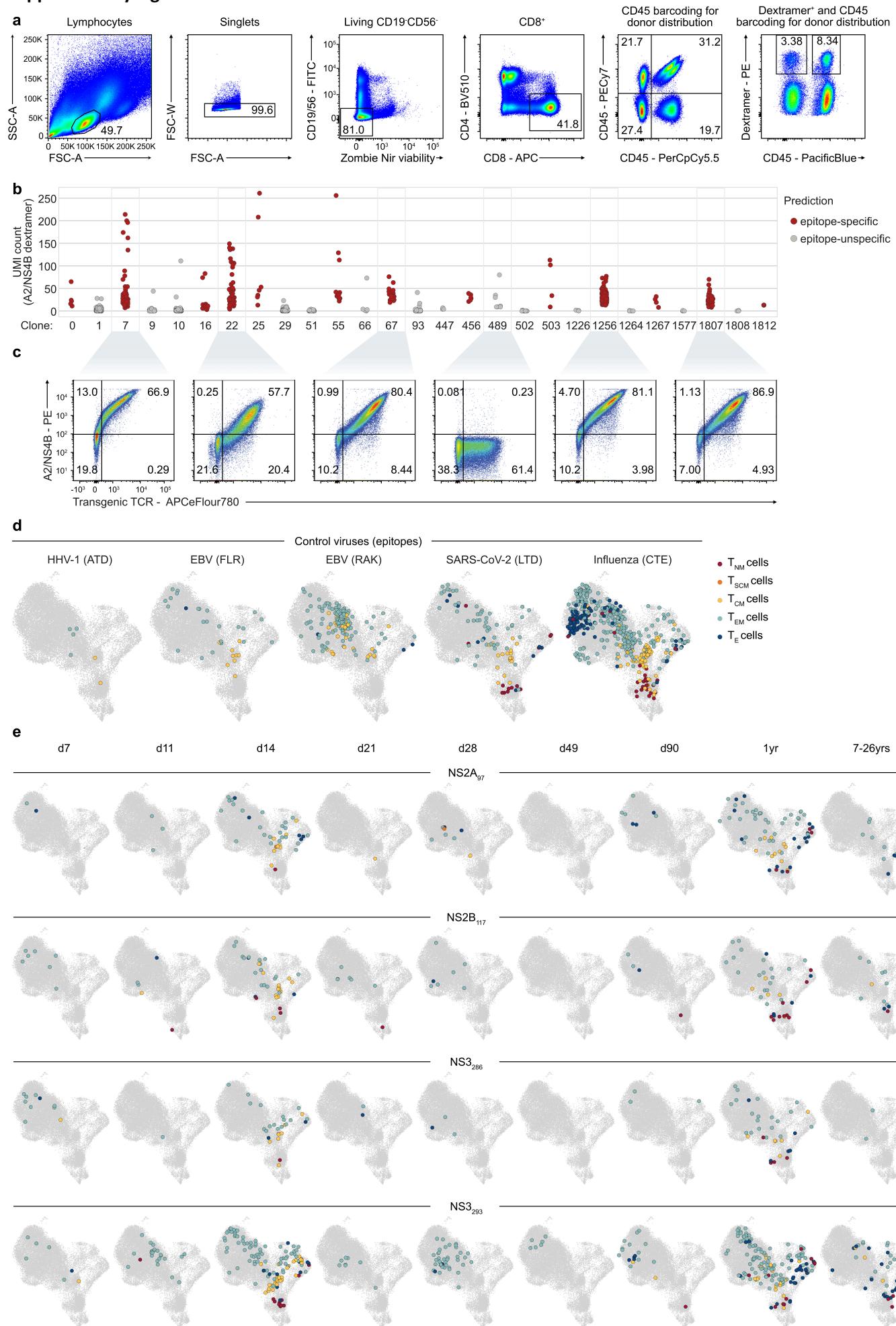


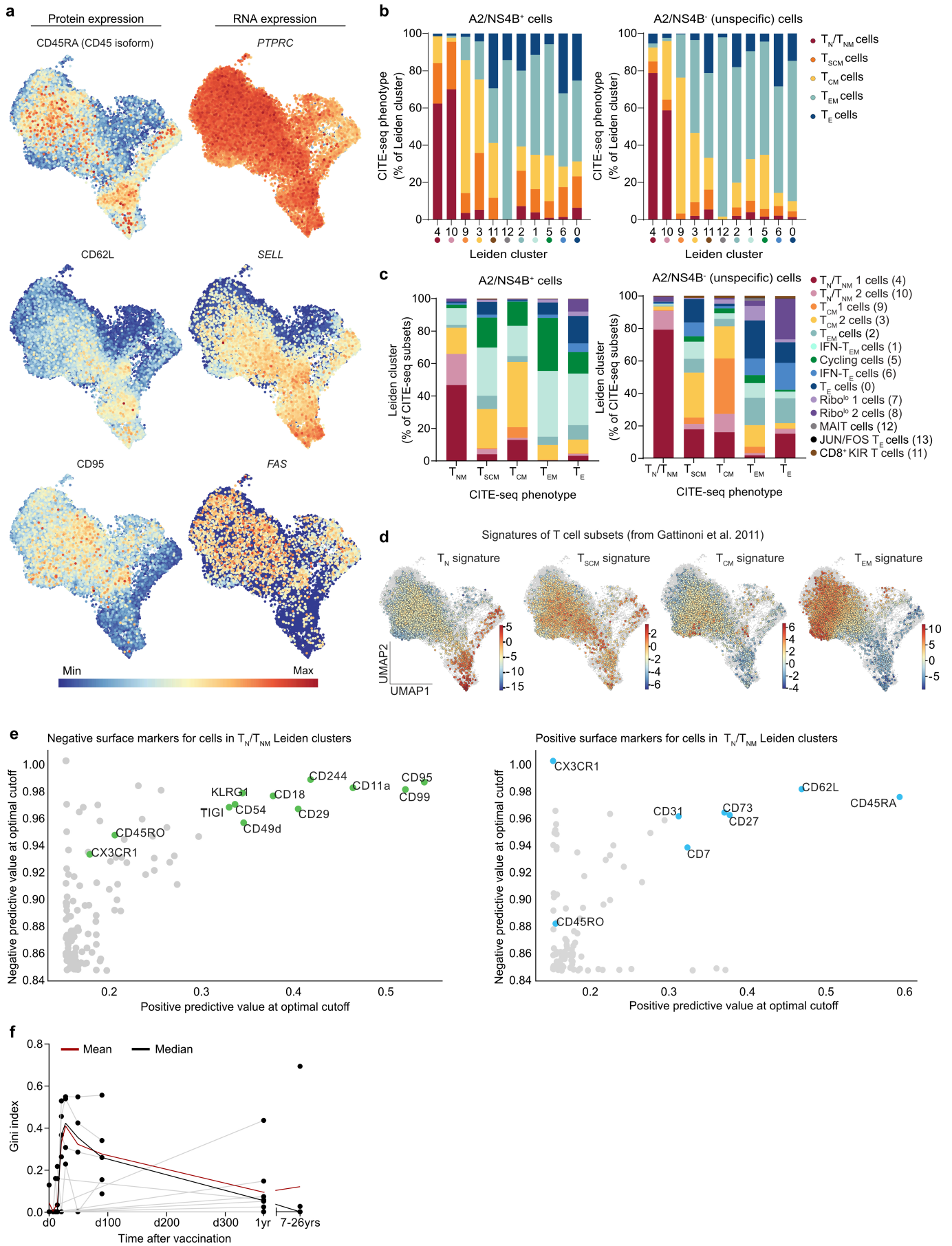
Metabolic quiescence of naive-like memory T cells precedes and maintains antigen-specific T cell memory

In the format provided by the
authors and unedited

Supplementary Figures:



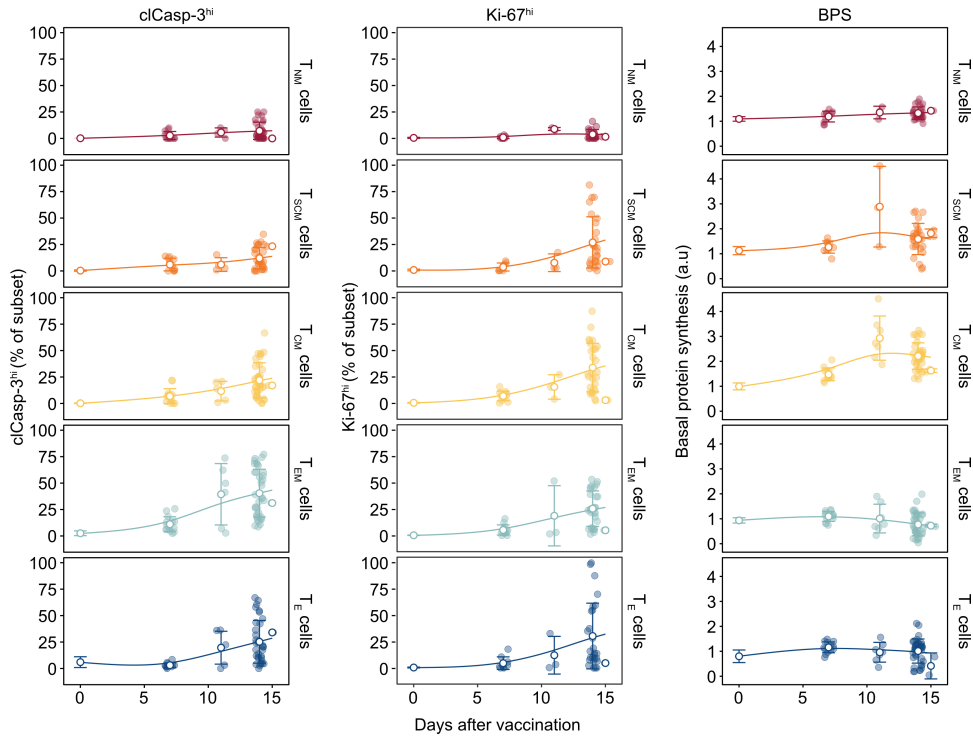
Supplementary Fig. 1: Identification of YFV-specific CD8⁺ T cell subsets by dextramer staining in scRNA-seq. **a**, Representative gating strategy to enrich dextramer⁺ YFV-specific CD8⁺ T cells for scRNA-seq via flow cytometric cell sorting. **b**, Strip plot showing A2/NS4B dextramer-derived UMI counts per cell for a selection of representative T cell clones analyzed in scRNA-seq. Specificity of individual clones was predicted by a computational algorithm (see methods). Specificity prediction, A2/NS4B-specific (red) or A2/NS4B-unspecific (grey). Data points, single cells. **c**, Flow cytometric plots for co-staining of TCR and A2/NS4B-pHLA in Jurkat cells expressing the highlighted TCR shown in **b**. **d**, scRNA-seq-derived UMAP with highlighted dextramer⁺CD8⁺ T cells, specific for control epitopes at individual timepoints post-YFV vaccination. Colors, CITE-seq phenotype (CD62L⁺CD45RA⁺CD95⁻ T_{NM}, CD62L⁺CD45RA⁺CD95⁺ T_{SCM}, CD62L⁺CD45RA⁻ T_{CM}, CD62L⁻CD45RA⁻ T_{EM}, CD62L⁻CD45RA⁺ T_E cells) of epitope-specific cells. Grey, cells without indicated epitope-specificity. **e**, scRNA-seq-derived UMAP with highlighted dextramer⁺CD8⁺ T cells, specific for different YFV epitopes at individual timepoints post-YFV vaccination. Legend as in **d**.



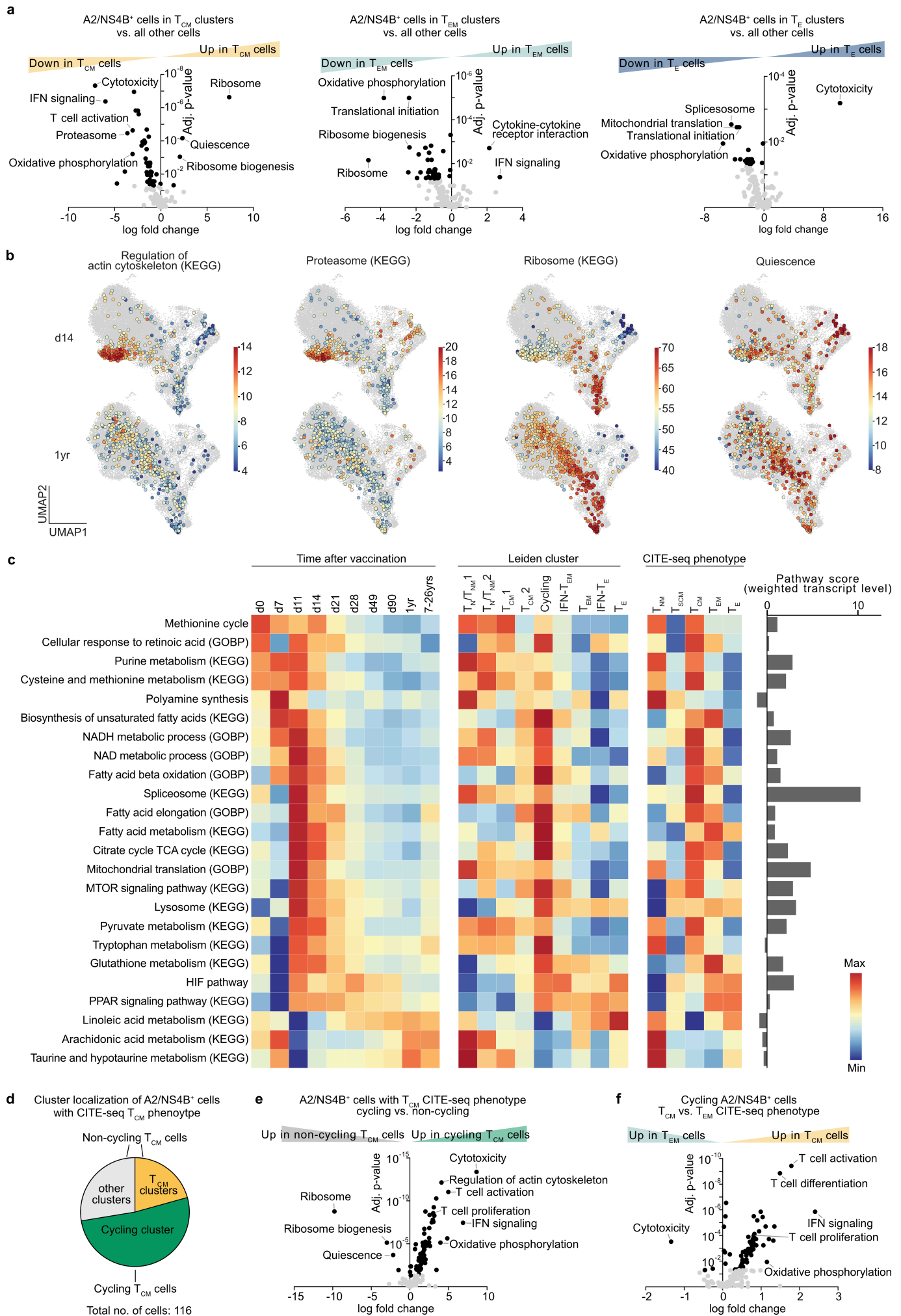
Supplementary Fig. 2: Surface protein and RNA expression of CD8⁺ T cells and diversity of A2/NS4B⁺ cells in scRNA-seq experiments.

a, Log-normalized expression of selected proteins, identified by CITE-seq, and corresponding RNA expression from scRNA-seq in all CD8⁺ T cells. **b**, Distribution of A2/NS4B⁺CD8⁺ and unspecific A2/NS4B⁺CD8⁺ T cell subsets (defined by CITE-seq phenotype (CD62L⁺CD45RA⁺CD95⁻T_{NM}, CD62L⁺CD45RA⁺CD95⁺T_{SCM}, CD62L⁺CD45RA⁻T_{CM}, CD62L⁻CD45RA⁻T_{EM}, CD62L⁻CD45RA⁺T_E cells)) per Leiden cluster (*MKI67*^{hi} cycling, *SELL*^{hi} T_N/T_{NM}, *IL7R*^{hi} T_{CM}, *GZMK*^{hi} T_{EM}, *CCL5*^{hi} T_E cells +/- IFN-sensing signature). X-axis label, Leiden cluster. Bars, mean over all donors. n=3-12 donors. **c**, Quantification of A2/NS4B⁺CD8⁺ and unspecific A2/NS4B⁺CD8⁺ T cell subsets (defined by Leiden cluster (defined in b) localization) per CITE-seq phenotype (defined in b). X-axis label, CITE-seq phenotype. Number in brackets, Leiden cluster number. Bars, mean over all donors. n=3-12 donors. **d**, Expression of gene signatures associated with indicated T cell subsets in A2/NS4B⁺CD8⁺ T cells. Signatures were derived from published differentially expressed genes of sorted T_N, T_{SCM}, T_{CM} and T_{EM} cells. Colors, scores per cell. Grey, A2/NS4B⁺CD8⁺ cells. **e**, Bioinformatic analysis on 130 surface markers analyzed by CITE-seq in an additional scRNA-seq dataset. For each marker, positive and negative predictive values were calculated to evaluate its performance as a (left) negative or (right) positive marker for cells in the transcriptionally defined T_{NM} Leiden clusters. **f**, Diversity of A2/NS4B⁺CD8⁺ T cell clones shown by the Gini index per timepoint. Data points, individual donors. Grey lines connect different timepoints across one donor. Red and black lines indicate mean and median across all donors, respectively.

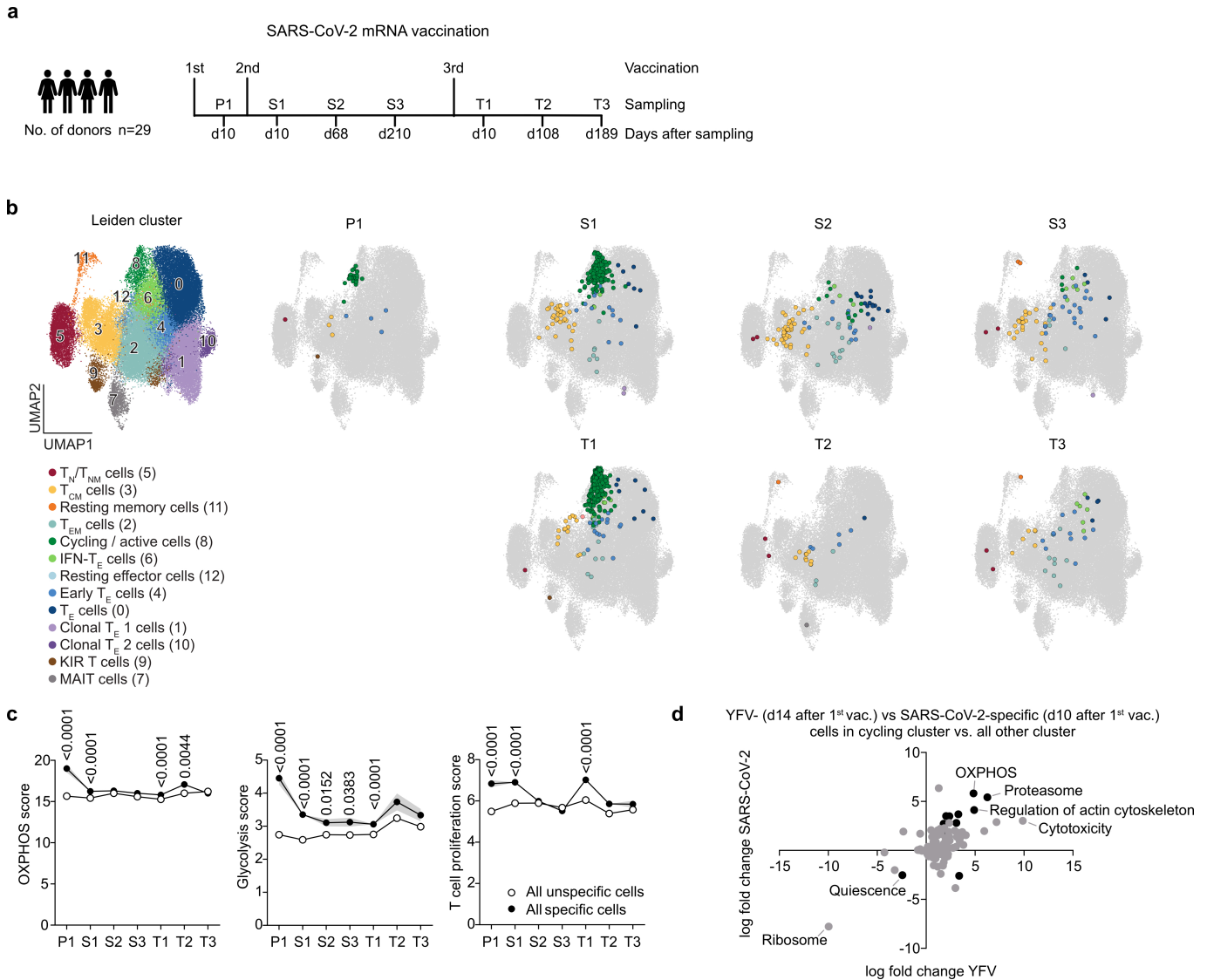
Supplementary Fig. 3: Mathematical modelling of early T cell differentiation dynamics post-YFV vaccination. **a-c**, Mathematical models were analyzed assuming constant parameters (a), Ki-67 (b) or BPS (c) to describe cell proliferation for the individual T cell subsets. Ki-67 and BPS data were derived from flow cytometry on A2/NS4B⁺CD8⁺ T cell subsets (CD62L⁺CD45RA⁺CD95⁻ T_{NM}, CD62L⁺CD45RA⁺CD95⁺ T_{SCM}, CD62L⁺CD45RA⁻ T_{CM}, CD62L⁻CD45RA⁻ T_{EM}, CD62L⁻CD45RA⁺ T_E cells) from 7 to 15 days post-YFV vaccination. Individual data points, donors. Whiskers, mean and SE. Solid lines, best model predictions. Shaded areas, 95%-prediction bands. **d**, Comparing model performance using the corrected Akaike Information Criterion (AICc) across different models (M1–M12), which additionally varied in the consideration of metabolic activity to describe cell proliferation (constant, BPS, Ki-67), revealed that one model with consideration of BPS performed best in explaining the data (M5, see methods for details). **e**, Model scheme depicting the best-performing model M5 given a linear differentiation pathway following T_{NM}–T_{SCM}–T_{CM}–T_{EM}–T_E, where ρ denotes the rate of cell proliferation, μ represents the rate of differentiation and δ defines the rate of cell death. M5 assumes subset-dependent variations in cellular turnover not covered by the metabolic dynamics, with similar values for the factors ρ , μ , and δ across the T_{NM} and T_{SCM} subsets, as well as for T_{CM}, T_{EM}, and T_E, with the exception of a distinct proliferation factor ρ_{EM} for T_{EM}. Cell death and proliferation additionally depend on the subset-specific metabolic dynamics considering cCasp-3 ($f_\delta(t)$) and BPS ($f_\rho(t)$) (see Supplementary Fig. 4). **f**, Estimated rates for cell proliferation, differentiation and cell death for the model used in c. **g**, Corresponding profile likelihoods of parameter estimates indicating identifiability of death and differentiation rates, while cell proliferation cannot be reliably quantified by the approach. N_0 represents the proportion of T_{NM} cells at the start of the modelling process and was fitted without affecting the model's accuracy.



Supplementary Fig. 4: Metabolic activity considered for the mathematical modelling of early T cell differentiation dynamics post-YFV vaccination. Measurements of metabolic activity of cCasp-3, BPS and Ki-67 expression levels (individual data points, donors. Whiskers, mean and SEM) were connected by fitting a cubic spline function, i.e., a polynomial of 3rd degree (solid line) to the data in order to approximate the metabolic dynamics over time. The spline functions for cCasp-3, and BPS or Ki-67 were considered within the mathematical models to describe changes in the cell death and proliferation dynamics, respectively.



Supplementary Fig. 5: Metabolic pathway transcription in YFV-specific CD8⁺ T cells with high resolution. **a**, Transcriptional analysis of metabolic and cellular pathways in A2/NS4B⁺CD8⁺ T cells in the combined clusters for T_{CM}, T_{EM} and T_E cells vs. all other on d14 post-YFV vaccination based on scRNA-seq using Vision. Black, pathways with statistically significant log fold changes (two-sided t-test with benjamini-hochberg correction for multiple testing, p<0.05). Grey, pathways with non-significant changes. **b**, scRNA-seq-derived UMAPs showing the transcriptional scores of selected KEGG pathways and a quiescence score in A2/NS4B⁺CD8⁺ T cells post-YFV vaccination. Colors, score per cell. Grey, unspecific cells. **c**, Transcriptional analysis of metabolic and cellular pathways in A2/NS4B⁺CD8⁺ T cells as in a sorted by timepoints after vaccination (left), localization within selected Leiden clusters (middle; (MKI67^{hi} cycling, SELL^{hi} T_N/T_{NM}, IL7R^{hi} T_{CM}, GZMK^{hi} T_{EM}, CCL5^{hi} T_E cells +/- IFN-sensing signature)) or CITE-seq phenotype (right; (CD62L⁺CD45RA⁺CD95⁻ T_{NM}, CD62L⁺CD45RA⁺CD95⁺ T_{SCM}, CD62L⁺CD45RA⁻ T_{CM}, CD62L⁻CD45RA⁻ T_{EM}, CD62L⁻CD45RA⁺ T_E cells)). Bar graph on the right, weighted transcript level per pathway (see methods section). Selected pathways with significant enrichment (two-sided t-test, p<0.05) in all A2/NS4B⁺CD8⁺ over all A2/NS4B⁺CD8⁺ cells are shown. **d**, Leiden cluster (defined in c) localization of A2/NS4B⁺CD8⁺ cells with T_{CM} CITE-seq phenotype (defined in c) on d14 post-YFV vaccination. Cycling cells, cells within the cycling cluster. Non-cycling cells, cells in T_{CM} cluster (T_{CM}1+T_{CM}2) or in any other cluster. **e**, Transcriptional analysis of metabolic and cellular pathways on d14 after vaccination as in a, in cycling vs. non-cycling A2/NS4B⁺CD8⁺ T cells with T_{CM} CITE-seq phenotype (see d) (two-sided t-test with benjamini-hochberg correction for multiple testing, p<0.05). Color legend as in a. **f**, Transcriptional analysis of metabolic and cellular pathways as in a, in A2/NS4B⁺CD8⁺ T cells in the cycling cluster with T_{CM} vs. T_{EM} CITE-seq phenotype on d14 post-YFV vaccination (two-sided t-test with benjamini-hochberg correction for multiple testing, p<0.05). Color legend as in a.



Supplementary Fig. 6: Transcriptional dynamics of metabolic and cellular pathways in SARS-CoV-2-specific CD8⁺ T cells after SARS-CoV-2 mRNA vaccination. **a** Schematics of the CoVa-Adapt study for CD8⁺ T cell responses after SARS-CoV-2 mRNA vaccination, indicating the nomenclature for each sampling timepoint. Donors were vaccinated with the mRNA vaccine Comirnaty. **b** Transcriptional profile of CD8⁺ T cells enriched for SARS-CoV-2-specific pHLA dextramer-binding cells (n=53,907 cells) determined by scRNA-seq and depicted as UMAP with Leiden cluster annotation (left). Number in brackets, Leiden cluster in UMAP. SARS-CoV-2-specific T cells (HLA-A*01:01/LTD, HLA-A*02:01/YLQ and HLA-A*03:01/KCY combined) after SARS-CoV-2 mRNA vaccination are highlighted (right). Colors, cluster localization. Grey, unspecific CD8⁺ T cells. **c** Quantification of pathway scores for oxidative phosphorylation, glycolysis, and T cell proliferation in SARS-CoV-2-specific vs. unspecific CD8⁺ T cells at indicated timepoints after SARS-CoV-2 vaccination determined by scRNA-seq and Vision. Bold line, mean. Shaded areas, SEM. Statistical testing, two-way ANOVA with Sidák's multiple comparisons test. **d** Correlation of transcriptional signatures of metabolic and cellular pathways in A2/NS4B⁺CD8⁺ (YFV-specific) T cells in the cycling cluster on day 14 post-YFV vaccination and SARS-CoV-2-specific CD8⁺ T cells in the cycling cluster on day 10 post-primary SARS-CoV-2 vaccination. Transcriptional scores were derived from scRNA-seq using Vision. Log fold changes of pathways that were significantly enriched (p<0.05) in YFV-specific cycling cells over all other YFV-specific cells at day 14 post-YFV vaccination are shown for both antigen-specific T cell populations. Black, pathways with statistically significant log fold changes (p<0.05) in both datasets. Grey, with statistically significant log fold changes (p<0.05) only in the YFV dataset.

Supplementary methods:

Epitope assignment - model description: To separate specific antigen-binding cells from unspecific binding and non-binding cells, a Bayesian two-component negative binomial mixture model of the following form was used:

$$\begin{aligned}
 Y_{ij} &\sim z_{noise} \text{NegBinom}(q_{noise}, \alpha) \\
 &\quad + z_{signal} \text{NegBinom}(q_{signal}, \alpha) \quad \forall i \in N \\
 z &\sim \text{Cate}(w) \\
 w &\sim \text{Dir}([1,1]) \\
 q_b &\sim \text{LogN}(\mu, \sigma) \quad \forall b \in \{\text{noise}, \text{signal}\} \\
 \text{s.t. } q_{noise} &< q_{signal} \\
 \frac{1}{\sqrt{\alpha}} &\sim \text{HC}(10) \\
 \mu &\sim N(0, 10) \\
 \sigma &\sim \text{HC}(10),
 \end{aligned}$$

where $i \in N$ indexes UMI count of cell i for antigen j . *NegBinom* is a mean, inverse-dispersion parameterized Negative Binomial distribution. The mean of the Negative Binomial distribution follows a Log-Normal (*LogN*) prior, whose mean and variance follow a Normal (*N*) and Half-Cauchy (*HC*) hyperprior, respectively. An inequality constraint is imposed on q_{signal} and q_{noise} to break the unidentifiability of the model. The prior of the dispersion parameter α follows the recommendation of the Stan Development Team¹. z represents the class latent variable drawn from a categorical distribution (*Cate*), whose prior probability w is Dirichlet distributed (*Dir*) with uninformative hyperprior.

Inference: For parameter estimation, Blackbox Variational Inference (BBVI)² with an Auto-Normal Guide is performed. To this end, z is marginalized. Posterior class probabilities p are then calculated from the estimates for each cell:

$$p(z = 1 | y) = \frac{w_{signal} \cdot p(y | \theta)}{p(y)}$$

BBVI is performed over 5000 iterations with an exponentially decaying learning rate using gradient-clipped Adam (*clip=10*)³. The initial learning rate is set to $1e^{-2}$ and is decayed over *maxiter*/2 iterations with a decay rate of 99% reaching a final learning rate of $1e^{-7}$. Evidence lower bound (ELBO) is estimated with Monte Carlo sampling using 10 particles. Model parameters are initialized using the set with the lowest loss out of 100 initializations sampled from the priors.

Cell assignment: To assign cells to the signal or noise component, the posterior class probability p is compared to a threshold t . To determine t whilst controlling for false positive inflation due to multiple decision-making^{4,5}, Newton's direct posterior probability approach is employed⁶.

Briefly, if p_{ij} approximates the certainty of cell i belonging to the signal component of antigen j , meaning it is specifically binding antigen j , then $1 - p_{ij}$ approximates the probability of a type I error. For a threshold t , we then obtain a set of credible antigen-specific cells $J(t) = \{i | 1 - p_{ij} \leq t\}$. The false discovery rate is approximately given by:

$$\widehat{FDR}(t) = \frac{\sum_{i \in J(t)} 1 - p_{ij}}{|J(t)|}$$

For the desired FDR α , we find the optimal threshold t' that includes as many cells as possible whose approximated FDR does not exceed α .

$$t' = \min_{0 < c < 1; \widehat{FDR}(t) < \alpha} t$$

We finally obtain the set of credibly binding cells $J(t')$.

Implementation: The model, inference, and FDR-control were implemented in Python 3.10 using numpyro 0.16.1⁷, jax 0.4.38⁸, and optax 0.2.4⁹

Mathematical modelling: Mathematical models that describe T cell differentiation dynamics were extended to consider the metabolic dynamics of individual subsets in order to determine if this additional information improves identification and quantification of T cell subset turnover. We assumed a linear differentiation pathway following $T_{NM} \rightarrow T_{SCM} \rightarrow T_{CM} \rightarrow T_{EM} \rightarrow T_E$ comparable to previous approaches¹⁰. With ρ , μ and δ defining the rates of cell proliferation, differentiation and loss, respectively, the system of ordinary differential equations is given by:

$$\begin{aligned}
 \frac{dT_{NM}}{dt} &= (\rho - \delta - \mu)T_{NM} \\
 \frac{dT_{SCM}}{dt} &= \mu T_{NM} + (\rho - \delta - \mu)T_{SCM} \\
 \frac{dT_{CM}}{dt} &= \mu T_{SCM} + (\rho - \delta - \mu)T_{CM} \\
 \frac{dT_{EM}}{dt} &= \mu T_{CM} + (\rho - \delta - \mu)T_{EM} \\
 \frac{dT_E}{dt} &= \mu T_{EM} + (\rho - \delta)T_E
 \end{aligned}$$

To consider the metabolic dynamics measured within the individual cellular subsets, the constant rate parameters within the equations were replaced by $\rho = \rho f_{\rho,i}(t)$ and $\delta = \delta f_{\delta,i}(t)$ with $f_{x,i}(t)$ describing the dynamic changes of basal protein synthesis ($x = \rho$) and cleaved Caspase 3 ($x = \delta$) for each individual T cell subset, which were informed by the measured metabolic changes (see below).

We tested multiple different model structures that differed in the variation of the constant rate parameters ρ , μ and δ . Using the corrected Akaike information criterion to compare model performance in explaining the observed data. We found that a model that assumed similar values of ρ , μ and δ for T_{NM} and T_{SCM} , as well as T_{CM} , T_{EM} , and T_E , with an additional independent proliferation rate ρ for T_{EM} (i.e., termed model M5), performed best in explaining the data when considering metabolic changes for additional subset-specific variations in these parameters (see **Supplementary Fig. 3**). While a model assuming additional differences between the rates for T_{CM} , T_{EM} and T_E improved the explainability of the early increase in T_{EM} and T_E sub-populations, which was not captured by the best modelling approach, this led to larger uncertainty in parameter estimates, indicating that the short time courses considered are insufficient to reliably quantify the dynamics of later differentiation steps.

Fitting and parameter estimation: To parameterize the metabolic dynamics $f_{x,i}(t)$, for each cellular subset and metabolic process, spline-functions were fitted to the measured metabolic expression levels of reactive T cells with $f_{x,i}(0)$ defining the homeostatic level prior immunization given by the average metabolic expression level of corresponding non-reactive T cell subsets (see **Supplementary Fig. 4**). Spline-functions were fitted using the function `smooth.spline` in R with `spar = 0.35`. Complete mathematical models on T cell differentiation dynamics with or without the metabolic dynamics considered by $f_{x,i}(t)$ were developed in CoRC¹¹ and were fitted to the measured subset sizes, which were normalized for each donor by the percentage of all living lymphocytes, using Particle Swarm followed by gradient-based Levenberg-Marquardt algorithm. Optimization boundaries for each parameter were set to [0,5]. The number and type of fitted parameters depended on the assumed model variant (see above) with additional estimation of the precursor frequency of naïve cells for each model. Profile likelihood analyses were performed to assess parameter identifiability and confidence intervals of estimates¹². To visualize model prediction bands, time course simulations for all parameter combinations identified during profile likelihood analysis were performed and point-wise 95%-percentile ranges of model predictions were considered.

References

1. Stan Development Team. Stan Modeling Language Users Guide and Reference Manual. at <https://mc-stan.org>.
2. Wingate, D. & Weber, T. Automated Variational Inference in Probabilistic Programming. *arXiv* 1–7 (2013).
3. Kingma, D. P. & Ba, J. L. Adam: A method for stochastic optimization. *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.* 1–15 (2015).
4. Müller, P., Parmigiani, G. & Rice, K. FDR and Bayesian Multiple Comparisons Rules. *Bayesian Stat.* 8 359–380 (2023) doi:10.1093/oso/9780199214655.003.0014.
5. Scott, J. G. & Berger, J. O. Bayes and empirical-Bayes multiplicity adjustment in the variable-selection problem. *Ann. Stat.* **38**, 2587–2619 (2010).
6. Newton, M. A., Noueiry, A., Sarkar, D. & Ahlquist, P. Detecting differential gene expression with a semiparametric hierarchical mixture method. *Biostatistics* **5**, 155–176 (2004).
7. Phan, D., Pradhan, N. & Jankowiak, M. Composable Effects for Flexible and Accelerated Probabilistic Programming in NumPyro. *arXiv* 1–10 (2019).
8. Bradbury, J. *et al.* JAX: composable transformations of Python+NumPy programs. at <http://github.com/google/jax> (2018).
9. Deepmind *et al.* The DeepMind JAX Ecosystem. at <http://github.com/google-deepmind> (2020).
10. Buchholz, V. R. *et al.* Disparate individual fates compose robust CD8+ T cell immunity. *Science* **340**, 630–5 (2013).
11. Förster, J., Bergmann, F. T. & Pahle, J. CoRC: The COPASI R Connector. *Bioinformatics* **37**, 2778–2779 (2021).
12. Raue, A. *et al.* Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* **25**, 1923–1929 (2009).