproducible discovery rate (IDR)[39] 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 × 10^{-4}) 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) were combined. There was greater variation between cell lines, 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 ReSeq 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 ReSeq transcript annotation.

ATAC-seq peak atlas summary. We found a total of 76,899 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 systematic 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 \(^\text{40}\). 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. \text{41}). Differential accessibility of these peaks was then calculated for all pairwise comparisons of cell types using DESeq2 v1.6.3 (ref. \text{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 ±1 kb around the peak summit in 20 bp bins. To improve visibility, bins with read counts greater than the 75th percentile + 1.5 x 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 genometools \(^\text{42}\). For ATAC-seq samples, this coverage was calculated after shifting the reads to account for the 5′-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) \(^\text{43}\).

Transcription factor peak assignment. Using the MEME \(^\text{44}\)-curated CisBP \(^\text{45}\) transcription factor binding motif (TFBM) reference, we scanned the mouse ATAC-seq atlas and confirmed that the set of peaks mapped by bnMapper and by the UCSC liftOver tool were nearly identical (57,383 out of 76,869 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 few 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(FC) were calculated between human N and human EM, CM or PD1 \(^\text{46}\) TIL versus log(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 Trimmomatic \(v0.39\) \((\text{http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/})\), powered by CutAdapter v1.8.1 \((\text{http://dx.doi.org/10.14806/ej.17.1.200})\) and FastQC v0.11.3 \((\text{http://www.bioinformatics.babraham.ac.uk/projects/fastqc/})\). Paired-end reads were aligned using STAR \(v4.3.4\) against either mm10 or hg38. The ReSeq 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. \text{41}). Differential expression of genes across cell types was calculated using DESeq2 v1.6.3 (ref. \text{40}). FDR correction of 0.05 was imposed unless otherwise stated. A log2 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 \(\text{46}\). 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 0106020 (membrane proteins). The top 50 most up- and downregulated genes (size-factor normalized RPMK) when compared between L5–L7 and L14–L60 were plotted in a heat map (size-factor normalized). Proteome expression was assessed by flow cytometry for those membrane proteins for which monoclonal antibodies were available. Mouse targets (clone; supplier): CD3 (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. Mouse 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+ 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\textsubscript{TAG} CD8 T cells differentiating to effector and memory T cells during acute Listeria infection. Naive TCR\textsubscript{TAG} cells (N; Thy1.1\textsuperscript{+}) were transferred into B6 (Thy1.2\textsuperscript{+}) mice, which were immunized with Lm\textsubscript{TAG} one day later. At days 5, 7, and 60+ after Lm\textsubscript{TAG}, 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\textsubscript{α}, TBET, and GZMB expression directly \textit{ex vivo} (upper panel; inset numbers show MFI), and intracellular IFN\textsubscript{γ} and TNF\textsubscript{α} production and CD107 expression after 4 h of \textit{ex vivo} TAG peptide stimulation (lower panel). Flow plots are gated on CD8\textsuperscript{+} Thy1.1\textsuperscript{+} 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 ± 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 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 log2 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-TAG cells during acute LmTAG infection. c, d, Epigenetic and transcriptional regulation of early CD8 response genes in TCR-TAG 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-TAG 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 log2(FC) > 2, log2(FC) = 1–2, and log2(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. "Common" peaks were determined for all cell type comparisons, "variable" peaks were significant in at least one cell type comparison (FDR < 0.05, log2(FC) > 1).

c, Venn diagrams show the number of significantly changed peaks during the transition from naive (N) to day 5 effectors (E5) TCRTAG cells during acute listeria LmTAG infection versus N to L5 early malignant lesion-infiltrating TCRTAG cells (FDR < 0.05, log2(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. 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, log2(FC) > 1). Common: no change in any cell type comparison.

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.
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) TCRγδ T cells during acute listeria LmTAG infection versus N to L5 in pre-malignant lesions (FDR < 0.05, log₂(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, log₂(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 naive 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 ± s.e.m. shown; ****P ≤ 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 TAG cells (Thy1.1⁺) were transferred into AST-Cre-ERT2 (Thy1.2⁺) 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⁻¹ 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 TAG 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 TAG cells. c, Production of IFNγ and TNFα by TCR TAG 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 ± s.e.m.; P values calculated using unpaired two-tailed t-test. d, Representative flow cytometric analysis of CD38 and CD101 expression of TCR TAG cells (numbers indicate %); CD38, CD101, and CD5 expression. Each symbol represents an individual mouse. Data show mean ± 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 ± s.e.m. (bottom panel). **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 ± s.e.m. shown; *P = 0.03, **P = 0.002, ***P ≤ 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\textsubscript{TAG} cells differentiating to the dysfunctional state in solid tumours. a, PD1 and LAG3 expression and cytokine production of memory TCR\textsubscript{TAG} cells in liver tumours. Each symbol represents individual mouse. Mean ± 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\textsubscript{TAG} 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 (−) listeria \textit{Lm}\textsubscript{TAG} immunization; peaks opening (red), peaks closing (blue). c, Principal component analysis of peak accessibility during naive TCR\textsubscript{TAG} cells differentiation in acute infection (green), early tumorigenesis (blue), and memory TCR\textsubscript{TAG} cells in established hepatocellular carcinomas (red). Circles, with \textit{Lm}\textsubscript{TAG} immunization; diamonds, no \textit{Lm}\textsubscript{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, log2(FC) > 2). Shown are ±1 kb from the peak summit (2 kb total per region). e, ATAC-seq signal profiles of \textit{Pdcd1}, \textit{Ctla4}, \textit{Cd38}, \textit{Tcf7}, and \textit{Ifng} genes of naive (N; grey), memory (M; green), L7, L14, L35 (blue series), and ML7, ML14, and ML35 (red series) TCR\textsubscript{TAG} 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\textsubscript{TAG} cells compared to naive and memory TCR\textsubscript{TAG} cells. f, CD38, CD101, CD30L, and CD5 expression on ML7, ML14, ML21. Inset numbers show MFI.

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.
Extended Data Figure 10 | See next page for caption.
Extended Data Figure 10 | Chromatin states of human PD1\textsuperscript{hi} tumour-infiltrating CD8\textsuperscript{+} T cells and model for CD8 TST differentiation and dysfunction in tumours. 

\textbf{a}, Sorting scheme of peripheral blood lymphocytes for naive (N), effector memory (EM), central memory (CM) CD8 T cell populations (left), and PD1\textsuperscript{hi} CD8 TIL from patients with melanoma or non-small-cell lung cancer. 

\textbf{b}, Differentially accessible ATAC-seq peaks grouped by DESeq-defined differential accessibility pattern. Each column represents one biological replicate. Samples shown include CD45RA\textsuperscript{+}CD45RO\textsuperscript{−} (naive; grey), CD45RA\textsuperscript{−}CD45RO\textsuperscript{+}CD62L\textsuperscript{−} (effector memory; light green) and CD45RA\textsuperscript{−}CD45RO\textsuperscript{+}CD62L\textsuperscript{+} (central memory; dark green) peripheral blood CD8\textsuperscript{+} T cells from healthy donors, and CD45RA\textsuperscript{−}CD45RO\textsuperscript{+}PD1\textsuperscript{hi}CD8\textsuperscript{+} T cells isolated and flow-sorted from human melanoma and lung tumours (PD1\textsuperscript{hi} TIL; blue). Open, accessible chromatin regions are presented in red; inaccessible chromatin regions are presented in blue. 

\textbf{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. 

\textbf{d}, ATAC-seq signal profiles of IFNG, EGR2, CD5, and CTLA4. Pink and blue boxes highlight peaks that become accessible or inaccessible in PD1\textsuperscript{hi} TIL compared to naive or central memory, respectively. 

\textbf{e}, Model for tumour-specific CD8 T cell differentiation and dysfunction in tumours.