Low-dose IL-2 reduces IL-21^{+} T cells and induces a long-lived anti-inflammatory gene expression signature inversely modulated in COVID-19 patients

Jia-Yuan Zhang[1], Fiona Hamey[1], Dominik Trzupek[1], Marius Mickunas[2], Mercede Lee[1], Leila Godfrey[1], Jennie H.M. Yang[2], Marcin L Pekalski[1], Jane Kennet[3], Frank Waldron-Lynch[4], Mark Evans[5], Timothy I. M. Tree[2], Linda S. Wicker[1†], John A. Todd[1†] and Ricardo C. Ferreira[1†]

[1]JDRF/Wellcome Diabetes and Inflammation Laboratory, Wellcome Centre for Human Genetics, Nuffield Department of Medicine, NIHR Oxford Biomedical Research Centre, University of Oxford, Oxford, United Kingdom.

[2]Department of Immunobiology, King's College London, School of Immunology and Microbial Sciences, London, United Kingdom.

[3]Wellcome-MRC Institute of Metabolic Science-Metabolic Research Laboratories and Medical Research Council Metabolic Diseases Unit, University of Cambridge Addenbrooke's Hospital, Cambridge, United Kingdom.

[4]Vertex Pharmaceuticals, Vertex Cell & Gene Therapies, Boston, Massachusetts, United States of America.

[5]National Institute for Health Research, Cambridge Biomedical Research Centre, Addenbrooke's Hospital, University of Cambridge, Cambridge, United Kingdom.

[†]Co-senior authors

Abstract

Despite early clinical success, the mechanisms of action of low-dose interleukin-2 (LD-IL-2) immunotherapy remain only partly understood. Here, we examine the effects of interval administration of low-dose recombinant IL-2 (iLD-IL-2) using high-resolution, single-cell multiomics. We confirmed that iLD-IL-2 selectively expands thymic-derived $FOXP3^{+}HELIOS^{+}$ Tregs and $CD56^{br}$ NK cells, and provide new evidence for an IL-2-induced reduction of highly differentiated IL-21-producing $CD4^{+}$ T cells. We also discovered that iLD-IL-2 induces an anti-inflammatory gene expression signature, which was detected in all T and NK cell subsets even one month after treatment. The same signature was present in COVID-19 patients, but in the opposite direction. These findings indicate that the sustained Treg and $CD56^{br}$ NK cell increases induced by our 4-week iLD-IL2 treatment create a long-lasting and global anti-inflammatory environment, warranting further investigations of the potential clinical benefits of iLD-IL-2 in immunotherapy, including the possibility of reversing the pro-inflammatory environment in COVID-19 patients.

Keywords

Interval administration of low-dose Interleukin-2 (iLD-IL-2); aldesleukin; type 1 diabetes; autoimmunity; single-cell multiomics; Regulatory T cells (Tregs); IL-21; Natural killer (NK) cells; CD56; SARS-CoV-2 (COVID-19); Mucosal-associated invariant T (MAIT) cells.

Main

The use of much lower doses of interleukin-2 (IL-2) than employed in cancer therapy has shown clinical efficacy in a number of inflammatory and autoimmune conditions, by stimulating a subset of $CD4^+$ T cells expressing high levels of the high-affinity trimeric IL-2 receptor, designated as regulatory T cells (Tregs)[1–9]. In the autoimmune disease type 1 diabetes (T1D), the genetic association with the IL-2 pathway[10,11] has provided a strong rationale for the use of low-dose (LD)-IL-2 immunotherapy to correct the pathogenic IL-2 pathway defect. However, initial trials using different dosing protocols showed that, whilst being safe, the LD-IL-2 immunotherapy had limited clinical benefit and variable effects on the levels of Treg induction[12–15]. These findings highlighted the importance of better understanding the mechanisms of action of LD-IL-2 immunotherapy and how different dosing regimens affect treatment efficacy. To address this question, we have previously conducted an adaptive single-dose observational study (DILT1D), which demonstrated the sensitivity and specificity of IL-2 signalling *in vivo* and determined a dosing range of $0.2\text{-}0.47 \times 10^6$ IU/m^2 of recombinant IL-2 (aldesleukin/proleukin) to achieve Treg increases in blood in the order of 20-50% from baseline, without expanding effector T cells (Teffs)[16]. This study also shed light into the critical importance of a dosing interval, as Tregs were found to be desensitised to further IL-2 signalling within 24 h of IL-2 treatment – through an IL-2 dose-dependent downregulation of the signalling IL-2Rβ (CD122) subunit on the surface of Tregs[20]. This observation demonstrated the potential limitation of the commonly-practised daily dosing schedule in T1D[13–15], thereby supporting an interval administration of LD-IL-2 (iLD-IL-2) as a model to maximise the selective effects and potential therapeutic benefits of LD-IL-2 immunotherapy. The efficacy of iLD-IL-2 was tested in a subsequent multiple dosing study (DILfrequency), which determined an optimal dosing regimen of IL-2 administered every three days[17].

Here, we set out to explore in depth the effects of iLD-IL-2. We employed a recently developed targeted multiomics approach based on the BD Rhapsody system, which allowed us to simultaneously quantify 565 mRNA transcripts and 65 surface protein targets at the single-cell level[18]. We performed a detailed immunophenotypic characterization of both T and NK cell

compartments, with a particular emphasis on the $CD4^{+}$ $CD127^{low}CD25^{hi}$ T cell and $CD56^{br}$ NK cell subsets, which have been previously shown to be highly sensitive to iLD-IL-2 treatment[16,17]. Aside from the expected expansions of thymic-derived $FOXP3^{+}HELIOS^{+}$ Tregs and $CD56^{br}$ NK cells, we show that iLD-IL-2 inhibits the differentiation of IL-21-producing cells, a subset that has been previously implicated in the pathogenesis of T1D[19–21]. Unexpectedly, we find that iLD-IL-2 elicits a long-lasting anti-inflammatory gene expression signature in both T and NK cell subsets one month after cessation of treatment. We found the same long-lived anti-inflammatory signature in SARS-CoV2 virus-infected COVID-19 patients, but in the opposite direction to iLD-IL-2. Taken together, our findings indicate that iLD-IL-2 can reduce IL-21 during treatment, and in the longer-term induce an anti-inflammatory environment, which is compromised by SARS-CoV-2 infection.

Results

Interval administration of low-dose IL-2 (iLD-IL-2) induces and maintains increased frequencies of $CD4^+$ $CD127^{low}CD25^{hi}$ T cells and $CD56^{br}$ NK cells

Previously, we designed an adaptive study (DILfrequency) to investigate the safety and efficacy of interval administration of low-dose recombinant IL-2 (iLD-IL-2) to increase Treg numbers in blood, and established an optimal 3-day dosing interval, with doses ranging from 0.20-0.47 x 10^6 IU/m^2 (**Fig. 1a**)[17]. To gain further insight into the mechanism of action of iLD-IL-2, we selected 18 participants treated with this dosing schedule (**Supplementary Table 1**) and characterised the cellular alterations in blood by polychromatic flow-cytometry (FACS). We observed a sustained increase in the frequency of $CD4^+$ $CD127^{low}CD25^{hi}$ T cells (henceforth designated as the $CD4^+$ Treg gate) during the dosing phase, but no alterations in the frequency of the broader $CD4^+$ Teff or $CD8^+$ T cell populations (**Fig. 1b,c**), including in a subset of $CD45RA^-CD62L^-$ effector memory $CD4^+$ T cells (T_{EM}), known to be enriched for previously activated, and potentially self-reactive, effector T cells (**Extended Data Fig. 1a-c**).

Outside the T cell compartment, iLD-IL-2 also caused sustained increased frequencies of a subset of NK cells expressing the highest levels of CD56 ($CD56^{br}$ NK cells; **Fig. 1c**). This effect was even more pronounced than the increase in $CD4^+$ Tregs and is consistent with the relatively high affinity $CD56^{br}$ NK subset for IL-2, especially when compared to the bulk NK population[22,23]. We observed no alteration in the frequency of the classical $CD56^{dim}$ NK cells, representing the majority of $CD56^+$ NK cells in circulation. The IL-2-induced increase in $CD4^+$ Tregs and $CD56^{br}$ NK cells was recapitulated when assessing absolute numbers of cells in circulation, although their relative low abundance in blood resulted in no discernible alteration in the numbers of either $CD4^+$ and $CD8^+$ T cells or $CD56^+$ NK cells (**Fig. 1d,e**). The increase in both frequency and numbers of $CD56^{br}$ (but not $CD56^{dim}$) NK cells in blood was similarly observed in all three IL-2 doses (**Extended Data Fig. 1d-f**).

Investigating IL-2-induced cellular alterations in blood using single-cell multiomics

We next profiled peripheral blood mononuclear cells (PBMCs) from 13 selected participants treated with the three-day interval IL-2 dosing regimen (**Supplementary Table 1**) using a targeted single-cell multiomics approach based on the BD Rhapsody system, which enables the parallel quantification of specific mRNA and cell-surface proteins –via oligo-conjugated antibodies (AbSeq)– in each cell[24]. To increase the power of this approach to identify small, but potentially important, changes elicited by LD-IL-2, we have adopted a cell enrichment strategy to vastly increase and normalise the numbers of $CD4^+$ Tregs and $CD56^{br}$ NK cells profiled in this study, and designed a custom mRNA panel of 565 oligonucleotide probes, assessed in parallel with the expression of 65 cell-surface protein targets (**Supplementary Table 2**)[18]. We isolated fixed proportions of the five assessed immune populations by FACS (30% $CD4^+$ Tregs; 25% $CD4^+$ Teffs and $CD8^+$ T cells; 12% $CD56^{br}$ and 8% $CD56^{dim}$ NK cells) from each participant and profiled both unstimulated and *in vitro* stimulated cells at three specific time points: (i) Day 0, corresponding to the baseline pre-treatment visit; (ii) Day 27, three days after the penultimate IL-2 injection and immediately before the last IL-2 injection; and (iii) Day 55, four weeks after the last IL-2 dose (**Fig. 1f**).

Unsupervised clustering of the 482,531 unstimulated cells passing quality control (QC), and Uniform Manifold Approximation and Projection (UMAP) revealed a clear delineation of the sorted T and NK cell populations, identifying 15 distinct functional clusters (**Fig. 1g**). We observed good overlap among cells from each participant on UMAP plots, with similar representation in the five sorted immune populations (**Extended Data Fig. 2a-c**). A similar result was observed for the clustering of the 323,839 stimulated cells passing QC (**Extended Data Fig. 2d-f**), indicating minimal batch effects on this single-cell multiomics dataset that are usually prevalent in scRNA-seq data.

Interval administration of low-dose IL-2 selectively expands $FOXP3^+HELIOS^+$ Tregs

Next, we sought to investigate whether our iLD-IL-2 regimen altered the relative composition of the cells sorted from the $CD127^{low}CD25^{hi}$ Treg gate (**Extended Data Fig. 3**). Analysis of the expression of the canonical Treg transcription factors *FOXP3* and *IKZF2* (encoding HELIOS)

allowed the stratification of CD127^{low}CD25^{hi} Treg clusters into three functional groups according to their origin (**Extended Data Fig. 4** and Methods). Differential abundance analysis at Day 27 revealed an enrichment in the frequency of naïve FOXP3^{+}HELIOS^{+} Tregs and a concomitant depletion of CD25^{+} FOXP3^{-}HELIOS^{-} Teffs after the conclusion of the four-week IL-2 dosing phase (Day 27; **Fig. 2a,b**), which was replicated in stimulated cells (**Fig. 2c,d**). Given the observed increase in CD127^{low}CD25^{hi} T cells by FACS, we next investigated whether these changes were also reflected in the absolute numbers in circulation. Using data for the whole-blood CD127^{low}CD25^{hi} Treg counts, we converted the frequencies of subsets identified by single-cell analysis into absolute cell numbers and found that the IL-2 induced expansion at Day 27 was restricted to the thymic-derived FOXP3^{+}HELIOS^{+} Treg subsets (**Fig. 2e**). We found very good concordance between unstimulated and stimulated conditions, with naïve Tregs showing >2-fold increase and memory FOXP3^{+}HELIOS^{+} Tregs displaying approximately a 70% increase in cell numbers compared to baseline pre-treatment levels (**Fig. 2e**). In contrast, despite the noticeable increase in the frequency and number of CD127^{low}CD25^{hi} T cells, we obtained no evidence of alterations in the number of activated CD25^{+} Teffs contained in this population, even though they express the high affinity trimeric IL-2 receptor (**Fig. 2e**). Furthermore, we observed no changes in the *in vitro* suppressive capacity of CD4^{+} CD127^{low}CD25^{hi} T cells after IL-2 treatment on a per cell basis compared to the pre-treatment cells (**Fig. 2f**).

The alterations in the relative composition of CD127^{low}CD25^{hi} T cells were found to be transient and were not maintained one month after the last dose of IL-2 (Day 55) in either unstimulated or stimulated cells (**Extended Data Fig. 5a-d**). We also replicated the transient IL-2-induced increase in the frequency of naïve Tregs by FACS at Day 27 by measuring the ratio of naïve:memory Tregs (**Extended Data Fig. 5e-g**). We found concordance between both the ratios of naïve:memory cells, as well as in the frequency of naïve and memory Tregs between the FACS assessment and the single-cell data (**Extended Data Fig. 5h,i**). In contrast to the alterations in the composition of CD127^{low}CD25^{hi} T cells, we observed no notable changes in the relative abundance of CD4^{+} Teffs subsets after IL-2 treatment (**Extended Data Fig. 6**).

Differentiation of IL-21-producing T cells is inhibited by iLD-IL-2 treatment

Among the identified $CD25^+$ Teff clusters sorted from the $CD127^{low}CD25^{hi}$ gate, we noted an IL-2-induced depletion of a subset of cells with a marked T follicular helper (T_{FH}) cell profile, including the expression of the canonical T_{FH} receptors CXCR5, PD-1 and ICOS, and the transcription factor *MAF* ($CD25^+$ T_{FH} (cluster 8); **Extended Data Fig. 5a**). Although T_{FH} cells are rare in blood, within cells sorted from the $CD25^{-/low}$ Teff gate, we were able to characterise an analogous heterogeneous cluster of central memory (T_{CM}) $CXCR5^+$ cells (cluster 1; **Extended Data Fig. 6a,b**), which is known to be enriched with precursors of T_{FH} cells. Upon *in vitro* stimulation, which elicited the expression of the classical T_{FH} cytokine IL-21, we obtained a better discrimination of the T_{FH} clusters. In particular, within cells sorted from the $CD25^{-/low}$ Teff gate, this led to a better differentiation of a cluster of $CXCR5^+$ T_{CM} cells (cluster 2) and a cluster of $CXCR5^{low}$IL-21^+ T cells (cluster 6), reminiscent of a subset of IL-21-producing peripheral T helper cells recently found to be enriched in T1D patients[21] (**Fig. 3a,b and Extended Data Fig. 6d,e**). On the other hand, within the $CD127^{low}CD25^{hi}$ Treg gate, *in vitro* stimulation allowed the stratification of the activated $CD25^+$ T_{FH} cells into two separate clusters (cluster 10 and 11) based on their expression of IL-21 (**Fig. 3b**). We note that a related cluster of $FOXP3^+$ T follicular regulatory cells (T_{FR}, cluster 11) which, in addition to the T_{FH} signature, also featured a distinct Treg transcriptional profile as well as the upregulation of *POU2AF1* and *CCR9*, was not altered by iLD-IL-2 treatment (**Extended Data Fig. 5a**).

We next compared the transcriptional profile of the four identified T_{FH}-like clusters marked by the specific upregulation of IL-21 expression. We identified a continuum of T_{FH} cell differentiation among these clusters, with the $CD25^+$IL-21^+ T_{FH} subset representing the terminal stage of maturation in blood, as illustrated by their classical T_{FH} profile and increased IL-21 production (**Fig. 3a,b**). Furthermore, we found similar profiles between the $CXCR5^+$ T_{CM} and the $CD25^+$ T_{FH} clusters sorted from the $CD25^{-/low}$ Teff and $CD25^+$ Treg gates, respectively, with the latter likely representing an earlier stage of T_{FH}-cell differentiation. In agreement with this shared identity and developmental stage, the $CXCR5^+$ T_{CM} and $CD25^+$ T_{FH} subsets displayed largely overlapping UMAP embeddings among all stimulated $CD4^+$ T cells (**Fig. 3c,d**). In contrast to these three conventional T_{FH}-like subsets, which seem to share a similar differentiation pathway, the $CXCR5^{low}$IL-21^+ $CD25^{-/low}$ Teff cluster represented a much more heterogeneous cluster with

a scattered clustering profile and containing a subset of cells with activated/exhausted T_{FH} profile and increased IL-21 production (**Fig. 3c,d**). Differential abundance analysis revealed the depletion of both *in vitro* stimulated $CD25^+$ T_{FH} subsets identified from the $CD127^{low}CD25^+$ gate, with a particularly noticeable reduction of the more differentiated $CD25^+IL\text{-}21^+$ T_{FH} subset (cluster 11; **Extended Data Fig. 5b**). In addition, we found similar evidence for an IL-2-induced depletion of the $CXCR5^{low}IL\text{-}21^+$ $CD25^{-/low}$ Teff subset (cluster 6; **Extended Data Fig. 6f**), which was the only cluster within the $CD25^{-/low}$ Teff gate that showed evidence of alteration with IL-2 treatment.

To better discriminate the effect of IL-2 on the differentiation of IL-21-producing cells, we then stratified the cells in these four T clusters according to their expression of IL-21 to define activated T_{FH} profiles (**Fig. 3e**). Consistent with the results from the differential abundance analysis, we found evidence for a reduction in the relative proportion of $IL\text{-}21^+$ cells at Day 27 in the two subsets with higher frequencies of $IL\text{-}21^+$ cells (**Fig. 3f**), which further supports a broader previously uncharacterized effect of iLD-IL-2 immunotherapy in inhibiting the differentiation of IL-21-producing cells. In addition to the depletion of IL-21-producing cells, we also obtained some evidence for the reduction of another subset marked by the expression of many pro-inflammatory cytokines, including GM-CSF and IL-2, as well as the classical type-2 inflammatory cytokines IL-4 and IL-13 ($CD25^+$ T_{H2} T_{EM}, cluster 6).

Effects of iLD-IL-2 on $CD8^+$ T and $CD56^+$ NK cells

In the $CD8^+$ T cell compartment, differential abundance analysis revealed no discernible differences in the relative composition of the conventional αβ $CD8^+$ T cell subsets after IL-2 treatment. However, we identified a previously uncharacterised reduction of two subsets of innate-like $CD8^+$ T cells, a subset of mucosal-associated invariant T cells (MAIT; cluster 4) and a rare subset of $V_{\gamma9}V_{\delta2}$ T cells (cluster 9; **Fig. 4a,b and Extended Data Fig. 7**). In contrast to the IL-2 induced alterations in the $CD4^+$ T cell compartment, the reduction of circulating MAIT and $V_{\gamma9}V_{\delta2}$ T cells was more long-lived and there was still evidence of a reduced frequency in blood one month after the last dose of IL-2 (**Fig. 4c**).

Finally, in the CD56^{+} NK cell compartment, despite the lower level of functional heterogeneity and differentiation (**Fig. 4d,e**), we found a large increase in the proportion of CD56^{br} NK cells after IL-2 treatment, particularly in the more activated HLA-II^{+} subset (**Fig. 4f**), but no alterations in the relative composition of CD56^{dim} subsets (**Fig. 4g** and **Extended Data Fig. 8**). In agreement with the CD4^{+} T cell compartment, the alterations in the CD56^{br} NK cells were also transient and were not observed one month after treatment (**Extended Data Fig. 8b,c**).

Differential expression analysis reveals a selective modulation of CD25^{hi} Tregs and CD56^{br} NK cells by iLD-IL-2 treatment

We next investigated whether in addition to cellular alterations, we could detect IL-2-induced gene expression changes at Day 27. We identified 40 differentially-expressed genes in unstimulated and stimulated cells (**Fig. 5a,b**), mostly found in cells sorted from the CD127^{low}CD25^{hi} Treg and CD56^{br} NK cell gates, indicating that the transcriptional changes induced by one month of IL-2 treatment are restricted to the more IL-2-sensitive subsets. In contrast, we observed little or no changes in the transcriptional profile of either CD4^{+} and CD8^{+} effector T cells, or in CD56^{dim} NK cells, including in stimulated cells (**Fig. 5a,b**). The differential expression analysis at Day 27 revealed three distinct signatures: (i) a Treg signature marked by the upregulation of Treg-specific markers and a downregulation of effector markers; (ii) a CD56^{br} NK cell signature marked by the upregulation of CD56^{br}-specific markers and a downregulation of classical CD16^{+}CD56^{dim} NK cells cytotoxic markers; and (iii) and a cell-cycle signature marked by the specific downregulation of cell-cycle genes. In addition, the pattern of differential expression and the identified gene expression signatures were consistent between unstimulated and stimulated cells (**Fig. 5c**).

Whilst the Treg and CD56^{br} NK cell-specific signatures were restricted to the same respective subsets, the cell-cycle signature was shared between CD25^{+} Tregs and CD56^{br} NK cells and reflects a systematic downregulation of cell-cycle genes at Day 27. These results recapitulated the observed depletion in cycling CD127^{low}CD25^{hi} Tregs (**Extended Data Fig. 5a**) and CD56^{br} NK cells (**Fig. 4f**) detected through the differential abundance analysis. Together with previous results showing increased survival of *ex vivo* expanded Tregs in T1D patients treated with LD-

IL-2[25], these data indicate that the expansion of the FOXP3^{+}HELIOS^{+} Tregs and CD56^{br} NK cells induced by IL-2 treatment is maintained through increased/preferential survival and not active proliferation of these cell subsets after the four-week IL-2 dosing phase.

Interval low-dose IL-2 immunotherapy induces a systemic and long-lasting anti-inflammatory gene expression signature

Contrary to the differential expression on Day 27, which was observed exclusively in the IL-2-sensitive Treg and CD56^{br} subsets, differential expression on Day 55, one month after the last IL-2 injection, was widespread and displayed a consistent pattern in all five assessed immune subsets (**Fig. 6a** and **Extended Data Fig. 9a,b**). These findings indicate that our iLD-IL-2 dosing regimen induces the same "Day 55" gene expression signature in different cell types, which featured most prominently the upregulation of Cytokine Inducible SH2 Containing Protein (*CISH*), a gene encoding a well-characterised negative regulator of cytokine signalling, and the downregulation of Amphiregulin (*AREG*), a secreted TNF-inducible protein with pleiotropic roles in inflammation[26]. We also noted an enrichment of genes associated with cytokine signalling, most notably in the tumour necrosis factor (TNF) signalling pathway, which in addition to *CISH* and *AREG*, also included *TNFSF14*, *TNFSF10*, *STAT1*, *SGK1*, *NFKBIZ*, *NFKBIA*, *DUSP2*, *DUSP4*, *DUSP5*, *RGS1* and *TNFAIP3* (**Fig. 6a**). Furthermore, we observed a downregulation of several TNF-induced genes that play a central role in curtailing TNF signalling, including the inhibitors of NFkB, *NFKBIZ* and *NFKBIA*, *RGS1*, *TNFAIP3* and *AREG*. Consistent with this observation, Oncostatin M (*OSM*), a pro-inflammatory cytokine shown to be increased in inflammatory bowel disease patients with poor response to anti-TNF therapy[27], was also downregulated by iLD-IL-2. Increased expression of *TNFSF14*, which encodes LIGHT, is notable since decreased expression is associated with higher susceptibility to multiple sclerosis[28] and in a mouse model of multiple sclerosis, LIGHT expression in the brain limits disease severity[29]. The widespread detection of this signature in all five T and NK cell populations assessed in this study indicates that this effect is not driven by IL-2 signalling directly but is likely the result of a long-lasting regulatory environment induced by iLD-IL-2 immunotherapy.

Although this Day 55 expression signature was detected in 11 of the 12 participants with Day 55 data, we observed heterogeneity among participants for the expression levels of the same genes on Day 27. Some participants exhibited concordant expression changes on Day 27 and Day d5, compared to Day 0, while others showed discordant expression profiles, with *CISH* being downregulated and *AREG* upregulated on Day 27 (**Fig. 6b**). Moreover, by principal component analysis (PCA), we found that this concordant-discordant axis, captured by the first principal component (PC1), was able to largely explain the inter-individual heterogeneity of Day 55 signature genes on Day 27 and appeared to be associated with IL-2 dose (**Extended Data Fig. 9c**). Considering this, we derived a linear score to quantify the changes in the Day 55 signature for each participant, referred to as the Day 55 signature score, with higher scores reflecting higher expression of signature genes upregulated on Day 55 (e.g. *CISH*), and lower expression of signature genes downregulated on Day 55 (e.g. *AREG*). As expected, the Day 55 signature score showed considerable inter-individual variation (**Fig. 6c**) that was consistent across different cell types (**Extended Fig. 9d**). This enabled us to correlate changes in signature score, which indicated the direction and amplitude of the modulation of Day 55 genes, with IL-2 dose level. We confirmed that participants receiving higher doses of IL-2 were more likely to have increased signature scores on Day 27 and Day 55 compared the baseline, indicating concordant expression changes, whilst participants receiving lower doses tended to show discordant changes (**Fig. 6d,e**). The dose-dependence of the Day 55 signature reached statistical significance only when comparing to the participant-specific baseline, but not when inspecting the cross-sectional data at each time point (**Extended Fig. 9e**). This underscores the importance of our longitudinal study design that enhanced our statistical power to identify differential expression patterns.

The long-lived anti-inflammatory Day 55 signature is inversely modulated in COVID-19 patients

To extend the evidence for the existence of the Day 55 gene expression signature, we sought independent evidence of a similar transcriptional response with another immune challenge. Recently, the Oxford COVID-19 Multi-omics Blood Atlas (COMBAT) study[30] has generated a large single-cell dataset illustrating the changes induced by COVID-19 infection in blood. A

COVID-19-induced gene expression signature identified in this study, designated as Component 187, featured the differential expression of several Day 55 signature genes, including *CISH* and *AREG*. Of the 1,419 genes contributing to the COVID-19 Component 187, 77 were present in our transcriptional panel. Notably, we found that virtually all signature genes upregulated on Day 55 had negative loading scores in the COVID-19 Component 187, and vice versa (**Fig. 6f**), indicating a strong inverse correlation between the differential expression induced by IL-2 treatment and COVID-19. Differential expression analysis of Day 55 signature genes in the COMBAT dataset confirmed this inverted signature across multiple immune cell types, including T and NK cells (**Extended Data Fig. 9f**), which implies related physiological mechanisms underlying the regulation of similar sets of target genes in these two very different immune challenges. The effect sizes of differential expression after IL-2 treatment were typically smaller than after SARS-CoV-2 infection, although for four participants (P4, P5, P11 and P12), two of whom from the 0.47×10^6 IU/m^2 dose group, the effect sizes were comparable, with CISH and AREG showing over two-fold changes (**Fig. 6b; Extended Data Fig. 9f**). Furthermore, among the 77 COVID-19 patients with single-cell data available (**Supplementary Table 3**), we observed a positive correlation between the Component 187 sample loading scores and time since recording of first symptoms in the COVD-19 patients (**Fig. 6g**), indicating that the observed gene expression changes are a long-term feature of both iLD-IL-2 and COVID-19, but in opposite directions.

To further validate the IL-2-induced dose-dependent expression changes of Day 55 signature genes, we re-analysed our dataset using Sparse Decomposition of Arrays (SDA)[31], which was applied in the COMBAT study[30] to identify COVID-19-related gene expression. From 18 robust components identified by SDA, we observed a single component, Component 48, that was induced specifically in Day 55 samples across most participants (**Extended Data Fig. 10a**). In agreement with the transcriptional changes on Day 55, Component 48 was observed predominantly in unstimulated rather than stimulated cells, and ubiquitously in all major cell types (**Extended Data Fig. 10b**). Furthermore, akin to the Day 55 signature scores, the Component 48 sample loading scores exhibited variable changes on Day 27, but consistent induction on Day 55 (**Extended Data Fig. 10c**), with the direction and magnitude of the changes correlating with IL-2 dose (**Extended Data Fig. 10d**). Moreover, the gene loading scores of Component 48 were also correlated with the loading scores from Component 187 in the

COMBAT dataset, revealing additional genes that were likely differentially expressed, such as *KBTBD7* and *IL12A* (**Extended Data Fig. 10e**).

Discussion

Our findings not only extend previous studies[13–15,32] but also provide new insight into the mechanism of action of LD-IL-2 that has practical implications for its clinical use: firstly, we show that LD-IL-2 administered every three days for a period of one month is able to selectively increase the frequency and number of thymic-derived FOXP3^{+}HELIOS^{+} Tregs, while showing no detectable impact on the expansion or activation of conventional effector T or CD56^{dim} NK cells. We observed a particularly pronounced increase in the frequency of naïve FOXP3^{+}HELIOS^{+} Tregs after iLD-IL-2 therapy, indicating that the Treg increases are caused by an increased representation of thymic-derived Tregs, and not the proliferation of peripherally-induced Tregs. These findings are consistent with the increased clonal diversity in T cells following an escalating LD-IL-2 regimen in graft-versus-host disease patients[33,34]. Secondly, we also noted that our iLD-IL-2 schedule did not induce a detectable cytotoxic gene expression signature in any T cell subset, indicating that among T cells iLD-IL-2 has a Treg-specific effect and does not lead to the activation of a pro-inflammatory response. In addition to Tregs, the CD56^{br} NK cell population is also probably contributing to the overall anti-inflammatory effects of iLD-IL-2, using an immunoregulatory function previously described[35,36].

Thirdly, iLD-IL-2 was able to block the differentiation of IL-21-producing T cells. Although there is ample literature supporting the role of IL-2 signalling in inhibiting the differentiation of T_{FH} cells[37], to our knowledge, this is the first report providing direct evidence that LD-IL-2 immunotherapy can reduce the differentiation of IL-21-producing cells *in vivo*. Previously, we have shown that a two-fold increase in IL-2, but not IL-21, was responsible for the protective effect conferred by the *Idd3* locus - containing *IL2* and *IL21* - in the nonobese diabetic (NOD) mouse model of T1D[38], and independently that a reduction in IL-21 correlated with increased IL-2[39]. In humans, IL-21 has been implicated in the pathogenesis of several autoimmune diseases[40,41], particularly T1D[19–21]. Together, these observations lend further support for the use of iLD-IL-2 in T1D as well as in other diseases associated with increased IL-21 production, such as systemic lupus erythematosus, rheumatoid arthritis, and psoriasis. In support of this therapeutic strategy, an anti-IL-21 monoclonal antibody showed promising results in preserving beta-cell function[42]. Note that we observed no evidence for the alteration of the frequency of

FOXP3$^+$ T_{FR} cells, which has been previously shown to play a role in the regulation of T_{FH} cells and the germinal centre reaction[43,44].

In addition to these cellular alterations, in this study we show that iLD-IL2 also induces a long-lived gene expression signature. The transient nature of the cellular alterations reported in this study, together with the widespread detection of this gene expression signature in all assessed immune subsets, supports a novel immunoregulatory mechanism of iLD-IL-2, whereby the induction of a prolonged anti-inflammatory environment leads to the remodelling of the immune transcriptional profile, perhaps associated with a more regulatory response. We note that the identification of this long-term gene expression signature was only possible due to the inclusion of samples taken one month after the last dose of IL-2, which have not been systematically profiled in other studies of LD-IL-2 immunotherapy. Although the effects of this signature are not yet understood, it is marked by the distinct upregulation of *CISH*, a well-characterised negative regulator of cytokine signalling whose expression is known to be induced by IL-2, as well as other cytokines, signalling through the activation of STAT5[45]. In addition, *CISH* expression has been reported to be involved in the regulation of homeostatic cytokine responses, and loss of *CISH* leads to a dysregulated immune response and increased cytokine signalling that is consistent with its role as a central regulator of the immune response[46]. In contrast, among the downregulated genes, we observed several negative regulators of an inflammatory response at Day 55, including key TNF-induced genes such as *TNFAIP3*, *RGS1* and *AREG*, suggesting that iLD-IL-2 can decrease the homeostatic levels of TNF for at least one month past the cessation of dosing. Plasma levels of TNF coupled to transcriptomics in different disease settings should be examined in future studies.

Validation of the physiological relevance of the iLD-IL2-induced long-lived gene expression signature was obtained from analysis of the Oxford COMBAT dataset[30], where we identified an analogous gene expression signature, showing consistent opposite direction of expression following infection with SARS-CoV-2. The appearance of this signature was most pronounced within a group of recovered COVID-19 patients sampled at least one week after the onset of symptoms. These data support a model whereby an adaptation to a sustained regulatory (iLD-IL-2 treatment), or inflammatory (severe virus infection), environment can remodel the immune transcriptional profile. In contrast to iLD-IL-2, we observed a marked upregulation of the same

negative regulators of a pro-inflammatory response in COVID-19 patients. This observation is consistent with the increased levels of plasma pro-inflammatory mediators such as IL-6, GM-CSF and TNF during the acute phase of the infection[30]. Although the levels of these plasma markers are normalised in the community COVID-19 group, the induction of the gene expression signature is more pronounced, therefore indicating the stability and longevity of this transcriptional signature.

Further longitudinal analyses of both patients treated with iLD-IL-2 and convalescent COVID-19 patients are required to measure the longevity and clinical significance of this signature. One possible explanation for the longevity of the signature is that TNF and IL-2 are among the cytokines known to bind to the extracellular matrix[47] and that the retention of the cytokines within tissues extends far beyond their initial increases. The establishment of the Day 55 IL-2 signature is dose-dependent at Day 27, possibly reflecting a dose-dependent accumulation of injected IL-2 on the extracellular matrix within the interstitial space following subcutaneous dosing. A long-lasting immune alteration that also possibly reflects a continuing inflammatory stimulus or long-lived pro-inflammatory cytokines on the extracellular matrix has been reported in mild COVID-19 patients, with infection inducing a pro-inflammatory response in monocyte-derived macrophages that continues to be detected 3-5 months following SARS-CoV-2 infection[48]. In a separate study, pro-inflammatory markers such as IL-8 and sTIM-3 were still elevated four months after the cessation of COVID-19 symptoms in patients having had mild or moderate disease, but the increases resolved eight months post-infection except in patients with long COVID[49]. These data support the homeostatic regulation of the threshold required to elicit a pro-inflammatory response, which is dependent on the interaction between the immune cells and the local environment. Our findings suggest that chronic immune challenges – for example iLD-IL-2 or SARS-CoV-2 infection – can transiently modify this homeostatic regulation, leading to the observed systemic alteration of the transcriptional profile of the immune cells and their ability to respond to subsequent immune challenges. These findings are also consistent with our observation that the pro-inflammatory gene expression signature is progressively induced after infection, and that it could contribute to the long-lasting systemic alterations associated with the persistence of clinical manifestations in long COVID patients.

Taken together, our findings provide a better understanding of the mechanism of action of LD-IL-2 immunotherapy and open several lines of evidence supporting the clinical benefit of using an interval dosing approach instead of the current daily dosing regimen. We note that iLD-IL-2 has recently shown efficacy in systemic lupus erythematosus[6], likely by avoiding the IL-2-induced Treg desensitisation caused by the daily dosing scheme[16]. In the past there have been concerns over the very short half-life of recombinant IL-2 *in vivo* and that NK cell expansion might be detrimental, which have led to the development of genetically-modified IL-2 muteins with increased half-life and preferential binding to the high-affinity trimeric IL-2 receptor in order to reduce $CD56^{br}$ NK cell stimulation. Our results demonstrating a long-lived gene expression signature associated with potential anti-inflammatory effects, detectable even one month after the end of aldesleukin treatment, alleviate some of these concerns and support the therapeutic potential of IL-2 that is not engineered for longer *in vivo* half-life or preferential binding to the high affinity IL-2 receptor, given intermittently and at low doses. Our study adds to the good track record of using low-dose recombinant IL-2 and confirms its very strong safety profile, which is critical for the potential long-term use of LD-IL-2 immunotherapy for diseases such as T1D where the target patient population will consist mainly of young children, including the possible future use of iLD-IL-2 treatment in the prevention of T1D diagnosis, as demonstrated in recent teplizumab (anti-CD3) trials[50] and for T1D treatment with monoclonal antibodies such as golimumab (anti-TNF)[51]. Additionally, iLD-IL-2 treatment could be trialled in recovering COVID-19 patients to determine if such treatment hastens the restoration of normal immune homeostasis and perhaps reduces long COVID occurrence.

Methods

Study design and participants

Study participants comprised 18 T1D patients enrolled in the DILfrequency study, a single-centre, response-adaptive trial of repeated doses of recombinant IL-2 (aldesleukin) administered using an interval dosing approach[17]. All 18 study participants were included in the 3-day interval group and were treated with IL-2 with doses ranging from 0.2 to 0.47 × 10^6 IU/m^2. Study participants' characteristics and immune profiling at baseline are summarised in **Supplementary Table 1**.

T and NK cell Immunophenotyping

Flow cytometric characterisation of the DILfrequency study participants was performed at each visit on whole blood samples collected within 4 h of phlebotomy, as previously described[20]. Briefly, 150 ul whole blood were incubated for 45 min at room temperature with fluorochrome-conjugated antibodies (**Supplementary Table 2**). Red blood cells were then lysed (BD FACS Lysing solution) and immunostained samples were acquired on a BD Fortessa (BD Biosciences) flow cytometer with FACSDiva software (BD Biosciences) and analysed using FlowJo (Tree Star, Inc.).

In vitro Treg suppression assay

A total of 12 participants were selected for this analysis, including nine participants that were also included in the single-cell analyses. For each participant, cells sampled from three time points were used: Day 0, Day 24 and Day 55. Cryopreserved PBMCs were thawed at 37°C and resuspended drop-by-drop in X-VIVO 15 (Lonza) with 1% heat-inactivated, filtered human AB serum (Sigma). After washing, cells were stained with a cocktail of monoclonal antibodies (**Supplementary Table 2**) for cell sorting. Suppression assays were established in V-bottom 96-well plates by sorting 500 $CD4^+CD25^{-/lo}CD127^+$ effector T cells (Teffs) in the presence or

absence of $CD4^{+}CD25^{high}CD127^{low}$ Tregs at various ratios (Treg:Teff 0:1, 1:4, 1:2 and 1:1) with 1×10^{3} $CD19^{+}$ B cells, using the BD FACSAria III flow cytometer automated cell deposition unit (BD Biosciences). Cells were sorted into in 150 µl X-VIVO 15 supplemented with Penicillin-Streptomycin-Fungizone (Gibco) and 10% heat-inactivated, filtered human AB sera, stimulated with PHA (4 µg/ml; Alere) and incubated at 37°C, 5% CO_2, for 5 days. Proliferation was assessed by the addition of 0.5 µCi/well [3H] thymidine (PerkinElmer) for the final 20 h of co-culture. Conditions were run in 5 replicates, and proliferation readings (counts per minute [CPM]) averaged. Two samples with averaged proliferation less than 1,000 CPM from the Teff wells alone were excluded from the analysis. The percentage suppression in each culture was calculated using the following formula: percent suppression = 100 − [(CPM in the presence of Tregs ÷ CPM in the absence of Tregs) × 100]. All time points from a participant were analysed concurrently.

Single-cell multiomics: cell preparation and capture

Single-cell multiomics profiling was performed using the BD Rhapsody system on PBMC aliquots from 13 of the DILfrequency participants treated with the 3-day interval dosing schedule (**Supplementary Table 1**) collected at the baseline visit (Day 0), 3 days after the last dose of IL-2, immediately prior to the last IL-2 infusion of the dosing phase (Day 27) and one month after the last dose of IL-2 (Day 55). Cryopreserved PBMCs were thawed at 37°C and resuspended drop-by-drop in X-VIVO 15 with 1% heat-inactivated, filtered human AB serum. After washing in PBS, cells were incubated with Fixable Viability Dye eFluor 780 (eBioscience) for 10 min at room temperature. Cells were then washed with cold PBS + 2% FBS and incubated for 30 min at 4ºC with fluorochrome conjugated antibodies for sorting (**Supplementary Table 2**). Cells were then washed two times and resuspended in cold PBS + 1% FBS for cell sorting at 4ºC in a BD FACSAria Fusion sorter (BD Biosciences). The following T and NK cell subsets were sorted from each participant at each visit: (i) $CD3^{-}CD56^{hi}$ ($CD56^{br}$ NK cells); (ii) $CD3^{-}CD56^{+}$ ($CD56^{dim}$ NK cells); (iii) $CD3^{+}CD8^{+}$ ($CD8^{+}$ T cells); (iv) $CD3^{+}CD4^{+}$ $CD25^{-/low}$ ($CD4^{+}$ Teffs); and (v) $CD3^{+}CD4^{+}$ $CD127^{low}CD25^{hi}$ ($CD4^{+}$ Tregs). We aimed to sort a total of 130,000 cells for the Day 0 and Day 55 visits and 160,000 for the Day 27 visit, with the following proportion of each of

the sorted cell populations: 30% $CD4^+$ Tregs; 25% $CD4^+$ Teffs and $CD8^+$ T cells; 12% $CD56^{br}$ NK cells and 8% $CD56^{dim}$ NK cells.

Following cell sorting each sorted subset was labelled with the respective sample barcoding antibody (sample multiplexing kit; BD Biosciences) for 20 min at 4ºC to allow for the identification of the sorting gate and visit of each captured single cell. To minimise day-to-day variation, we processed the three PBMC samples corresponding to the three visits from each participant on the same day. As the number of barcoded oligo-conjugated antibodies available in the kit was less than the total number of sorted subsets, we combined the $CD4^+$ Tregs and $CD56^{dim}$ NK cell subsets as well as the $CD4^+$ Teffs and $CD56^{br}$ NK cells from each visit and labelled the pooled cells with the same barcode. These combined subsets were chosen for their highly differentiated transcriptional and proteomics profile, which allowed robust separation of the respective original subsets at the analysis stage. Cells were then washed three times with BD Sample Buffer (BD Biosciences) at 4ºC to remove any residual unbound barcoding antibody and then pooled together for downstream processing.

To profile unstimulated cells, half of the pooled cells volume (corresponding to ~210,000 cells) were initially incubated at 4ºC for 5 min with human Fc block (BD Biosciences) and then for a further 45 min with a master mix of 65 oligo-conjugated AbSeq antibodies (BD Bioscience; **Supplementary Table 2**). Cells were then washed three times with BD Sample Buffer at 4ºC to remove residual unbound oligo-conjugated AbSeq antibodies, resuspended in 500 ul cold BD Sample buffer and filtered through a sterile 50 um filcon cup-type filter (BD Biosciences) for cell counting. Samples were then resuspended in 620 µl of cold BD sample buffer at a concentration of 40 cells/µl - for an estimated capture rate of ~20,000 single-cells / cartridge - and immediately loaded on two BD Rhapsody cartridges (BD Biosciences) for single-cell capture.

For *in vitro* stimulation, the remaining volume from the pooled cells was washed at 4ºC and resuspended in X-VIVO 15 + 10% heat-inactivated, filtered human AB serum at a final concentration of 2×10^6 cells/ml. A volume of 100 µl (~200,000 cells) was then incubated in one well of a round-bottom 96-well plate at 37ºC for 90 minutes with a PMA and ionomycin cell stimulation cocktail (eBioscience), in the absence of protein transport inhibitors. Cells were

harvested into a FACS tube, washed with PBS + 2% FBS and processed for cell capture, as detailed above for the unstimulated cells.

We note that the pre-purification of the Treg and CD56^{br} NK subsets was central to the resolution of this approach, providing greater power to dissect their functional heterogeneity and to investigate how their relative composition was modified by LD-IL-2 immunotherapy. A consequence of the cell-subset enrichment strategy used in this study is that the frequency of the T and NK cell subsets described in this dataset do not directly represent their frequency in whole blood. The aim of this study was not to identify changes in the global composition of PBMCs, but to specifically focus on the T and NK cell subsets.

Single-cell multiomics: cDNA library preparation and sequencing

Single-cell capture and cDNA library preparation was performed using the BD Rhapsody Express Single-cell analysis system (BD Biosciences), according to the manufacturer's instructions. cDNA was initially amplified for 11 cycles (PCR1) using the pre-designed Human T-cell Expression primer panel (BD Biosciences) containing 259 primer pairs, together with a custom designed primer panel containing 306 primer pairs (BD Biosciences), described previously[18]. A total of 565 primer pairs targeting 534 different genes were targeted for amplification with this targeted panel (**Supplementary Table 2**).

The resulting PCR1 products were purified using AMPure XP magnetic beads (Beckman Coulter) and the respective mRNA and AbSeq/sample tag products were separated based on size selection, using different bead ratios (0.7× and 1.2×, respectively). The purified mRNA and sample tag PCR1 products were further amplified (10 cycles) using a nested PCR approach and the resulting PCR2 products purified by size selection (0.8× and 1.2× for the mRNA and sample tag libraries, respectively). The concentration, size and integrity of the resulting PCR products was assessed using both Qubit (High Sensitivity dsDNA kit; Thermo Fisher) and the Agilent 4200 Tapestation system (Agilent). The final products were normalised to 2.5 ng/µl (mRNA), 1.1 ng/µl (sample tag) and 0.35 ng/µl (AbSeq) and underwent a final round of amplification (six cycles for mRNA and sample tag and 7 cycles for AbSeq) using indexes for Illumina sequencing

to prepare the final libraries. Final libraries were quantified using Qubit and Agilent Tapestation and pooled (~29/67/4% mRNA/AbSeq/sample tag ratio) to achieve a final concentration of 5 nM. Final pooled libraries were spiked with 15% PhiX control DNA to increase sequence complexity and sequenced (150 bp paired-end) on a NovaSeq 6000 sequencer (Illumina).

Pre-processing of single-cell sequencing data

The single-cell sequencing reads were processed following instructions from BD Biosciences described earlier[24], with slight modifications. First, low-quality read pairs were removed based on read length, mean base quality score and highest single-nucleotide frequency. High-quality R1 reads were analysed to extract cell labels and unique molecular identifier (UMI) sequences. High-quality R2 reads were aligned to the reference panel sequences using Bowtie2[52]. Reads having the identical cell labels and UMI sequences and aligned to the same gene were collapsed into a single molecule. To correct sequencing and PCR errors, the obtained counts were adjusted by a recursive substitution error correction (RSEC) algorithm developed by BD Biosciences. The sampling time point and sorting gate of origin of each cell were labelled using information from barcoded oligo-conjugated antibodies. Cells without barcoding information were labelled as “tag N/A”, and cells with two or more barcodes were labelled as “multiplet”. We note that sample barcoding and AbSeq staining efficiency was compromised in $CD56^{br}$ NK cells. As the quality of the main mRNA library was not affected by this technical issue, $CD56^{br}$ NK cells without barcoding information were included in the quality control and clustering step, but excluded from subsequent analyses dependent on sampling time point (see below). This led to a significant reduction in the number of $CD56^{br}$ cells analysed and a concomitant reduction in statistical power to identify IL-2-induced alterations. Nevertheless, we also note that our cell-enrichment strategy compensated for this cell loss due to technical reasons and provided sufficient resolution to dissect the heterogeneity of the rare population of $CD56^{+}$ NK cells.

Doublet scoring

Doublet scoring was performed on each sample separately using Scrublet[53]. Briefly, this method simulates synthetic doublets by sampling pairs of cells and then scores each cell by its similarity to the simulated doublets. For each sample, count matrices for RNA and AbSeq data were concatenated and used as the input. The method was modified to exclude known multiplet-tagged cells from the doublet simulation while retaining them for scoring. The doublet score was used to help identify clusters of doublets in subsequent quality control stages.

Normalisation, integration, dimensionality reduction and clustering

Data normalisation was performed using Seurat[54]. The RNA count matrices were normalised by log normalisation. Specifically, the RNA counts for each cell were (1) divided by the total counts for that cell, (2) multiplied by a scale factor of 10,000, and (3) natural-log transformed after adding one pseudo count. The AbSeq counts for each cell were considered compositional and normalised by a centred log ratio transformation[55]. Normalised RNA and AbSeq expression matrices were then centred and scaled. Specifically, each feature was linearly regressed against selected latent variables, and the resulting residuals were mean-centred and divided by their standard deviations.

Integration of the scaled AbSeq expression matrices from different participants were performed using a combination of canonical correlation analysis (CCA) and identification of mutual nearest neighbours (MNNs), implemented in Seurat[54]. After integration was performed, the updated AbSeq expression matrices produced by the integration algorithm were used in subsequent dimensionality reduction and clustering.

Dimensionality reduction and clustering were performed using Seurat[54]. Specifically, principal components (PCs) were calculated separately for each assay using the top variable features of the scaled or integrated expression matrices, and a weighted nearest neighbour (WNN) graph as well as a weighted shared nearest neighbour (WSNN) graph were generated on the ten top PCs from each assay, unless otherwise specified. We used $k = 30$ nearest neighbours for each cell. Uniform Manifold Approximation and Projection (UMAP)[56] embeddings were then calculated based on

the WNN graph. The Louvain algorithm[57] was applied to identify clusters of cells based on a WSNN graph, at the selected resolution level. Clusters with fewer than ten cells were removed. The resulting clusters were manually annotated based on marker genes and sorting gate information to elucidate cell types and label suspicious clusters.

Data quality control

The BD Rhapsody single-cell sequencing data generated in this study was processed through a pipeline consisting of multiple quality control steps as described below. Methodological details of each step are specified in previous sections. The single-cell sequencing data of each sample (i.e. the collection of stimulated and unstimulated cells for a single participant) were pre-processed to obtain two annotated count matrices (RNA and AbSeq). We performed data normalisation, dimensionality reduction and clustering for each sample. Specifically, 30 RNA principal components (PCs) and 30 AbSeq PCs were used to generate the WNN and WSNN graphs, and clustering was performed at resolution 0.5.

Within each sample, cells meeting one of the following criteria were removed: (1) cells that are labelled as “multiplet” or in clusters with high doublet scores; (2) cells labelled as “tag N/A”, except those in clusters annotated as $CD56^{br}$; (3) cells in clusters with very low RNA counts and very low or very high AbSeq counts; (4) cells in clusters annotated as monocyte or B cell contamination; (5) cells labelled as DN T cells.

The filtered cells from each participant were merged according to the stimulation status into two datasets (unstimulated cells and *in vitro* stimulated cells). Both stimulated and unstimulated cells were then processed in the same way unless otherwise specified. We performed data normalisation, dimensionality reduction and clustering for each merged dataset. Specifically, clustering was performed at resolution 0.7 (unstimulated cells) or 1.5 (stimulated cells). Doublet clusters were manually annotated and removed. Data normalisation, dimensionality reduction and clustering were performed again. Specifically, clustering was performed at resolution 0.5.

Cells in each filtered dataset were subsequently partitioned into four major groups ($CD4^+$ Tregs, $CD4^+$ Teffs, CD8+ T cells and $CD56^+$ NK cells) based on cluster annotations and sorting gates

of origin. In addition, unstimulated cells in clusters marked by genes related to cell cycle were grouped as cycling cells. Doublet clusters were manually annotated and removed. Data normalisation, dimensionality reduction, integration and clustering were performed. Specifically, clustering was performed at resolution 0.5. Cells from the CD8 gate were removed from the unstimulated CD4 groups, and cells from non-CD8 gates were removed from the unstimulated CD8 group. Finally, Data normalisation, dimensionality reduction, and clustering were performed again.

Due to the aforementioned technical limitations with AbSeq staining in $CD56^{br}$ NK cells, AbSeq data was not used in dimensionality reduction and clustering of NK cells. The following resolutions were used for the final clustering: unstimulated Tregs, 0.6; unstimulated Teffs, 0.5; unstimulated CD8, 0.5; unstimulated NK, 0.3; unstimulated cycling cells, 0.5; stimulated Tregs, 0.5; stimulated Teffs, 0.4; stimulated CD8: 0.5; stimulated NK: 0.5. The resulting partitioned datasets were used for subsequent analyses. We note that one PBMC aliquot (corresponding to the Day 55 visit of Participant 8) yielded very low cell numbers (N = 603 unstimulated and N = 67 stimulated cells) and displayed a compromised FACS staining profile. These cells were included in quality control and clustering procedures but excluded from downstream analyses. All other 38 samples corresponding to 38 visits yielded the expected cell frequencies and numbers.

Single-cell proliferation scores

Single-cell proliferation scores were calculated using a previously described approach[58], which is based on the average normalised expression levels of a pre-selected set of 11 mRNA features related to cell cycle (*AURKB*, *HMGB2*, *HMMR*, *MCM4*, *MKI67*, *PCLAF*, *PCNA*, *TK1*, *TOP2A*, *TYMS*, and *UBE2C*), subtracted by the aggregated expression of a control set of 50 randomly selected mRNA features.

The proliferation scores were used to identify cycling clusters from the 15 functional T and NK cell subsets identified in blood from the initial clustering (**Fig. 1g**). We identified two clusters characterised by the distinct expression of cell-cycle genes, corresponding to clusters of cycling

T and NK cells, respectively (**Fig. 1h**). We note that the distinct co-expression of these cell-cycle genes present in our targeted transcriptional panel superseded more subtle functional differences and aggregated all cycling T and NK cells into two respective clusters, regardless of their original sorting gate. Given this lack of functional differentiation and to avoid the potential overwhelming effect of these cell-cycle genes on the resulting clustering visualisation, we opted to separate these cycling T and NK cells from the rest of the cells included in subsequent clustering steps to assess IL-2-induced changes in their relative frequency in blood.

Cluster annotation

Functional annotation of the resulting clusters was performed manually based on the cluster-specific differential expression of the mRNA and protein markers (summarised in **Supplementary Data 1**). Sample tagging information, corresponding to sorting gate information, was used to sub-cluster the 5 main immune populations (ie. $CD4^{+}$ $CD127^{low}CD25^{hi}$ Tregs; $CD4^{+}$ $CD25^{-/low}$ Teffs; $CD8^{+}$ T cells; $CD56^{br}$ NK cells and $CD56^{dim}$ NK cells). The use of a bespoke targeted single-cell approach combined with a cell-subset enrichment strategy provided a high-resolution map of the heterogeneity of these cell populations – and particularly of the rare circulating $CD4^{+}$ Treg and $CD56^{br}$ NK cell subsets.

We note that a consequence of the cell-subset enrichment strategy used in this study is that the frequency of the T and NK cell subsets described in this dataset do not directly represent their frequency in whole blood. The aim of this study was not to identify changes in the global composition of PBMCs, but to specifically focus on the T and NK cell subsets. The pre-purification of the Treg and $CD56^{br}$ NK subsets was central to the resolution of this approach and provided greater power to dissect their functional heterogeneity and investigate how their relative composition was modified by LD-IL-2 immunotherapy

Functional annotation of $CD4^{+}$ $CD127^{low}CD127^{hi}$ T cell clusters

Owing to our cell-subset enrichment strategy, we obtained a high-resolution map depicting the heterogeneity of CD127^{low}CD25^{hi} T cells in blood, with 14 distinct clusters identified in unstimulated cells and 12 in *in vitro* stimulated cells (**Extended Data Fig. 3**). Overall, we found consistency in the functional annotation of the identified subsets in stimulated and unstimulated cells, with a clear demarcation of the naïve and memory subsets.

Expression of the canonical Treg transcription factors *FOXP3* and *IKZF2* (encoding HELIOS) was used to identify three distinct groups of clusters, corresponding to different populations of CD127^{low}CD25^{hi} T cells: (i) CD45RA^{+} FOXP3^{+}HELIOS^{+} clusters (depicted in green) corresponding to thymic-derived naïve Tregs; (ii) CD45RA^{-} FOXP3^{+}HELIOS^{+} clusters (depicted in blue) corresponding to memory Tregs with suppressive function; and (iii) CD25^{+} FOXP3^{-}HELIOS^{-} clusters (depicted in red) corresponding to a small subset of memory Teffs expressing CD25 likely due to recent activation (**Extended Data Fig. 4**). Additional clusters showing variable expression of FOXP3 but lack of HELIOS expression (FOXP3^{+}HELIOS^{-} clusters) were depicted in black, and represent a more heterogeneous group of Treg subsets, containing a mix of thymic-derived and induced Treg subsets.

As expected, most of the cells within the CD127^{low}CD25^{hi} Treg gate showed concomitant expression of FOXP3 and HELIOS, which is consistent with their largely demethylated *FOXP3* Treg-specific demethylated region (TSDR)[59,60], the hallmark of *bona fide* thymic Tregs. Furthermore, we were able to identify a trajectory of thymic Treg differentiation, originating from the naïve FOXP3^{+}HELIOS^{+} Tregs (cluster 0). Memory FOXP3^{+}HELIOS^{+} Tregs displayed more heterogeneity and included four distinct subsets in unstimulated cells, including a cluster of CD80^{+} Tregs characterised by elevated expression of tissue-homing receptors (cluster 7) and a cluster of effector Tregs (cluster 1) typically marked by the expression of HLA class II and other effector Treg markers such as CD39 and GITR (**Extended Data Fig. 3**). We have also identified a subset of CD45RA^{+} naïve cells (cluster 2) showing much lower expression of FOXP3 and HELIOS compared to the classical naïve Treg cluster. This likely corresponds to a subset of homeostatically-expanded CD45RA^{+} T cells upregulating CD25 that we have previously characterised[61]. Consistent with this non-Treg profile, these CD25^{+} naïve T cells upregulated the expression of IL-2 upon *in vitro* stimulation (**Extended Data Fig. 4f**). The expression of IL-2 in stimulated cells was also a hallmark of the CD25^{+} FOXP3^{-}HELIOS^{-} Teff clusters and allowed

us to further refine their functional heterogeneity. Moreover, the relative frequency of CD45RA^{-} FOXP3^{-}HELIOS^{-} cells identified in this single-cell dataset is consistent with the frequency of cells identified by FACS as cytokine-producing FOXP3^{-}HELIOS^{-} cells in the CD127^{low}CD25^{hi} Treg gate that we have previously shown to contain a methylated TSDR[60], further indicating that they correspond to a small proportion of activated CD25^{+} Teffs captured in the CD127^{low}CD25^{hi} Treg gate.

Functional annotation of CD56^{+} NK cell clusters

In contrast to T cells, we found much lower heterogeneity in NK cells (**Extended Data Fig. 8**). In addition, the technical issue affecting the efficiency of the AbSeq staining on CD56^{br} NK cells, precluded the use of protein data for the clustering and annotation of the CD56^{+} NK cells, which also contributed to a lower resolution of the clusters. Nevertheless, our cell-enrichment strategy allowed us to generate a comprehensive single-cell map of NK cell heterogeneity, most notably of the rare, IL-2-sensitive CD56^{br} NK population. We observed differentiation of CD56^{br} NK cells from a rare progenitor subset marked by the expression of genes associated with haematopoiesis such as SOX4 and KIT, and characterization of two functional subsets associated with a trajectory of CD56^{br} NK cell maturation, which was marked by HLA class II gene expression in the more differentiated cells. In contrast, CD56^{dim} NK cells showed a distinct transcriptional profile of two distinct subsets, reflecting a gradient of NK cell activation associated with the acquisition of a cytotoxic profile (**Fig. 4d** and **Extended Data Fig. 8**). These findings support a distinct functional differentiation between the CD56^{dim} and CD56^{br} subsets, which allowed clear identification a fraction of CD56^{dim} NK cells sorted in the CD56^{br} gate, as well as a fraction of CD56^{br} NK cells sorted within the CD56^{dim} gate (**Fig. 4e**)

Differential abundance analysis

To identify specific cell frequency changes across different sampling time points, we calculated the frequency of each cell type annotated. A cell type was defined as a group of cells from the same cluster or several similar clusters, after removing cells not sorted from the corresponding

gate. The frequency of each cell type for each participant and each time point was calculated as the proportion of that cell type within all cells from the same participant, time point and sorting gate of origin. For each cell type, a Wilcoxon signed-rank test was performed to compare the frequencies at Day 0 with the frequencies at Day 27 or Day 55 in the 13 participants selected for single-cell analysis. Benjamini-Hochberg correction was applied after pooling all resulting *P* values. For each cell type and each participant, the $\log_2$ fold change values between corresponding frequency values were calculated and visualised. We note that cells were compared only if they were from the same sorting gate, as cells from different sorting gates were pooled in designated proportions that are not biologically relevant.

Generation and normalisation of pseudo-bulk expression data

Unstimulated and stimulated cells were separately partitioned into five groups (Tregs, Teffs, CD8, $CD56^{br}$ and $CD56^{dim}$), after removing cells not sorted from the corresponding gate. In each group, cells from the same participant, time point and stimulation status were aggregated into one pseudo-bulk sample by summarising raw counts of each mRNA or AbSeq feature. The resulting pseudo-bulk expression matrices were normalised by dividing raw counts with sample-specific scale factors calculated using the median-of-ratios method previously described[62].

Differential expression analysis

Differential expression analysis was performed separately for RNA and AbSeq data from raw pseudo-bulk counts using DESeq2[63]. Specifically, the likelihood ratio test was used, with a full model including time points and the participants as independent variables, and a reduced model including only participants as the independent variable. The apeglm method[64] was applied to shrink the resulting fold change values, and the Benjamini-Hochberg FDR correction was applied after pooling all resulting *P* values (**Supplementary Data 2**). Genes with either (i) absolute $\log_2$ fold change values equal to or greater than 0.4 and FDR-adjusted *P* values less than 0.01, or (ii) FDR-adjusted *P* values less than 1×10^{-10}, were selected as significantly differentially

expressed genes, unless otherwise noted. Participant-specific $\log_2$ fold change values were calculated directly from the normalised pseudo-bulk expression data.

Day 55 signature scores

Forty-one genes that are significantly upregulated or downregulated in at least one unstimulated cell population comparing Day 55 with Day 0 were defined as Day 55 signature genes. Day 55 signature scores were calculated from the normalised pseudo-bulk expression data of Day 55 signature genes. For each pseudo-bulk sample, the Day 55 signature score was defined as the average normalised expression level of upregulated signature genes subtracted by that of downregulated signature genes. For each participant and cell type, the Δ signature scores were defined as the signature score differences between Day 27 or Day 55 and Day 0, as indicated.

Validation of Day 55 signature in the COMBAT dataset

To confirm the differential expression of Day 55 signature genes after COVID-19 infection, we applied DESeq2 to pseudo-bulk RNAseq data from the COMBAT dataset, comparing 13 community COVID-19 patients with 10 healthy controls. Only the Day 55 signature genes were included. Each assessed cell type was analysed independently, using the likelihood ratio test with participant group as the independent variable, followed by ageglm shrinkage of fold change values. Samples with fewer than 10 pseudo-bulk counts per gene on average were removed.

Sparse Decomposition of Arrays (SDA)

The SDA analysis was performed based on pseudo-bulk data as previously described[31], with slight modifications. For this analysis, each biological sample for a single participant, from a single time point, was considered an SDA sample, and unstimulated or stimulated condition for each major cell type was considered an SDA context. Accordingly, we constructed two three-dimensional tensors for RNA and AbSeq data, each containing 38 samples and 10 contexts, from

the raw pseudo-bulk count matrix. The following pre-processing and normalisation steps were then performed: (i) removal of 28 RNA and 0 AbSeq features that had zero counts in over 20% of samples in all 10 contexts, (ii) normalisation within each sample and context by the total count, (iii) quantile normalisation across samples within each context, and (iv) rank-based transformation of each feature onto the standard normal distribution. The resulting normalised RNA and AbSeq tensors were decomposed into 60 components. Components identified from 20 independent runs were pooled and hierarchically clustered using 1 - absolute Pearson correlation coefficient as the distance metric. Clustering was terminated after the max distance between any two components from different clusters reached 0.6. Clusters spanning less than 50% of runs were removed. Components from each of the remaining clusters were combined by taking the mean of the sample, context and gene loading scores and median posterior inclusion probabilities (**Supplementary Data 3**).

Acknowledgements

We thank all participants in DILfrequency for their generous contribution to this study. The authors also acknowledge the support of the NIHR Cambridge Biomedical Research Centre and the Cambridge Clinical Trial Unit for trial coordination; the NIHR/Wellcome Trust Clinical Research Facility and Addenbrooke's Centre for Clinical Investigation for clinical facilities; the Department of Clinical Immunology, Addenbrooke's Hospital; the JDRF/Wellcome Diabetes and Inflammation laboratory sample processing team (led by Helen Stevens) and information technology and administration team (led by Judy Brown) and data teams at the Cambridge Institute for Medical Research. We thank Shannah Donhou, Sarune Kacinskaite and Heather McMurray, University of Oxford for sample collection and preparation, and members of the Diabetes and Inflammation Laboratory for critical discussion. We also thank Justin Whalley, Georgina Kerr, Brian Marsden and Julian Knight, University of Oxford, for providing access to clinical data collected from the COMBAT cohort.

Funding

This work was supported by the Sir Jules Thorn Trust (13/JTA (OCT2013/DR/1044)), the JDRF (1-SRA-2019-657-A-N), and the NIHR Cambridge Biomedical Research Centre. The Diabetes and Inflammation Laboratory was supported by a strategic award from the Wellcome (107212/A/15/Z) and the JDRF (4-SRA-2017-473-A-A). JYZ was supported by the China Scholarship Council-University of Oxford Scholarship. The research was supported by the Wellcome Trust Core Award Grant Number 203141/Z/16/Z with funding from the NIHR Oxford BRC. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Competing interests

FWL is employed by Vertex Pharmaceuticals; JAT is a member of the GSK Human Genetics Advisory Board.

References

1. Koreth, J. *et al.* Interleukin-2 and Regulatory T Cells in Graft-versus-Host Disease. *New England Journal of Medicine* **365**, 2055–2066 (2011).
2. Saadoun, D. *et al.* Regulatory T-Cell Responses to Low-Dose Interleukin-2 in HCV-Induced Vasculitis. *New England Journal of Medicine* **365**, 2067–2077 (2011).
3. Castela, E. *et al.* Effects of Low-Dose Recombinant Interleukin 2 to Promote T-Regulatory Cells in Alopecia Areata. *JAMA Dermatology* **150**, 748–751 (2014).
4. von Spee-Mayer, C. *et al.* Low-dose interleukin-2 selectively corrects regulatory T cell defects in patients with systemic lupus erythematosus. *Annals of the Rheumatic Diseases* **75**, 1407–1415 (2016).
5. Zhao, T. X. *et al.* Low-dose interleukin-2 in patients with stable ischaemic heart disease and acute coronary syndromes (LILACS): protocol and study rationale for a randomised, double-blind, placebo-controlled, phase I/II clinical trial. *BMJ Open* **8**, e022452 (2018).
6. He, J. *et al.* Efficacy and safety of low-dose IL-2 in the treatment of systemic lupus erythematosus: a randomised, double-blind, placebo-controlled trial. *Annals of the Rheumatic Diseases* **79**, 141–149 (2020).
7. Klatzmann, D. & Abbas, A. K. The promise of low-dose interleukin-2 therapy for autoimmune and inflammatory diseases. *Nature Reviews Immunology* **15**, 283–294 (2015).
8. Graßhoff, H. *et al. Low-Dose IL-2 Therapy in Autoimmune and Rheumatic Diseases*. *Frontiers in Immunology* **12**, e648408 (2021).
9. Kolios, A. G. A., Tsokos, G. C. & Klatzmann, D. Interleukin-2 and regulatory T cells in rheumatic diseases. *Nature Reviews Rheumatology* **17**, 749–766 (2021).
10. Todd, J. A. Etiology of Type 1 Diabetes. *Immunity* **32**, 457–467 (2010).
11. Hulme, M. A., Wasserfall, C. H., Atkinson, M. A. & Brusko, T. M. Central Role for Interleukin-2 in Type 1 Diabetes. *Diabetes* **61**, 14–22 (2011).
12. Long, S. A. *et al.* Rapamycin/IL-2 Combination Therapy in Patients With Type 1 Diabetes Augments Tregs yet Transiently Impairs β-Cell Function. *Diabetes* **61**, 2340–2348 (2012).
13. Hartemann, A. *et al.* Low-dose interleukin 2 in patients with type 1 diabetes: a phase 1/2 randomised, double-blind, placebo-controlled trial. *The Lancet Diabetes & Endocrinology* **1**, 295–305 (2013).
14. Rosenzwajg, M. *et al.* Low-dose interleukin-2 fosters a dose-dependent regulatory T cell tuned milieu in T1D patients. *Journal of Autoimmunity* **58**, 48–58 (2015).
15. Rosenzwajg, M. *et al.* Low-dose IL-2 in children with recently diagnosed type 1 diabetes: a Phase I/II randomised, double-blind, placebo-controlled, dose-finding study. *Diabetologia* **63**, 1808–1821 (2020).
16. Todd, J. A. J. *et al.* Regulatory T cell responses in participants with type 1 diabetes after a single dose of interleukin-2: a non-randomised, open label, adaptive dose-finding trial. *PLoS Med* **13**, e1002139 (2016).
17. Seelig, E. *et al.* The DILfrequency study is an adaptive trial to identify optimal IL-2 dosing in patients with type 1 diabetes. *JCI Insight* **3**, e99306 (2018).
18. Trzupek, D. *et al.* Single-cell multi-omics analysis reveals IFN-driven alterations in T lymphocytes and natural killer cells in systemic lupus erythematosus. *Wellcome Open Research* **6**, (2021).

19. Kenefeck, R. *et al.* Follicular helper T cell signature in type 1 diabetes. *The Journal of Clinical Investigation* **125**, 292–303 (2015).
20. Ferreira, R. C. *et al.* IL-21 production by CD4+ effector T cells and frequency of circulating follicular helper T cells are increased in type 1 diabetes patients. *Diabetologia* **58**, 781–790 (2015).
21. Ekman, I. *et al.* Circulating CXCR5−PD-1hi peripheral T helper cells are associated with progression to type 1 diabetes. *Diabetologia* **62**, 1681–1688 (2019).
22. Bell, C. J. M. *et al.* Sustained in vivo signaling by long-lived IL-2 induces prolonged increases of regulatory T cells. *Journal of Autoimmunity* **56**, 66–80 (2015).
23. Yu, A. *et al.* Selective IL-2 Responsiveness of Regulatory T Cells Through Multiple Intrinsic Mechanisms Supports the Use of Low-Dose IL-2 Therapy in Type 1 Diabetes. *Diabetes* **64**, 2172–2183 (2015).
24. Trzupek, D. *et al.* Discovery of CD80 and CD86 as recent activation markers on regulatory T cells by protein-RNA single-cell analysis. *Genome Medicine* **12**, 55 (2020).
25. Dong, S. *et al.* The effect of low-dose IL-2 and Treg adoptive cell therapy in patients with type 1 diabetes. *JCI Insight* **6**, e147474 (2021).
26. Singh, S. S. *et al.* Amphiregulin in cellular physiology, health, and disease: Potential use as a biomarker and therapeutic target. *Journal of Cellular Physiology* **237**, 1143–1156 (2022).
27. West, N. R. *et al.* Oncostatin M drives intestinal inflammation and predicts response to tumor necrosis factor–neutralizing therapy in patients with inflammatory bowel disease. *Nature Medicine* **23**, 579–589 (2017).
28. Richard, A. C. *et al.* Targeted genomic analysis reveals widespread autoimmune disease association with regulatory variants in the TNF superfamily cytokine signalling network. *Genome Medicine* **8**, 76 (2016).
29. Maña, P. *et al.* LIGHT (TNFSF14/CD258) Is a Decisive Factor for Recovery from Experimental Autoimmune Encephalomyelitis. *The Journal of Immunology* **191**, 154–163 (2013).
30. Ahern, D. J. *et al.* A blood atlas of COVID-19 defines hallmarks of disease severity and specificity. *Cell* 185, 916-938 (2022).
31. Hore, V. *et al.* Tensor decomposition for multiple-tissue gene expression experiments. *Nature Genetics* **48**, 1094–1100 (2016).
32. Ito, S. *et al.* Ultra-low Dose Interleukin-2 Promotes Immune-modulating Function of Regulatory T Cells and Natural Killer Cells in Healthy Volunteers. *Molecular Therapy* **22**, 1388–1395 (2014).
33. Whangbo, J. S. *et al.* Functional analysis of clinical response to low-dose IL-2 in patients with refractory chronic graft-versus-host disease. *Blood Advances* **3**, 984–994 (2019).
34. Whangbo, J. S. *et al.* Dose-escalated interleukin-2 therapy for refractory chronic graft-versus-host disease in adults and children. *Blood Advances* **3**, 2550–2561 (2019).
35. Bielekova, B. *et al.* Regulatory CD56bright natural killer cells mediate immunomodulatory effects of IL-2Rα-targeted therapy (daclizumab) in multiple sclerosis. *Proceedings of the National Academy of Sciences* **103**, 5941–5946 (2006).
36. Melsen, J. E., Lugthart, G., Lankester, A. C. & Schilham, M. W. Human Circulating and Tissue-Resident CD56bright Natural Killer Cell Populations. *Frontiers in Immunology* **7**, 262 (2016).

37. Ballesteros-Tato, A. *et al.* Interleukin-2 Inhibits Germinal Center Formation by Limiting T Follicular Helper Cell Differentiation. *Immunity* **36**, 847–856 (2012).
38. Yamanouchi, J. *et al.* Interleukin-2 gene variation impairs regulatory T cell function and causes autoimmunity. *Nat Genet* **39**, 329–337 (2007).
39. Liu, S. M. *et al.* Differential IL-21 signaling in APCs leads to disparate Th17 differentiation in diabetes-susceptible NOD and diabetes-resistant NOD.Idd3 mice. *The Journal of Clinical Investigation* **121**, 4303–4310 (2011).
40. Crotty, S. T Follicular Helper Cell Biology: A Decade of Discovery and Diseases. *Immunity* **50**, 1132–1148 (2019).
41. Ren, H. M., Lukacher, A. E., Rahman, Z. S. M. & Olsen, N. J. New developments implicating IL-21 in autoimmune disease. *Journal of Autoimmunity* **122**, 102689 (2021).
42. von Herrath, M. *et al.* Anti-interleukin-21 antibody and liraglutide for the preservation of β-cell function in adults with recent-onset type 1 diabetes: a randomised, double-blind, placebo-controlled, phase 2 trial. *The Lancet Diabetes & Endocrinology* **9**, 212–224 (2021).
43. Sage, P. T. & Sharpe, A. H. T follicular regulatory cells. *Immunological Reviews* **271**, 246–259 (2016).
44. Jacobsen, J. *et al.* Expression of Foxp3 by T follicular helper cells in end-stage germinal centers. *Science* **373**, eabe5146 (2021).
45. Matsumoto, A. *et al.* CIS, a Cytokine Inducible SH2 Protein, Is a Target of the JAK-STAT5 Pathway and Modulates STAT5 Activation. *Blood* **89**, 3148–3154 (1997).
46. Shoger, K. E. *et al.* CISH attenuates homeostatic cytokine signaling to promote lung-specific macrophage programming and function. *Science Signaling* **14**, eabe5137 (2021).
47. Franitza, S. *et al.* TNF-α Associated with Extracellular Matrix Fibronectin Provides a Stop Signal for Chemotactically Migrating T Cells. *The Journal of Immunology* **165**, 2738–2747 (2000).
48. Bohnacker, S. *et al.* Mild COVID-19 imprints a long-term inflammatory eicosanoid- and chemokine memory in monocyte-derived macrophages. *Mucosal Immunology* (2022) doi:10.1038/s41385-021-00482-8.
49. Phetsouphanh, C. *et al.* Immunological dysfunction persists for 8 months following initial mild-to-moderate SARS-CoV-2 infection. *Nature Immunology* **23**, 210–216 (2022).
50. Sims, E. K. *et al.* Teplizumab improves and stabilizes beta cell function in antibody-positive high-risk individuals. *Science Translational Medicine* **13**, eabc8980 (2021).
51. Quattrin, T. *et al.* Golimumab and Beta-Cell Function in Youth with New-Onset Type 1 Diabetes. *New England Journal of Medicine* **383**, 2007–2017 (2020).
52. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).
53. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Systems* **8**, 281-291.e9 (2019).
54. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29 (2021).
55. Aitchison, J. *The Statistical Analysis of Compositional Data*. (Chapman & Hall, Ltd., 1986).
56. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv* **arXiv:1802** (2018).

57. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* **2008**, P10008 (2008).
58. Itay, T. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).
59. Rainbow, D. B. *et al.* Epigenetic analysis of regulatory T cells using multiplex bisulfite sequencing. *European Journal of Immunology* **45**, 3200–3203 (2015).
60. Ferreira, R. C. *et al.* Cells with Treg-specific FOXP3 demethylation but low CD25 are prevalent in autoimmunity. *Journal of Autoimmunity* **84**, 75–86 (2017).
61. Pekalski, M. L. *et al.* Postthymic expansion in human CD4 naive T cells defined by expression of functional high-affinity IL-2 receptors. *The Journal of Immunology* **190**, 2554–2566 (2013).
62. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biology* **11**, R106 (2010).
63. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 1–21 (2014).
64. Zhu, A., Ibrahim, J. G. & Love, M. I. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics* **35**, 2084–2092 (2019).

Figures

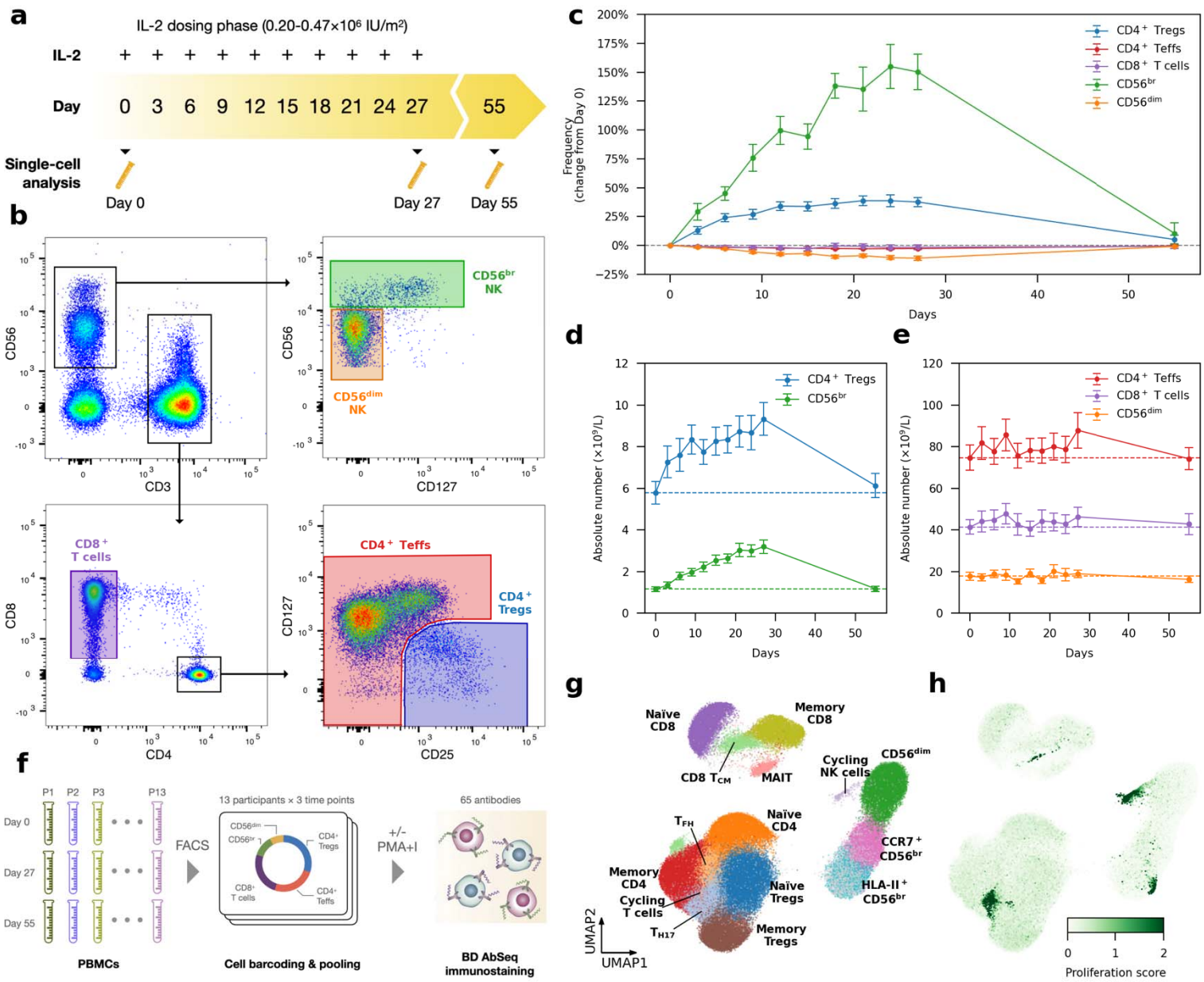


Figure 1. Interval low-dose IL-2 administered every three days sustains an increased frequency of CD4^{+} regulatory T cells and CD56^{br} NK cells in blood.

a, Schematic representation of the interval low-dose IL-2 (iLD-IL-2) dosing regimen used in this study. **b**, Gating strategy for the delineation of the five T and NK cell populations assessed by flow cytometry. **c**, Variation (depicted as the % change from baseline pre-treatment levels) in the frequency of the five assessed immune subsets in blood (denominator is total T cells for T cell subsets and total CD56^{+} NK cells for NK subsets). Data shown depict the average (± standard error of mean, SEM) variation at each visit in the 18 DILfrequency participants treated with the 3-day interval dosing schedule. **d-e**, Absolute

numbers of the assessed $CD4^+$ $CD127^{low}CD25^{hi}$ (Tregs) and $CD56^{br}$ NK cell (**d**), as well as the broader $CD4^+$and $CD8^+$ T and $CD56^+$ NK cell (**e**) subsets in blood. Data shown depict the average (± SEM) numbers at each visit in the 18 selected participants - treated with either 0.2 (*N* = 4), 0.32 (*N* = 12) or 0.47 (*N* = 2) × 10^6 IU/m^2 IL-2. Dotted lines indicate the average baseline pre-treatment numbers of each immune subset. **f**, Overview of the experimental design employed in this study. Simultaneous quantification of mRNA (*N* = 565 probes) and surface-expressed proteins (*N* = 65 AbSeq antibodies) was performed at the single-cell level using the BD Rhapsody system and employing a cell-enrichment strategy. Equal proportions of flow-sorted cells subsets (30% $CD4^+$ Tregs; 25% $CD4^+$ Teffs and $CD8^+$ T cells; 12% $CD56^{br}$ NK cells and 8% $CD56^{dim}$ NK cells) from 13 selected DILfrequency participants treated with the 3-day dosing interval were profiled at three distinct time points: (i) prior to IL-2 treatment (Day 0), (ii) three days after the penultimate dose of IL-2 (Day 27), and (iii) one month after the last dose of IL-2 (Day 55). **g**, Uniform Manifold Approximation and Projection (UMAP) plot depicting the 15 clusters identified from the clustering of the 482,531 unstimulated cells profiled in this study. Clusters were manually annotated based on the expression of key differentially expressed mRNA and protein markers. Light grey cells correspond to $CD56^{br}$ NK cells without sorting gate information. **h**, UMAP plot depicting the expression of the cell-cycle signature score in each unstimulated single-cell, calculated by summarising the expression levels of 11 pre-selected genes related to cell cycle.

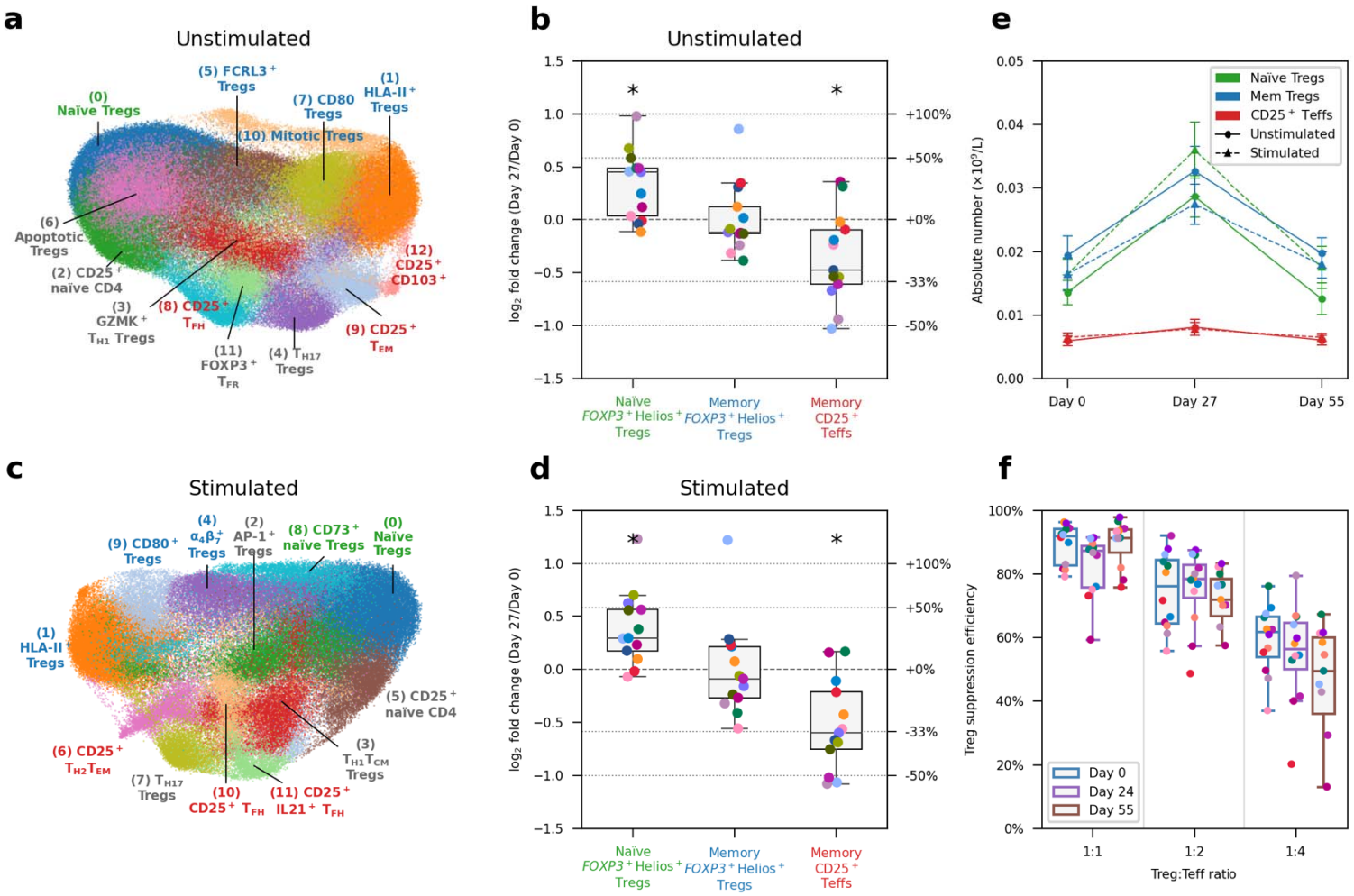


Figure 2. Interval low-dose IL-2 immunotherapy selectively expands FOXP3^{+}HELIOS^{+} Tregs.

a, UMAP plot depicting the 13 clusters identified from the clustering of the 149,202 unstimulated cells sorted from the CD127^{low}CD25^{hi} Treg gate profiled in this study. Clusters were manually annotated based on the expression of key differentially expressed mRNA and protein markers. Clusters corresponding to: (i) naïve Tregs (defined as CD45RA^{+} FOXP3^{+}HELIOS^{+}); (ii) memory Tregs (defined as CD45RA^{-} FOXP3^{+}HELIOS^{+}); or (iii) Memory CD25^{+} Teffs (defined as CD45RA^{-} FOXP3^{-}HELIOS^{-}) are annotated in green, blue and red, respectively. All other miscellaneous clusters not conforming to this definition are annotated in black. **b**, Relative frequency changes of naïve FOXP3^{+}HELIOS^{+} Tregs (green), memory FOXP3^{+}HELIOS^{+} Tregs (blue), and memory CD25^{+} Teff (red) clusters on Day 27 compared to the baseline pre-treatment (Day 0) levels, for unstimulated cells. Relative frequency of a certain cluster was calculated as the number of cells in that cluster divided by the total number of cells sorted from the Treg gate. **c**,**d** Corresponding UMAP plot (**c**) and relative frequency for the identified subsets (**d**) for the 115,153 *in vitro* stimulated CD4^{+} T cells sorted from the CD127^{low}CD25^{hi} Treg gate. **e**, Absolute cell

numbers of naïve FOXP3^{+}HELIOS^{+} Tregs (green), memory FOXP3^{+}HELIOS^{+} Tregs (blue), and memory CD25^{+} Teff (red) on Day 0, Day 27 and Day 35, for unstimulated (circles connected by solid lines) and stimulated (triangles connected by dashed lines) cells. Absolute cell numbers were calculated as the total absolute number of Tregs measured by FACS multiplied by the relative frequency of a certain cluster. **f**, Per-cell Treg suppression efficiency for each time point. Treg suppression efficiency was assessed *in vitro* by culturing CD4^{+}CD25^{hi}CD127^{low} Tregs with autologous CD4^{+}CD25$^{-/lo}$CD127^{+} Teffs at various ratios (x axis) and represented as the decrease in Teff proliferation compared with the control condition (y axis). PBMCs from 12 participants sampled on Day 0, Day 24 and Day 55 were used for this assay. *FDR-adjusted $P < 0.05$; two-sided Wilcoxon signed-rank test comparing the relative frequency of the subsets between Day 27 and Day 0. In **b**, **d** and **f**, each dot represents cells from a single participant. Dots with the same colour represent the same participant.

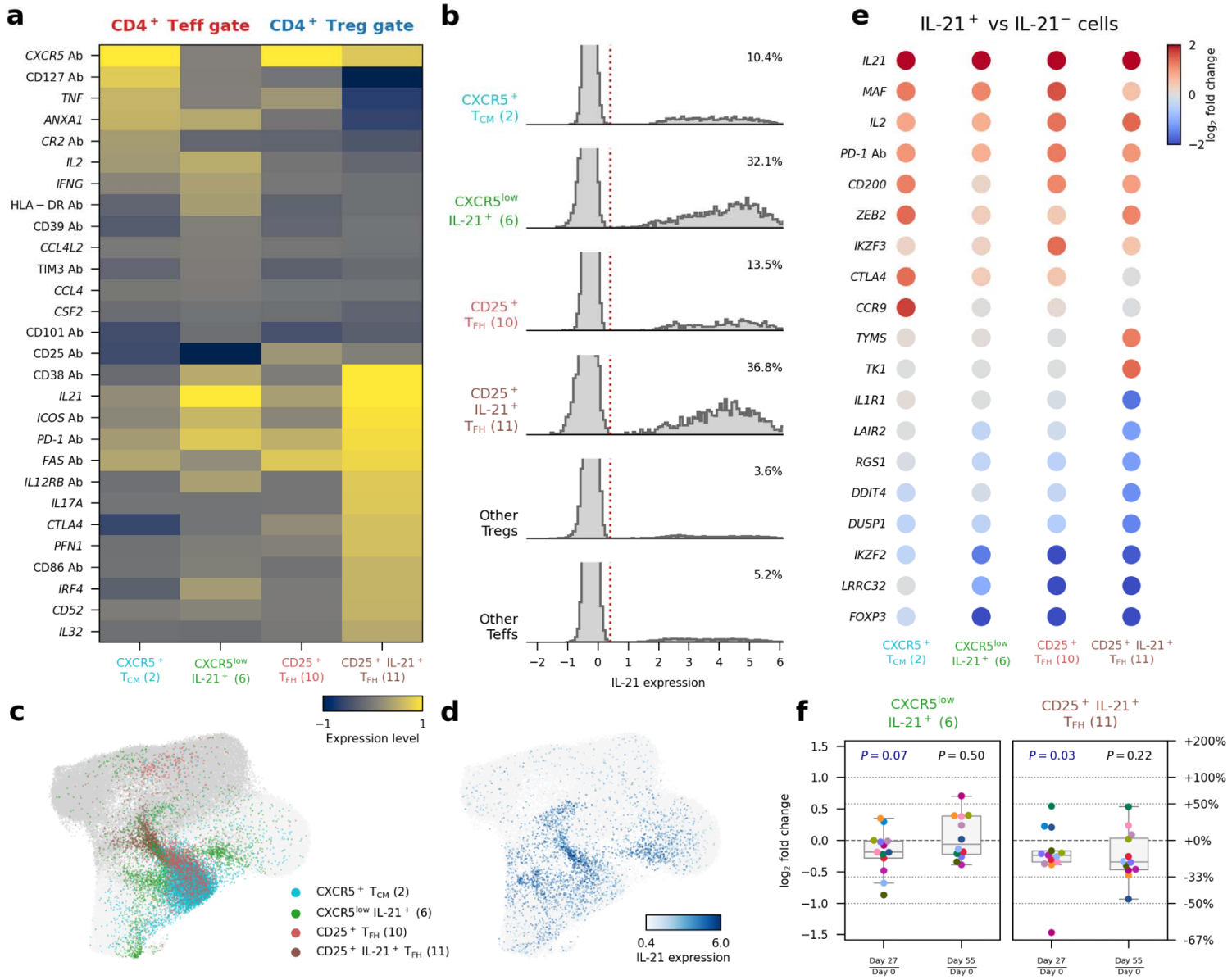


Figure 3. Low-dose IL-2 treatment inhibits the differentiation of IL-21-producing CD4$^+$ T cells

a, Expression levels of 28 differentially expressed markers in the four assessed IL-21-producing CD4$^+$ T cell subsets from *in vitro* stimulated CD4$^+$ T cells: CXCR5$^+$ T_{CM} and CXCR5^{low}IL-21$^+$ sorted from the CD25$^{-/low}$ Teff gate; and CD25$^+$ T_{FH} and CD25$^+$IL-21$^+$ T_{FH} sorted from the CD127^{low}CD25^{hi} Treg gate. Expression levels are based on normalised and scaled single-cell expression data. Gene markers were selected based on pairwise differential expression analysis among the four assessed clusters. **b**, Distribution of IL-21 expression levels in the four assessed IL-21$^+$ CD4$^+$ T cell subsets from *in vitro* stimulated cells. The IL-21 expression levels in the rest of cells from the CD25$^{-/low}$ Teff gate and from the CD127^{low}CD25^{hi} Treg gate are also shown for reference. **c**, UMAP plot depicting the clustering of *in vitro* stimulated CD4$^+$ T cells (N = 209,543 cells). Cells are coloured according to: the respective IL-21$^+$ cluster

assignment. Dark grey and light grey denote cells sorted from the CD127^{low}CD25^{hi} Treg gate and CD25$^{-/low}$ Teff gate, respectively, that are not in IL-21^{+} clusters (**c**); or by normalised and scaled IL-21 expression levels (**d**). **e**, Gene expression changes between IL-21^{+} and IL-21^{-} cells within each of the four assessed IL-21^{+} CD4^{+} T cell clusters. Fold change values were calculated from pseudo-bulk expression data. Genes with absolute $\log_2$ fold change values equal to or greater than 1.2 and FDR-adjusted *P* values less than 0.01 in at least one cluster were selected. **f**, Relative frequency changes of IL-21^{+} cells within the two CD4^{+} T cell subsets displaying higher expression levels of IL-21: the CXCR5^{low}IL-21^{+} subset (left panel) and the CD25^{+}IL-21^{+} T_{FH} subset (right panel), comparing Day 27 and Day 55 with Day 0, respectively. Each dot represents cells from a single participant. Dots with the same colour represent the same participant. *P* values were calculated by comparing the relative frequency values between the two time points using a two-sided Wilcoxon signed-rank test.

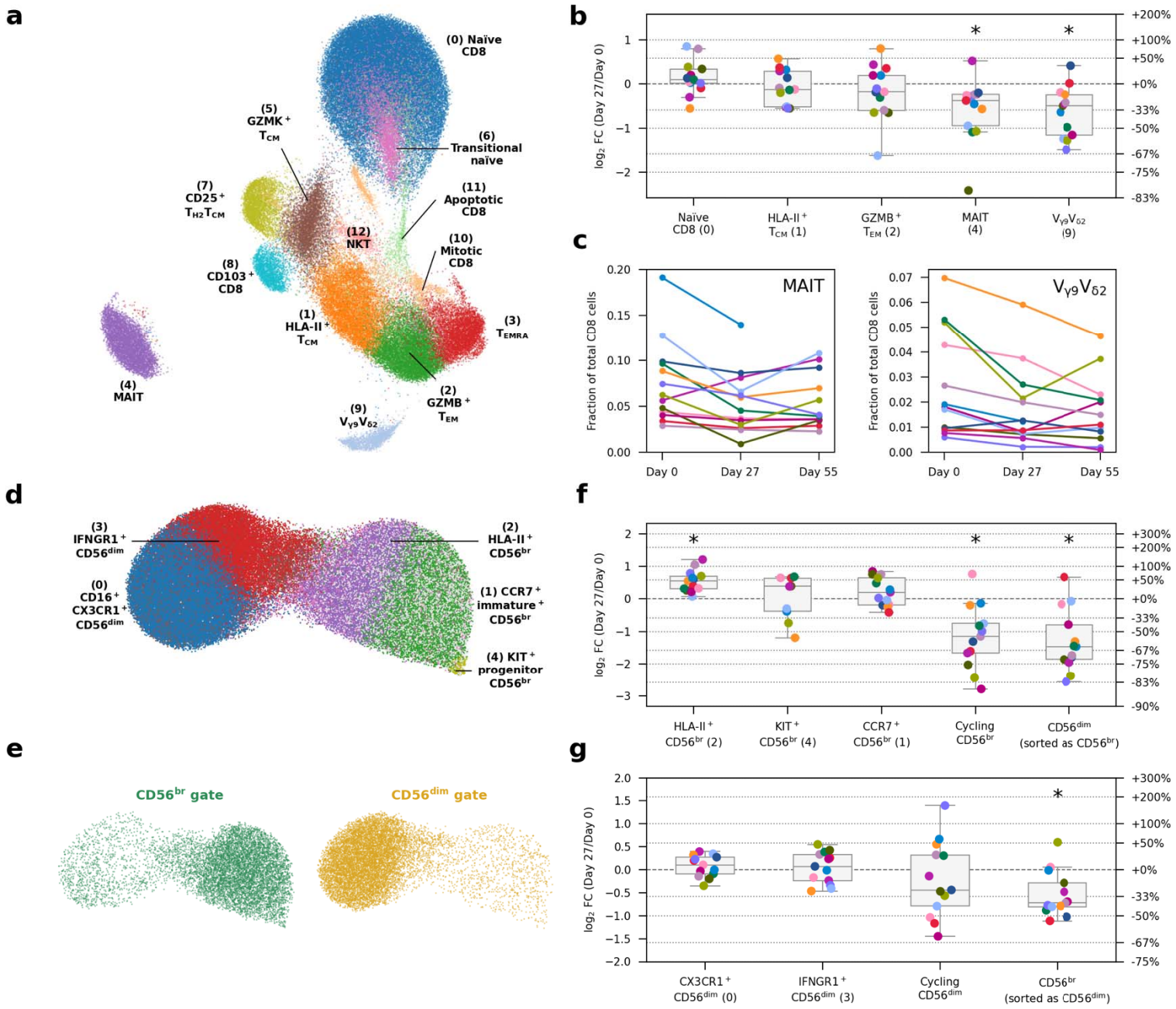


Figure 4. IL-2 treatment reduces the frequency of innate-like CD8$^+$ MAIT and V$_\gamma$9V$_\delta$2 T cells in circulation and selectively expands the IL-2 sensitive CD56^{br} NK cell subset.

a, UMAP plot depicting the clustering of the 93,379 unstimulated CD8$^+$ T cells profiled in this study. Clusters were manually annotated based on the expression of key differentially expressed mRNA and protein markers. **b**, Relative frequency changes of five selected CD8$^+$ T cell clusters on Day 27 compared to the baseline pre-treatment (Day 0) levels, for unstimulated cells. **c**, Relative frequency of the innate-like mucosal-associated invariant T cell (MAIT; left panel) and V$_{\gamma9}$V$_{\delta2}$ T cell (right panel) clusters on Day 0, Day 27, and Day 55. Each line represents cells from a participant. Lines with the same colour

represent the same participant. **d**, UMAP plot depicting the clustering of the 118,625 unstimulated NK cells profiled in this study. Clusters were manually annotated based on the expression of key differentially expressed mRNA and protein markers. **e**, UMAP plot as in **d**, split and coloured by FACS sorting gates of origin. **f-g**, Relative frequency changes of five selected $CD56^{br}$ (**f**) or $CD56^{dim}$ (**g**) NK cell clusters on Day 27 compared to the baseline pre-treatment (Day 0) levels, for unstimulated cells. "$CD56^{dim}$ cells sorted as $CD56^{br}$" represents the fraction of cells sorted from the $CD56^{br}$ NK cell gate that are found in clusters that predominantly consists of $CD56^{dim}$ NK cells (ie. Clusters 0 and 3). Conversely, "$CD56^{br}$ cells sorted as $CD56^{dim}$" represents the fraction of cells sorted from the $CD56^{dim}$ NK cell gate that are found in clusters that predominantly consists of $CD56^{br}$ NK cells (ie. Clusters 1, 2 and 4). Note that "Cycling $CD56^{br}$" and "Cycling $CD56^{dim}$" correspond to the fraction of cells in cell-cycle sorted from the $CD56^{br}$ and $CD56^{dim}$ NK cell gates, respectively (**Fig. 1h**), which are not represented in **d**. In **b**, **f**, and **g**, the relative frequency of a certain cluster is calculated as the number of cells in that cluster divided by the total number of cells sorted from the CD8 (**b**), $CD56^{br}$ (**f**) or $CD56^{dim}$ (**g**) NK cell gate. Each dot represents cells from a single participant. Dots with the same colour represent the same participant. Asterisks denote statistically significant differences (two-sided Wilcoxon signed-rank test) in cell population frequencies between Day 27 and Day 0.

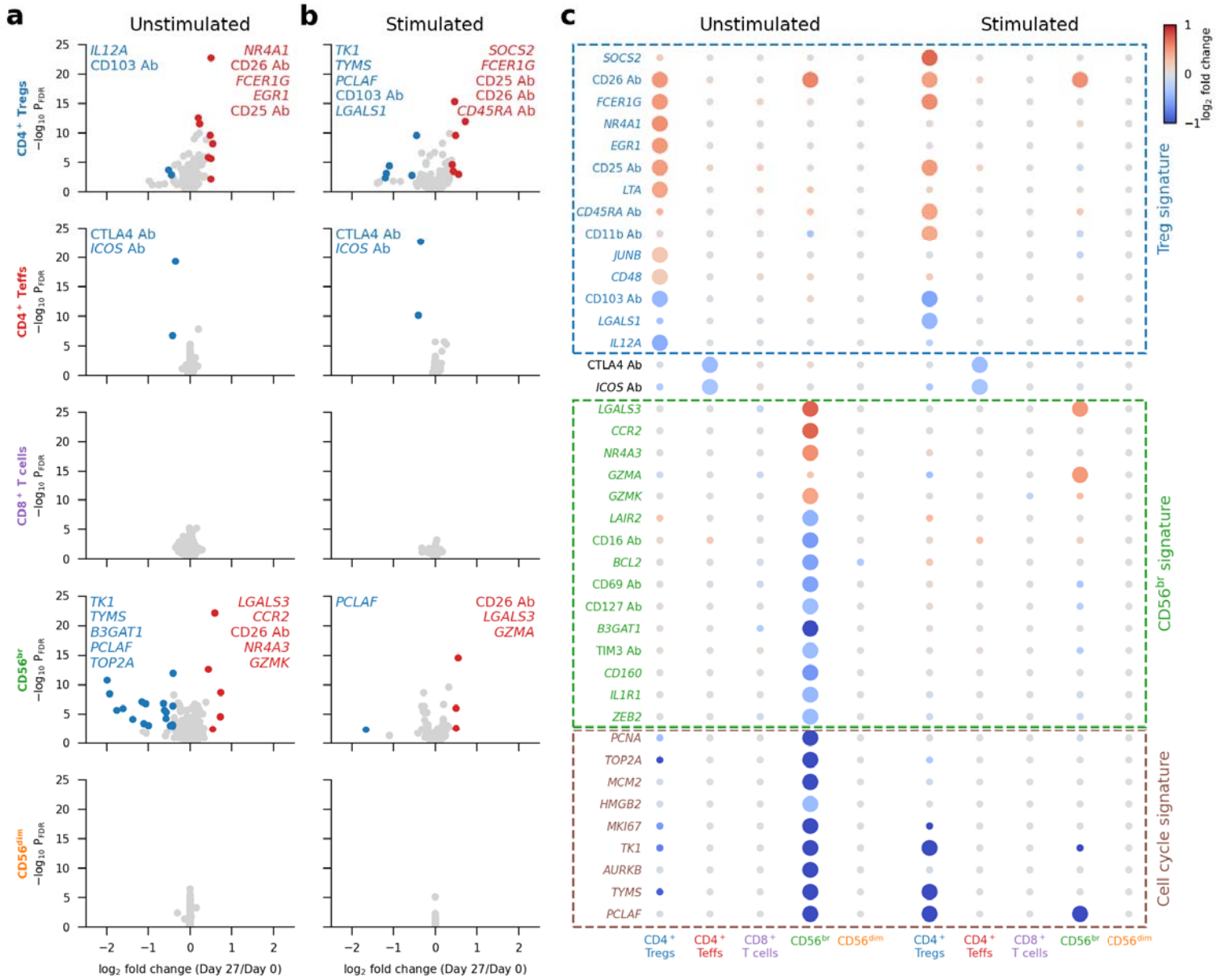


Figure 5. IL-2 treatment selectively modulates gene expression of CD25^{hi} Tregs and CD56^{br} NK cells at Day 27.

a-b, Volcano plots depicting gene expression changes between Day 0 and Day 27 for the five assessed immune populations, for unstimulated (**a**) and stimulated (**b**) cells. Significantly differentially expressed genes are coloured in red (upregulated genes) or blue (downregulated genes). Names of top five up- and downregulated genes as defined by fold change are labelled on each panel. **c**, Differential expression of Day 27 signature genes. Day 27 signature genes are defined as the 40 genes that are significantly differentially expressed in at least one cell population comparing Day 27 with Day 0. Dot colours represent $\log_2$ fold change between Day 0 and Day 27. Larger dots represent genes that are significantly differentially expressed in a certain cell population. Manually annotated signature gene subsets are labelled with dashed boxes.

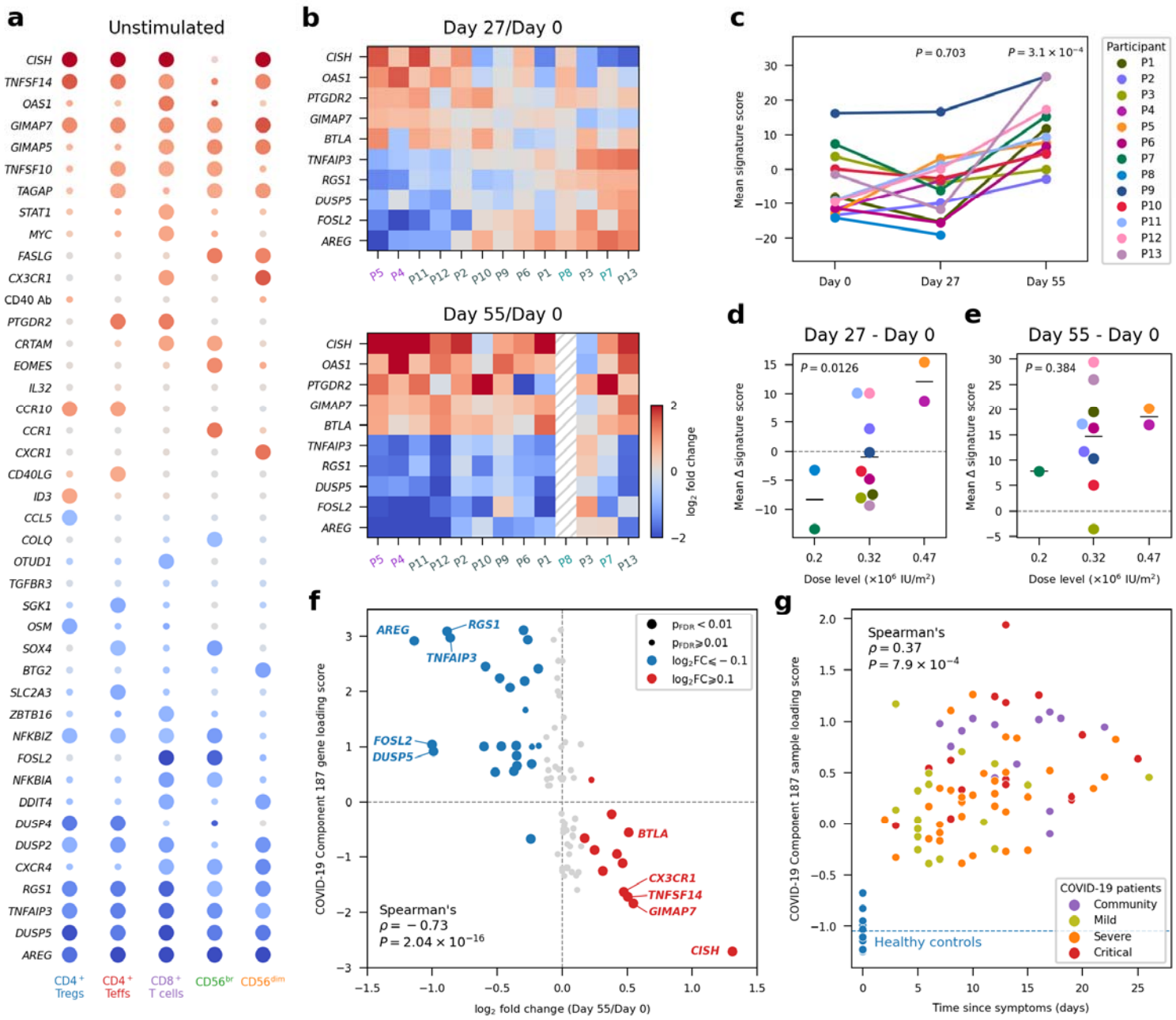
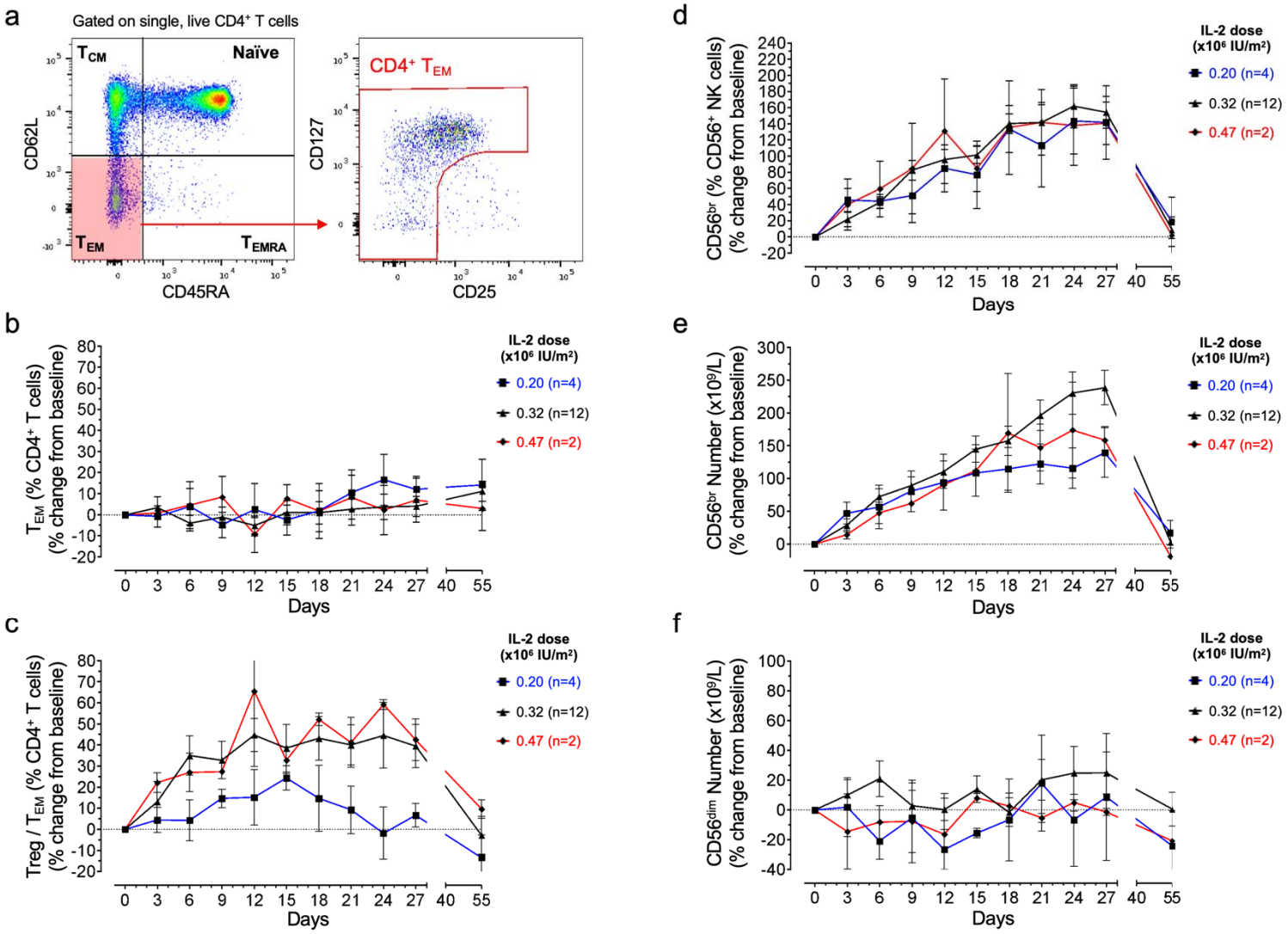


Figure 6. Interval low-dose IL-2 immunotherapy induces a long-lasting anti-inflammatory gene expression signature reversed in COVID-19 patients

a, Differential expression of Day 55 signature genes. Day 55 signature genes were defined as the 41 genes significantly differentially expressed in at least one cell population comparing Day 55 with Day 0. Dot colours represent $\log_2$ fold-change between Day 0 and Day 55. Larger dots represent significantly differentially expressed genes in at least one cell population. **b**, Participant-specific differential expression in $CD8^+$ T cells comparing Day 27 (top panel) or Day 55 (bottom panel) with Day 0. Top 5 up- and downregulated genes as defined by Day 55/Day 0 fold-change in $CD8^+$ T cells. Participant labels are

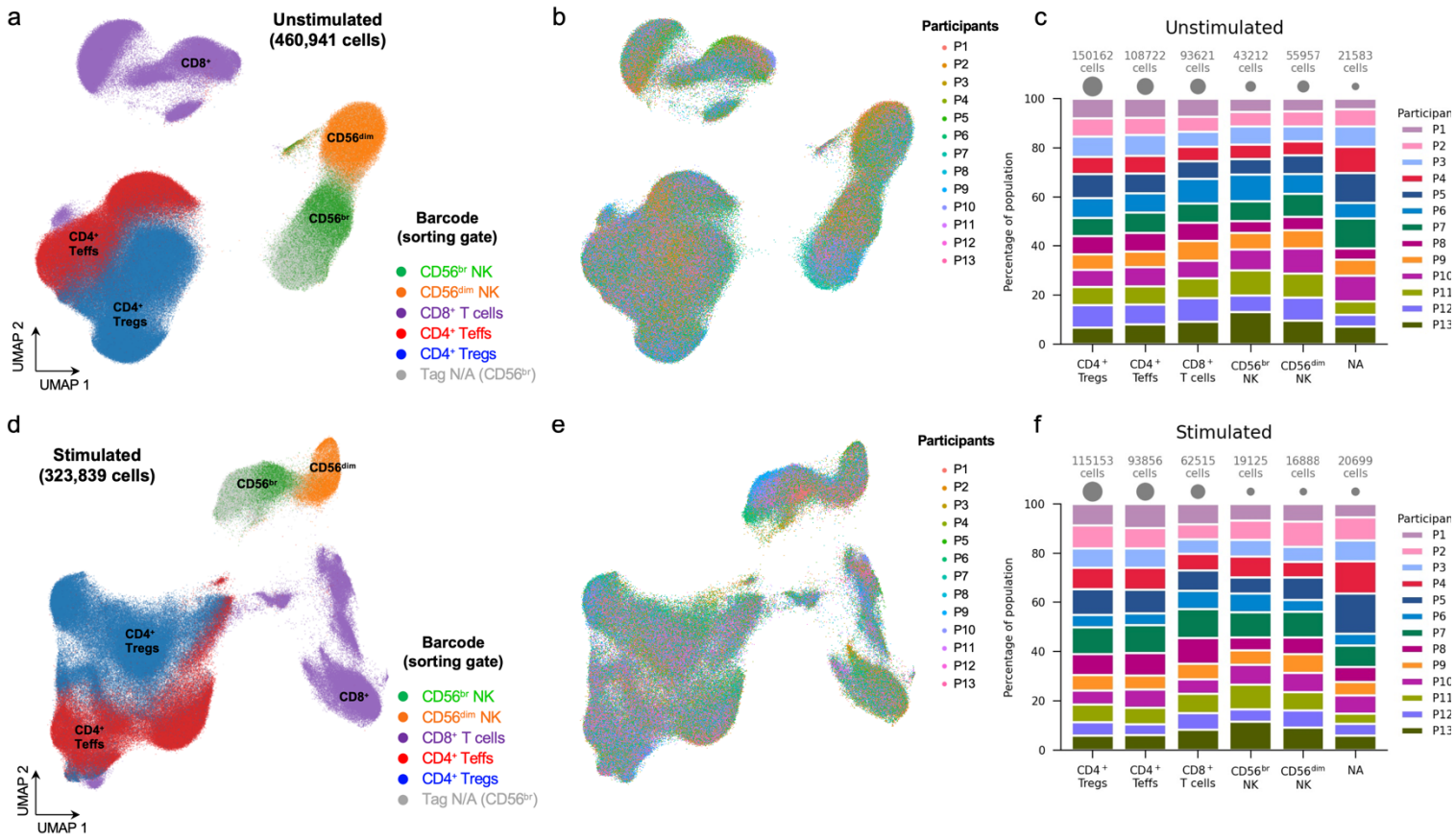
coloured by their respective IL-2 dose ($\times 10^6$ IU/m^2): 0.2 (cyan), 0.32 (gray), 0.47 (purple). **c**, Day 55 signature scores by participant for each time point. Each participant is represented by a different colour. *P* values were calculated using a two-tailed paired t test comparing Day 27 or Day 55 with Day 0, respectively. **d-e**, Variation in the expression of Day 55 signature genes (Δ signature score) between Day 27 and Day 0 (**d**) and between Day 55 and Day 0 (**e**) for each participant. Participants were stratified by IL-2 dose. *P* values were calculated using linear regression. In **c-e**, average values of the Day 55 signature in all five cell types are shown for each participant. **f**, Correlation between the Day 55 differential expression and Component 187 from the Oxford COMBAT study[30]. Of the 1,419 genes included in Component 187, a gene expression signature induced by COVID-19, 77 are also present in the target transcriptional panel used in this study. The effects of IL-2 immunotherapy and COVID-19 infection on the induction of these genes are compared by correlating Component loading score (y axis) with $\log_2$ fold change (Day 55/Day 0) in the CD8 population after IL-2 treatment (x axis). The positive and negative values in Component 187 loading scores represent the up- and downregulation of the corresponding gene, respectively, after COVID-19 infection. Larger dots represent genes with FDR-adjusted *P* values < 0.01, comparing Day 55 and Day 0 expression levels. **g**, Component 187 sample loading scores plotted against time since symptoms for COVID-19 patients. Each dot represents a patient. Patients are coloured by COVID-19 severity, according to World Health Organisation categorical scales (mild as no oxygen requirement, severe as ≤ 93% oxygen saturation, and critical as requiring mechanical ventilation). For 11 patients where multiple samples were available in the COMBAT dataset, only the earliest sample is shown. Community patients represent non-hospitalised health-care workers with mild symptoms. All other patients were hospitalised at the time of sampling. The blue dots represent the loading scores of the 10 healthy control participants, with a blue horizontal dashed line denoting the mean. In **f-g**, Spearman's rank correlation coefficient ρ and the corresponding *P* value are shown. In **g**, only COVID-19 patients are included in the correlation analysis.

Extended Data Figures



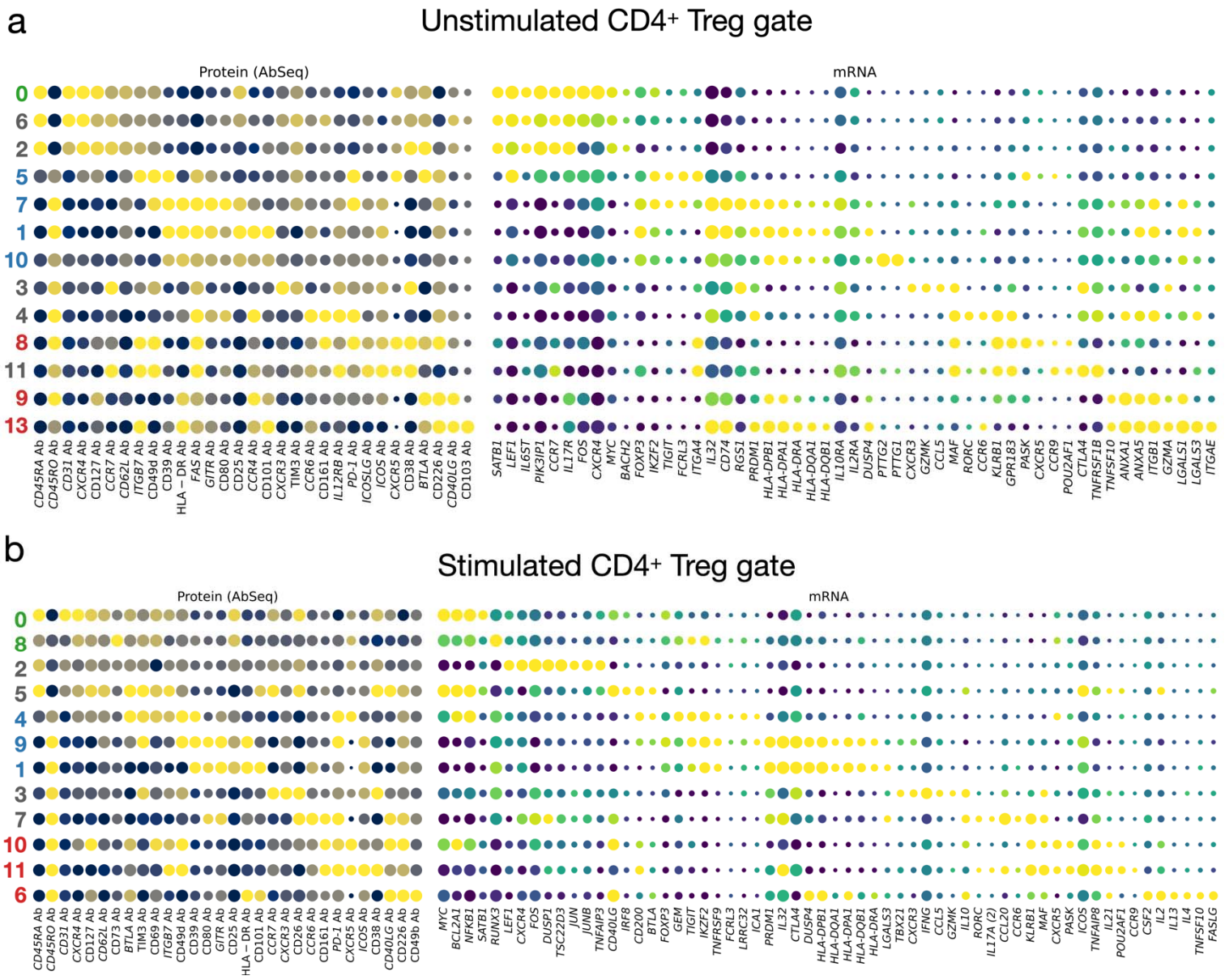
Extended Data Figure 1. Flow cytometric characterisation of the effector memory CD4$^+$ T and CD56$^+$ NK cell populations.

a, Gating strategy for the delineation of CD4$^+$ effector memory T cells (T_{EM}). **b-c**, Variation (depicted as the % change from pre-treatment baseline levels) in the frequency of CD4$^+$ T_{EM} (**b**), or the ratio between CD4$^+$ Tregs and T_{EM} cells (**c**) during the course of the study. **d-f,** Variation (depicted as the % change from pre-treatment baseline levels) in the relative frequency of CD56^{br} cells within the total CD3$^-$CD56$^+$ NK cell gate (**d**), and in the absolute number of CD56^{br} (**e**) and CD56^{dim} (**f**) NK cells in blood during the course of the study. Frequencies of the assessed immune subsets were determined by flow cytometry in the 18 DILfrequency participants treated with the 3-day interval dosing schedule. Data shown depict the average (± SEM) variation of the assessed immune subsets at each visit and were stratified according to the IL-2 dose.



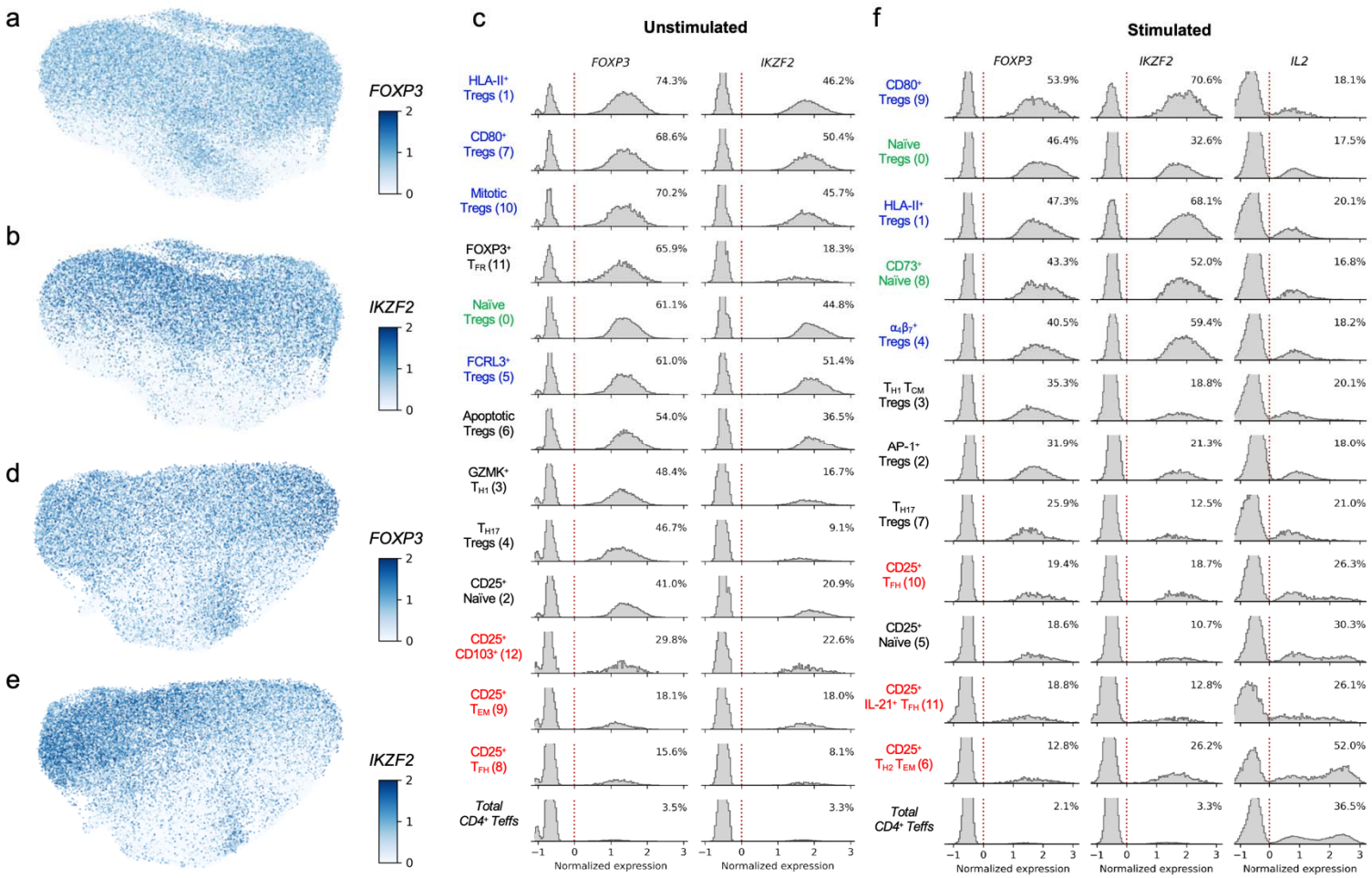
Extended Data Figure 2. Integration of the single-cell multiomics data.

a-b, Uniform Manifold Approximation and Projection (UMAP) plot depicting the clustering of the 460,941 unstimulated cells profiled in this study. Cells were coloured according to the corresponding oligo-conjugated barcode indicative of the respective FACS sorting gate (**a**), or according to the participant of origin (**b**). Untagged cells in **b** (labelled as 'Tag N/A') correspond to CD56^{br} NK cells, where we observed a suboptimal sample tagging efficiency, resulting in a significant fraction (37.2% in unstimulated cells) of CD56^{br} NK cells with missing sample barcode information. **c**, Relative proportion of cells from each individual participant in each of the flow-sorted unstimulated immune cell populations. Areas of grey circles represent total cell numbers in each immune cell population. **d-e**, UMAP plots depicting the clustering of the 323,839 cells profiled in this study following a short (90 min) *in vitro* stimulation with PMA + ionomycin. Cells were coloured according to the corresponding oligo-conjugated barcode indicative of the respective FACS sorting gate (**d**), or to the donor information (**e**). Untagged cells in panel a (labelled as 'Tag N/A') correspond to CD56^{br} cells, where we observed a suboptimal sample tagging efficiency resulting in a significant fraction (44.3% in the *in vitro* stimulated cells) of CD56^{br} NK cells with missing sample barcode information. **f**, Relative proportion of cells from each individual participant in each of the flow-sorted stimulated immune cell populations. Areas of grey circles represent total cell numbers in each immune cell population.



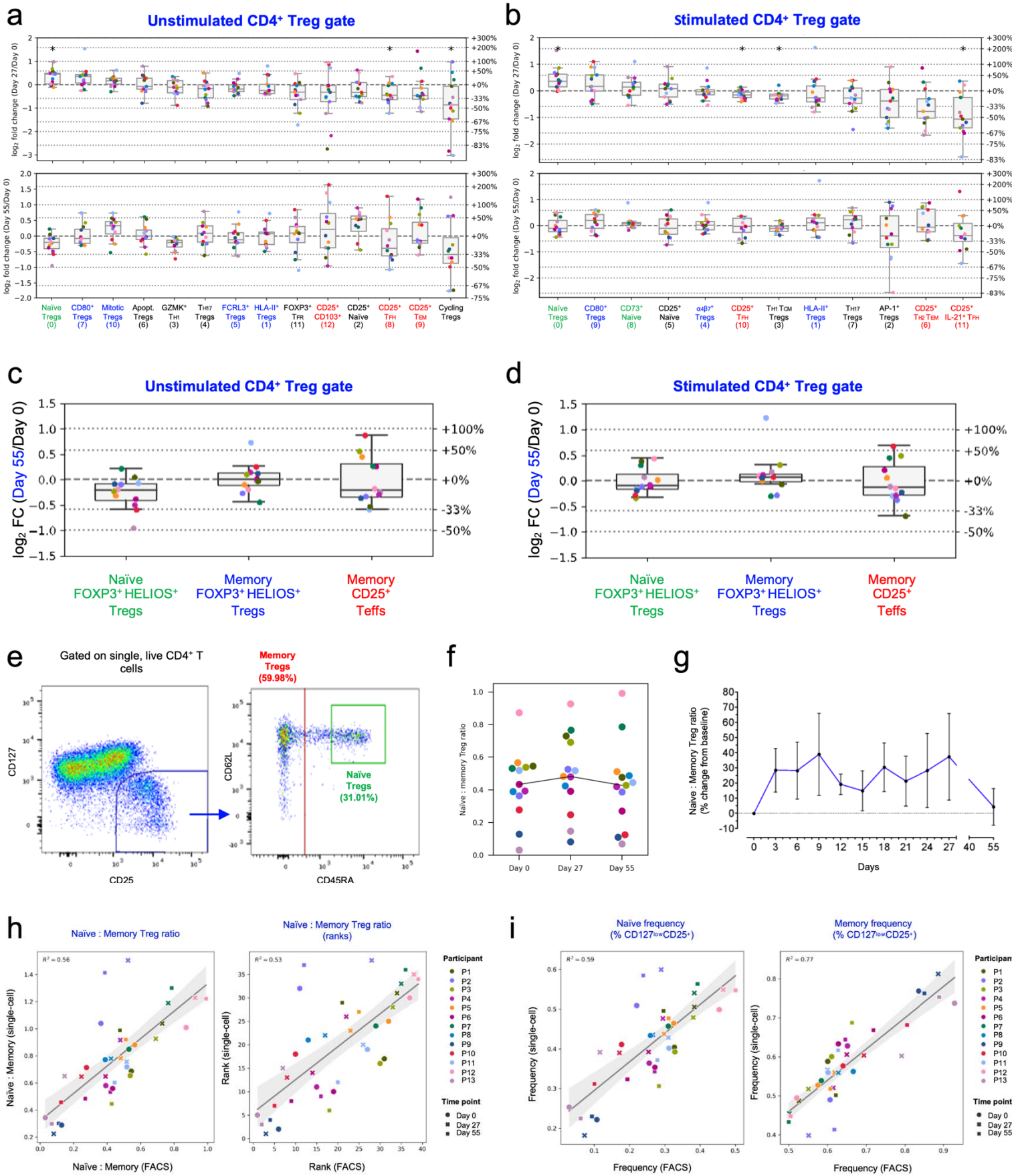
Extended Data Figure 3. Functional annotation of the CD4^{+} CD127^{low}CD25^{hi} T cell subsets.

a-b, The expression levels of selected differentially expressed markers delineating the functional subsets in the unstimulated (**a**) and *in vitro* stimulated (**b**) CD4^{+} T cells sorted from the CD127^{low}CD25^{hi} Treg gate. The size of the dots represents the frequency of the marker in the respective cluster and the colour brightness represents the relative expression level. Protein markers (AbSeq) are shown on the left section of the plots and mRNA markers on the right section. Clusters corresponding to: (i) naïve Tregs (defined as CD45RA^{+} FOXP3^{+}HELIOS^{+}); (ii) memory Tregs (defined as CD45RA^{-} FOXP3^{+}HELIOS^{+}); or (iii) Memory CD25^{+} Teffs (defined as CD45RA^{-} FOXP3^{-}HELIOS^{-}) are annotated in green, blue and red, respectively. All other miscellaneous clusters not conforming to this definition are annotated in black.



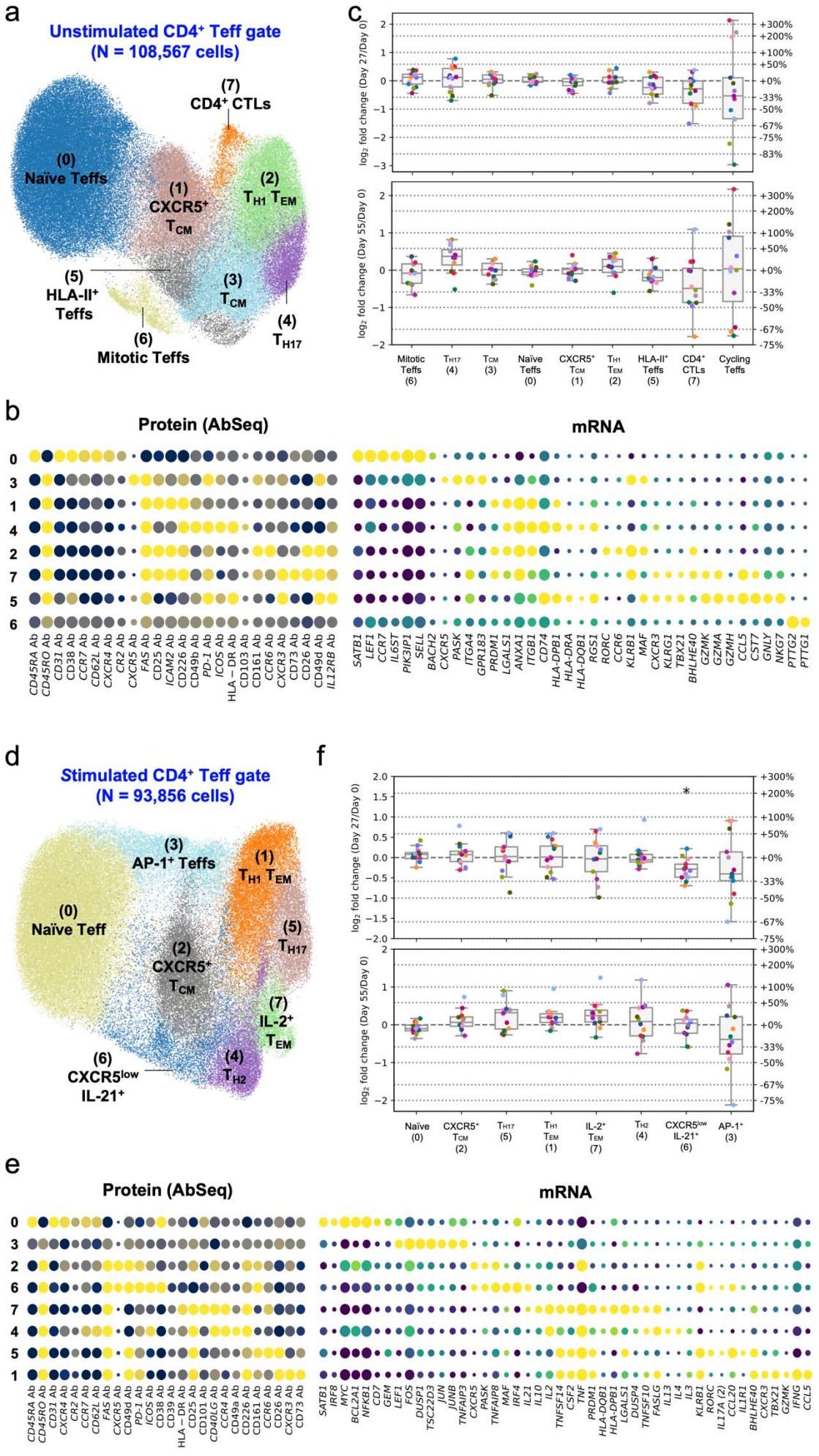
Extended Data Figure 4. Expression of *FOXP3*, *IKZF2* (HELIOS) and *IL2* differentiate functional CD4$^+$ Treg and CD4$^+$ Teffs subsets.

a-b, UMAP plots depicting the expression of the canonical Treg transcription factors *FOXP3* (**a**) and *IKZF2* (encoding HELIOS; **b**) in the unstimulated cells sorted from the CD4$^+$ CD127^{low}CD25^{hi} Treg gate. **c**, Distribution of expression levels of *FOXP3* and *IKZF2* in the identified unstimulated CD127^{low}CD25^{hi} CD4$^+$ T cell clusters. **d-e**, UMAP plots depicting the expression of the canonical Treg transcription factors *FOXP3* (**d**) and *IKZF2* (**e**) in the *in vitro* stimulated cells. **f**, Histograms depicting the expression of *FOXP3*, *IKZF2* and *IL2* in the identified *in vitro* stimulated clusters. Clusters corresponding to: (i) naïve Tregs (defined as CD45RA$^+$ FOXP3$^+$HELIOS$^+$); (ii) memory Tregs (defined as CD45RA$^-$ FOXP3$^+$HELIOS$^+$); or (iii) Memory CD25$^+$ Teffs (defined as CD45RA$^-$ FOXP3$^-$HELIOS$^-$) are annotated in green, blue and red, respectively. All other miscellaneous clusters not conforming to this definition are annotated in black.


a
Unstimulated CD4+ Treg gate
b
Stimulated CD4+ Treg gate
log2 fold change (Day 27/Day 0)
log2 fold change (Day 55/Day 0)
+300%
+200%
+100%
+50%
+0%
-33%
-50%
-67%
-75%
-83%
Naïve Tregs (0)
CD80+ Tregs (7)
Mitotic Tregs (10)
Apopt. Tregs (6)
GZMK+ TH1 (3)
TH17 Tregs (4)
FCRL3+ Tregs (5)
HLA-II+ Tregs (1)
FOXP3+ TFR (11)
CD25+ CD103+ (12)
CD25+ Naïve (2)
CD25+ TFH (8)
CD25+ TEM (9)
Cycling Tregs
CD73+ Naïve (8)
CD25+ Naïve (5)
αvβ7+ Tregs (4)
CD25+ TFH (10)
TH1 TCM Tregs (3)
HLA-II+ Tregs (1)
TH17 Tregs (7)
AP-1+ Tregs (2)
CD25+ TH2 TEM (6)
CD25+ IL-21+ TFH (11)
c
Unstimulated CD4+ Treg gate
d
Stimulated CD4+ Treg gate
log2 FC (Day 55/Day 0)
+100%
+50%
+0%
-33%
-50%
Naïve FOXP3+ HELIOS+ Tregs
Memory FOXP3+ HELIOS+ Tregs
Memory CD25+ Teffs
e
Gated on single, live CD4+ T cells
Memory Tregs (59.98%)
Naïve Tregs (31.01%)
CD127
CD25
CD62L
CD45RA
f
Naïve : memory Treg ratio
Day 0
Day 27
Day 55
g
Naïve : Memory Treg ratio (% change from baseline)
Days
h
Naïve : Memory Treg ratio
Naïve : Memory Treg ratio (ranks)
Naïve : Memory (single-cell)
Naïve : Memory (FACS)
Rank (single-cell)
Rank (FACS)
Participant
P1
P2
P3
P4
P5
P6
P7
P8
P9
P10
P11
P12
P13
Time point
Day 0
Day 27
Day 55
i
Naïve frequency (% CD127lowCD25+)
Memory frequency (% CD127lowCD25+)
Frequency (single-cell)
Frequency (FACS)

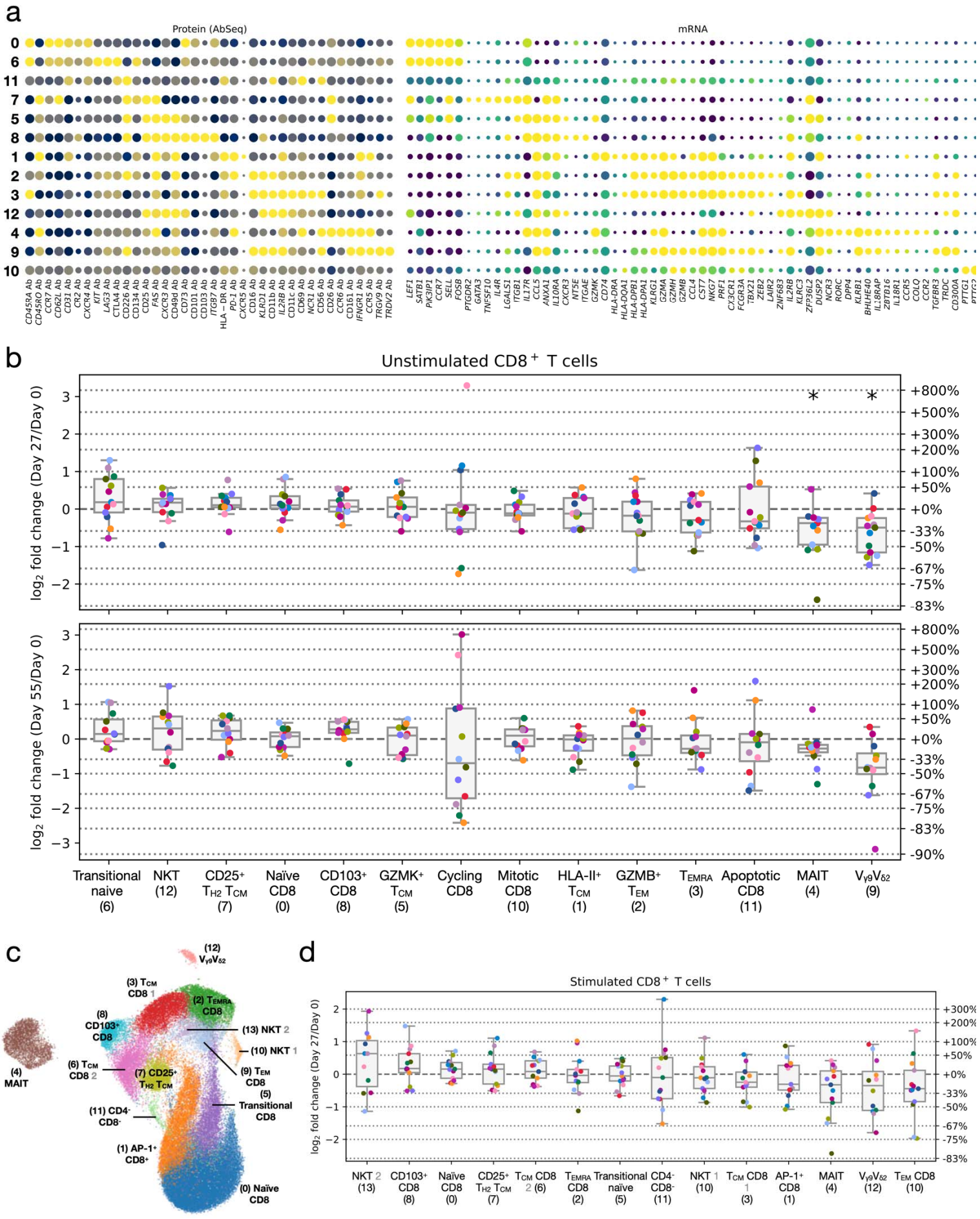
Extended Data Figure 5. Differential abundance analysis reveals a transient increased frequency of FOXP3^{+}HELIOS^{+} Tregs following ld-IL2 treatment.

a-b, Relative frequency changes of all the identified unstimulated (**a**) or *in vitro* stimulated (**b**) CD127^{low}CD25^{hi} T cell clusters between Day 0 and Day 27 (top panels) or Day 55 (bottom panels). Clusters corresponding to: (i) naïve Tregs (defined as CD45RA^{+} FOXP3^{+}HELIOS^{+}), (ii) memory Tregs (defined as CD45RA^{-} FOXP3^{+}HELIOS^{+}), and (iii) Memory CD25^{+} Teffs (defined as CD45RA^{-} FOXP3^{-} HELIOS^{-}) are annotated in green, blue and red, respectively. All other miscellaneous clusters not conforming to this definition are annotated in black. Asterisks denote statistically significant differences (two-sided Wilcoxon signed-rank test) in cell population frequencies between the two time points. **c-d**, Relative frequency changes of unstimulated (**c**) or *in vitro* stimulated (**d**) naïve FOXP3^{+}HELIOS^{+} Treg (green), memory FOXP3^{+}HELIOS^{+} Treg (blue), and memory CD25^{+} FOXP3^{-}HELIOS^{-} (red) Teff clusters in blood between Day 0 and Day 55. In **a-d**, each dot represents cells from a single participant. Dots with the same colour represent the same participant. **e**, Gating strategy for the delineation of the naïve (CD45RA^{+}) and memory (CD45RA^{-}) compartment by flow cytometry within cells sorted from the CD4^{+} CD127^{low}CD25^{hi} Treg gate. **f**, Ratio between naïve and memory cells within the CD4^{+} CD127^{low}CD25^{hi} Treg gate at the three assessed timepoints (Day 0, Day 27 and Day 55). The connecting line represents the median naïve : memory ratio in the 13 participants selected for single-cell analysis at each time point. **g**, Variation (depicted as the % change from baseline pre-treatment levels) in the naïve : memory ratio within the CD4^{+} CD127^{low}CD25^{hi} Treg gate assessed by flow cytometry. Data shown depict the average (±SEM) variation at each visit for the 13 participants selected for single-cell analysis. **h**, Correlation between the naïve : memory ratio within CD4^{+} CD127^{low}CD25^{hi} T cells (left panel), or the corresponding ranks (right panel), obtained from FACS data (x axis) and that estimated from single-cell sequencing data (y axis). **i**, Correlation between the frequency of naïve (CD45RA^{+}; left panel) and memory (CD45RA^{-}; right panel) cells within CD127^{low}CD25^{hi} T cells obtained from FACS data (x axis) and that obtained from single-cell sequencing data (y axis). In **f**, **h** and **i**, each dot represents a sample, with colours and shapes representing participants and time points of origin, respectively. In **h-i**, linear regression lines are shown in black, and 95% confidence intervals of the regression estimate are shown in grey shades.


a
Unstimulated CD4+ Teff gate
(N = 108,567 cells)
(0) Naïve Teffs
(1) CXCR5+ TCM
(2) TH1 TEM
(3) TCM
(4) TH17
(5) HLA-II+ Teffs
(6) Mitotic Teffs
(7) CD4+ CTLs
c
log2 fold change (Day 27/Day 0)
log2 fold change (Day 55/Day 0)
+300%
+200%
+100%
+50%
+0%
-33%
-50%
-67%
-75%
-83%
Mitotic Teffs (6)
TH17 (4)
TCM (3)
Naïve Teffs (0)
CXCR5+ TCM (1)
TH1 TEM (2)
HLA-II+ Teffs (5)
CD4+ CTLs (7)
Cycling Teffs
b
Protein (AbSeq)
mRNA
CD45RA Ab
CD45RO Ab
CD31 Ab
CD38 Ab
CCR7 Ab
CD62L Ab
CXCR4 Ab
CR2 Ab
CXCR5 Ab
FAS Ab
CD25 Ab
ICAM2 Ab
CD226 Ab
CD49b Ab
PD-1 Ab
ICOS Ab
HLA – DR Ab
CD103 Ab
CD161 Ab
CCR6 Ab
CXCR3 Ab
CD73 Ab
CD26 Ab
CD49d Ab
IL12RB Ab
SATB1
LEF1
CCR7
IL6ST
PIK3IP1
SELL
BACH2
CXCR5
PASK
ITGA4
GPR183
PRDM1
LGALS1
ANXA1
ITGB1
CD74
HLA-DPB1
HLA-DRA
HLA-DQB1
RGS1
RORC
CCR6
KLRB1
MAF
CXCR3
KLRG1
TBX21
BHLHE40
GZMK
GZMA
GZMH
CCL5
CST7
GNLY
NKG7
PTTG2
PTTG1
d
Stimulated CD4+ Teff gate
(N = 93,856 cells)
(0) Naïve Teff
(1) TH1 TEM
(2) CXCR5+ TCM
(3) AP-1+ Teffs
(4) TH2
(5) TH17
(6) CXCR5low IL-21+
(7) IL-2+ TEM
f
log2 fold change (Day 27/Day 0)
log2 fold change (Day 55/Day 0)
*
Naïve (0)
CXCR5+ TCM (2)
TH17 (5)
TH1 TEM (1)
IL-2+ TEM (7)
TH2 (4)
CXCR5low IL-21+ (6)
AP-1+ (3)
e
Protein (AbSeq)
mRNA
CD45RA Ab
CD45RO Ab
CD31 Ab
CXCR4 Ab
CR2 Ab
CCR7 Ab
CD62L Ab
FAS Ab
CXCR5 Ab
CD49d Ab
PD-1 Ab
ICOS Ab
CD38 Ab
CD39 Ab
HLA – DR Ab
CD25 Ab
CD101 Ab
CD40LG Ab
CCR4 Ab
CD49a Ab
CD226 Ab
CD161 Ab
CCR6 Ab
CD26 Ab
CXCR3 Ab
CD73 Ab
SATB1
IRF8
MYC
BCL2A1
NFKB1
CD7
GEM
LEF1
FOS
DUSP1
TSC22D3
JUN
JUNB
TNFAIP3
CXCR5
PASK
TNFAIP8
MAF
IRF4
IL21
IL10
IL2
TNFSF14
CSF2
TNF
PRDM1
HLA-DQB1
HLA-DPB1
LGALS1
DUSP4
TNFSF10
FASLG
IL13
IL4
IL3
KLRB1
RORC
IL17A (2)
CCL20
IL1R1
BHLHE40
CXCR3
TBX21
GZMK
IFNG
CCL5

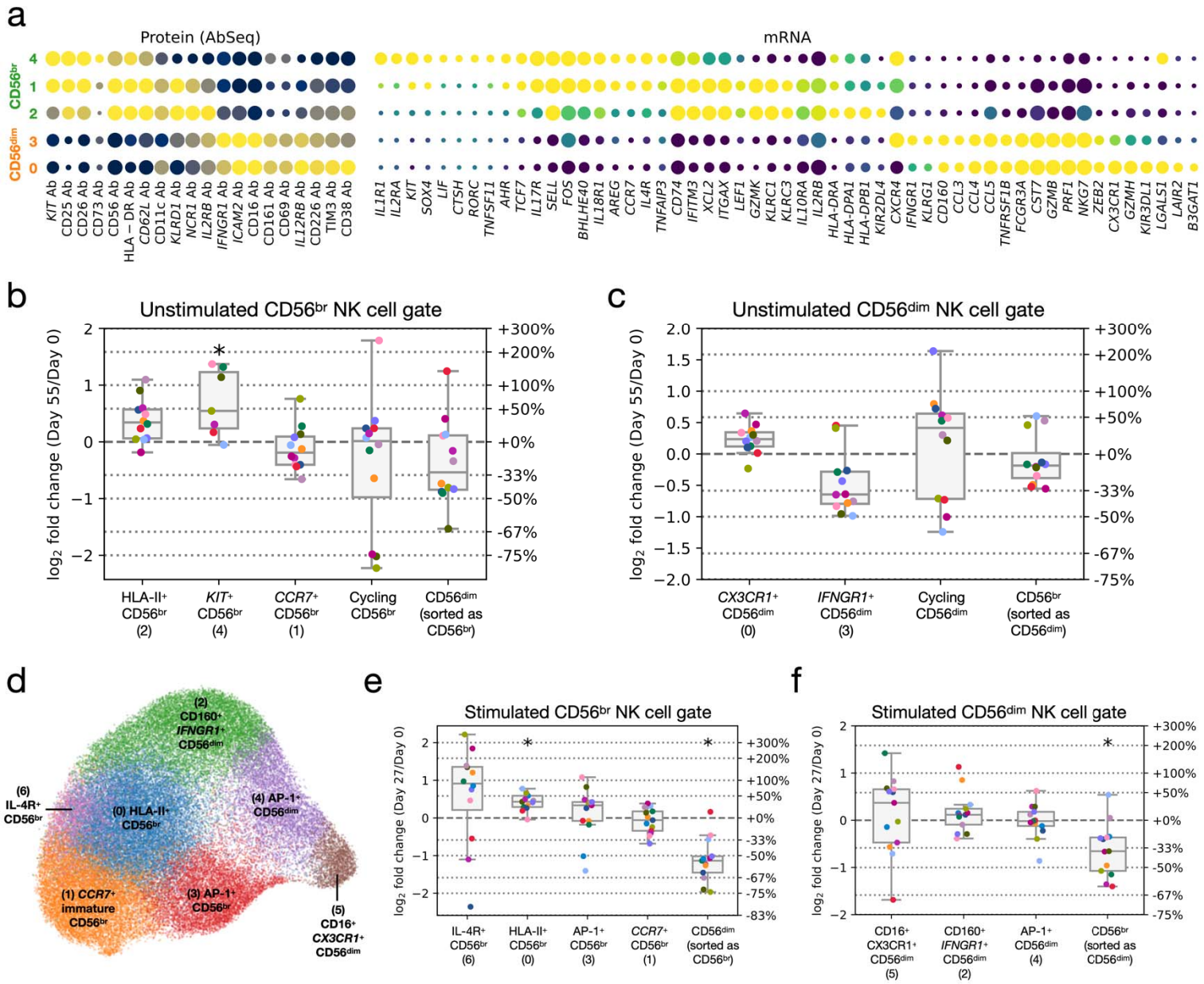
Extended Data Figure 6. Frequency of CD4^{+} CD25$^{-/low}$ effector T cells in blood is not affected by low-dose IL-2 immunotherapy.

a, UMAP plot depicting the clustering of the 108,722 unstimulated CD4^{+} T cells sorted from the CD25$^{-/low}$ effector T cell (Teff) gate profiled in this study. Clusters were manually annotated based on the expression of key differentially expressed mRNA and protein markers. **b**, The expression levels of selected differentially expressed markers delineating the functional subsets in the unstimulated CD4^{+} T cells sorted from the CD25$^{-/low}$ effector T cell (Teff) gate. The size of the dots represents the frequency of the marker in the respective cluster and the colour brightness represents the relative expression level. Protein markers (AbSeq) are shown on the left section of the plots and mRNA markers on the right section. **c,** Relative frequency changes of all identified unstimulated CD4^{+} CD25$^{-/low}$ Teff clusters in blood at Day 27 (top panel) or Day 55 (bottom panel) compared to the baseline pre-treatment (Day 0) levels. **d-f**, Plots corresponding to **a-c**, for the 93,856 *in vitro* stimulated CD4^{+} T cells sorted from the CD25$^{-/low}$ Teff gate. Note that "Cycling Teffs" correspond to the fraction of cells in cycle sorted from the CD4^{+} CD25$^{-/low}$ Teff gate (**Fig. 1h** and **methods**), and were not represented in the UMAP plots.


a
Protein (AbSeq)
mRNA
0
6
11
7
5
8
1
2
3
12
4
9
10
b
Unstimulated CD8+ T cells
log2 fold change (Day 27/Day 0)
log2 fold change (Day 55/Day 0)
+800%
+500%
+300%
+200%
+100%
+50%
+0%
-33%
-50%
-67%
-75%
-83%
-90%
Transitional naive (6)
NKT (12)
CD25+ TH2 TCM (7)
Naïve CD8 (0)
CD103+ CD8 (8)
GZMK+ TCM (5)
Cycling CD8
Mitotic CD8 (10)
HLA-II+ TCM (1)
GZMB+ TEM (2)
TEMRA (3)
Apoptotic CD8 (11)
MAIT (4)
Vγ9Vδ2 (9)
c
(12) Vγ9Vδ2
(3) TCM CD8 1
(2) TEMRA CD8
(8) CD103+ CD8
(13) NKT 2
(10) NKT 1
(4) MAIT
(6) TCM CD8 2
(7) CD25+ TH2 TCM
(9) TEM CD8
(5) Transitional CD8
(11) CD4- CD8-
(1) AP-1+ CD8+
(0) Naïve CD8
d
Stimulated CD8+ T cells
log2 fold change (Day 27/Day 0)
+300%
+200%
+100%
+50%
+0%
-33%
-50%
-67%
-75%
-83%
NKT 2 (13)
CD103+ CD8 (8)
Naïve CD8 (0)
CD25+ TH2 TCM (7)
TCM CD8 2 (6)
TEMRA CD8 (2)
Transitional naïve (5)
CD4- CD8- (11)
NKT 1 (10)
TCM CD8 1 (3)
AP-1+ CD8 (1)
MAIT (4)
Vγ9Vδ2 (12)
TEM CD8 (10)

Extended Data Figure 7. Functional annotation of CD8$^+$ T cells.

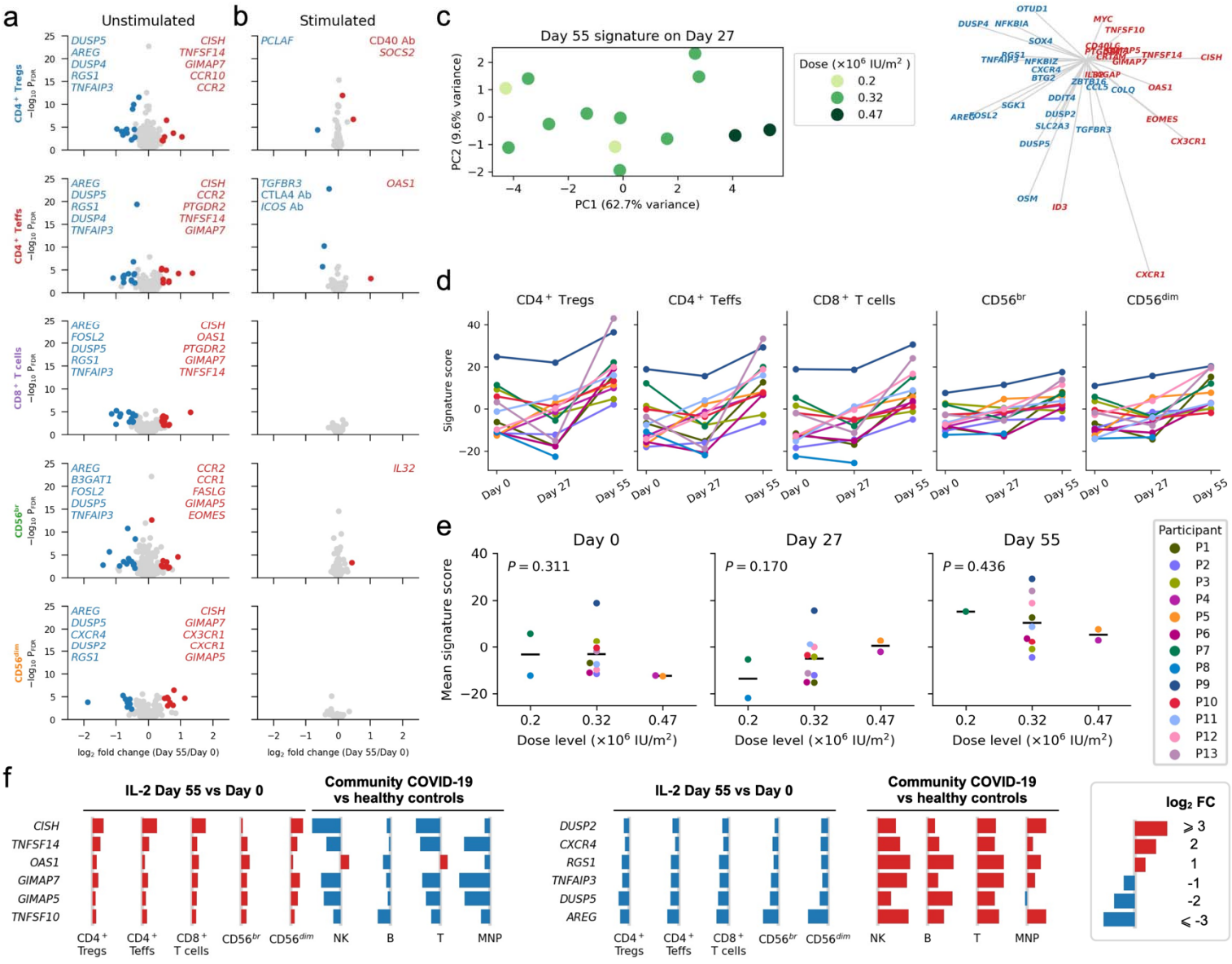
a, Expression of selected differentially expressed markers delineating the functional subsets identified from the 13 clustering of the 93,621 unstimulated CD8$^+$ T cells profiled in this study. The size of the dots represents the frequency of the marker in the respective cluster and the colour brightness represents the relative expression level. Protein markers (AbSeq) are shown on the left section of the plots and mRNA markers on the right section. **b,** Relative frequency changes of all identified CD8$^+$ T cell clusters on Day 27 (top panel) or Day 55 (bottom panel) compared to the baseline pre-treatment levels. Note that the "Cycling CD8" cluster corresponds to the fraction of cells in cycle sorted from the respective sorting gates (**Fig. 1h** and **methods**), and were not represented in the UMAP plots. **c**, Uniform Manifold Approximation and Projection (UMAP) plot depicting the clustering of *in vitro* stimulated CD8$^+$ T cells. Clusters were manually annotated based on the expression of key differentially expressed mRNA and protein markers. **d**, Relative frequency changes of the 14 identified stimulated CD8$^+$ T cell clusters at Day 27 compared to the baseline pre-treatment (Day 0) levels.



Extended Data Figure 8. Functional annotation of CD56^{+} NK cells.

a, Expression of selected differentially expressed markers delineating the 5 functional subsets identified from the clustering of the 99,169 unstimulated CD56^{+} NK cells profiled in this study. The size of the dots represents the frequency of the marker in the respective cluster and the colour brightness represents the relative expression level. Protein markers (AbSeq) are shown on the left section of the plots and mRNA markers on the right section. The dotted red line separates the clusters corresponding to either CD56^{br} (top) and CD56^{dim} NK cells (bottom). **b-c**, Relative frequency changes of all identified unstimulated CD56^{br} (**b**) and CD56^{dim} (**c**) NK cell clusters on Day 55 compared to the baseline pre-treatment levels. Note that "Cycling CD56^{br}" and "Cycling CD56^{dim}" clusters correspond to the fraction of cells in cycle

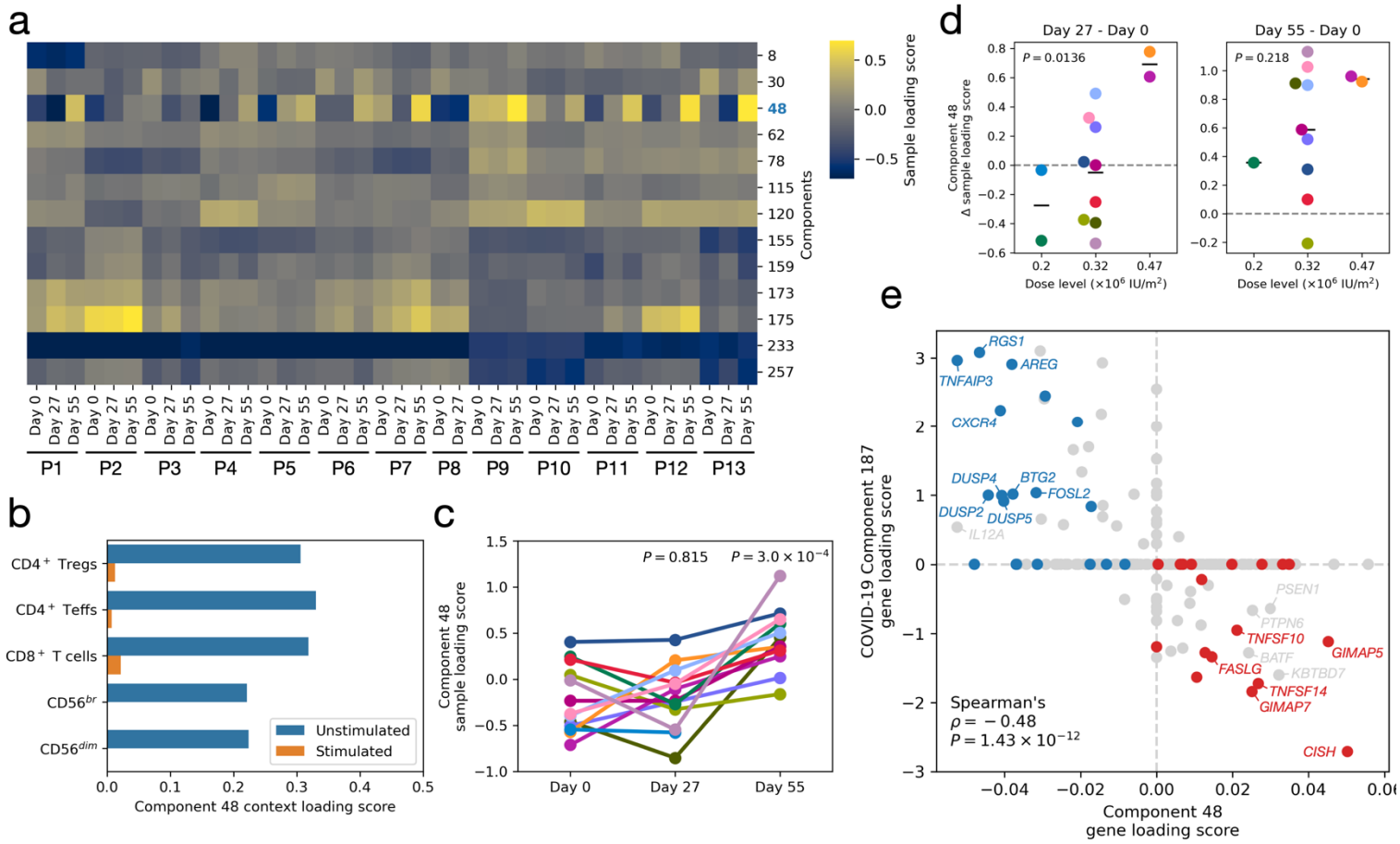
sorted from the respective sorting gates (**Fig. 1h** and **methods**), and were not represented in the UMAP plots. **d**, UMAP plot depicts the clustering of *in vitro* stimulated $CD56^{+}$ NK cells. **e-f** Relative frequency changes of all identified clusters of cells sorted from either the $CD56^{br}$ (**e**) or $CD56^{dim}$ (**f**) NK cell gates following *in vitro* stimulation on Day 27 compared to the baseline pre-treatment (Day 0) levels.



Extended Data Figure 9. Induction of the anti-inflammatory gene expression signature at Day 55 is IL-2 dose-dependent and is detected in all T and NK cell populations

a-b, Volcano plots depicting gene expression changes between Day 0 and Day 55 for the five assessed immune populations, for unstimulated (**a**) and stimulated (**b**) cells. Significantly differentially expressed genes are coloured in red (upregulated genes) or blue (downregulated genes). The top five up- and downregulated genes (as defined by fold change) are labelled on each panel. **c**, Principal component analysis (PCA) correlating gene expression changes with dose levels. PCA was performed on $\log_2$ fold-change values of Day 55 signature genes on Day 27 in $CD8^+$ T cells. Each dot represents a participant, with colours indicating IL-2 dose levels. Vectors representing the gene loading scores of PC1 and PC2 are shown on the right panel, with genes up- and downregulated on Day 55 shown in red and blue, respectively. **d**, Day 55 signature scores by participant for each time point. Each participant is represented by a different colour. **e**, Day 55 signature scores plotted against IL-2 dose. The average value of the Day

55 signature score in the five cell types is shown for each participant. **f**, Expression changes of the top 5 up- and downregulated Day 55 signature genes were compared in the context of iLD-IL2 treatment (comparing Day 55 and Day 0) and in the context of recent SARS-CoV-2 infection (comparing community COVID-19 patients and healthy controls). Data were stratified according to the assessed immune populations.



Extended Data Figure 10. Sparse Decomposition of Arrays (SDA) analysis replicates the association of the anti-inflammatory Day 55 signature with IL-2 dose and COVID-19.

a, Heatmap showing sample loading scores normalised usage of each gene expression component identified through Sparse Decomposition of Arrays (SDA) analysis. **b**, Component 48 context loading scores for each unstimulated or stimulated cell population. **c**, Component 48 sample loading scores by participant for each time point. Each participant is represented by a different colour. **d**, Difference in the Component 48 sample loading scores between Day 0 and Day 27 (left panel) or Day 55 (right panel), plotted against IL-2 dose. *P* value was calculated using linear regression. **e**, Correlation between gene loading scores of Component 48 and Component 187 from the Oxford COMBAT study[30]. Up- and downregulated Day 55 signature genes are shown in red and blue, respectively. Other genes are shown in grey. Spearman's rank correlation coefficient ρ and the corresponding *P* value are shown.

Supplementary Information

Supplementary Table 1. Baseline characteristics of study participants

Supplementary Table 2. Targeted mRNA, protein (AbSeq) and immunostaining panels used in this study

Supplementary Table 3. Characteristics of COMBAT study participants with single-cell data available

Supplementary Data 1. Summary of the cluster-specific differentially expressed markers

Supplementary Data 2. Differential expression between Day 0 and Day 27 or Day 55 analysed using DESeq2

Supplementary Data 3. Gene loading scores and posterior inclusion probabilities (PIP) for each component identified using Sparse Decomposition of Arrays (SDA)