## Supplemental Information

## Innate Lymphoid Cells Are Depleted Irreversibly

## during Acute HIV-1 Infection

## in the Absence of Viral Suppression

Henrik N. Kløverpris, Samuel W. Kazer, Jenny Mjösberg, Jenniffer M. Mabuka, Amanda Wellmann, Zaza Ndhlovu, Marisa C. Yadon, Shepherd Nhamoyebonde, Maximilian Muenchhoff, Yannick Simoni, Frank Andersson, Warren Kuhn, Nigel Garrett, Wendy A. Burgers, Philomena Kamya, Karyn Pretorius, Krista Dong, Amber Moodley, Evan W. Newell, Victoria Kasprowicz, Salim S. Abdool Karim, Philip Goulder, Alex K. Shalek, Bruce D. Walker, Thumbi Ndung'u, and Alasdair Leslie
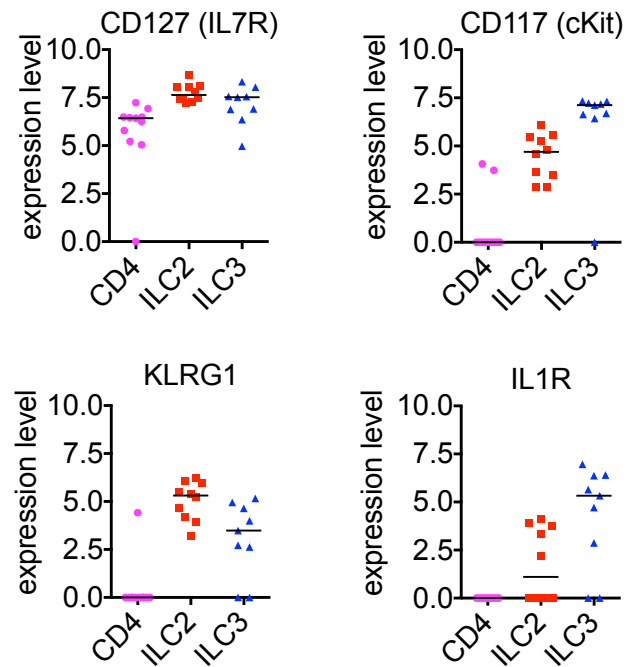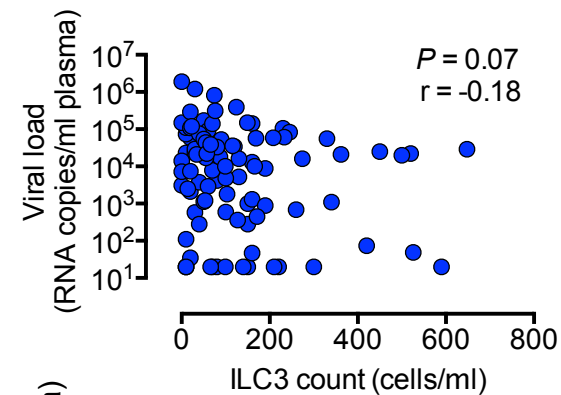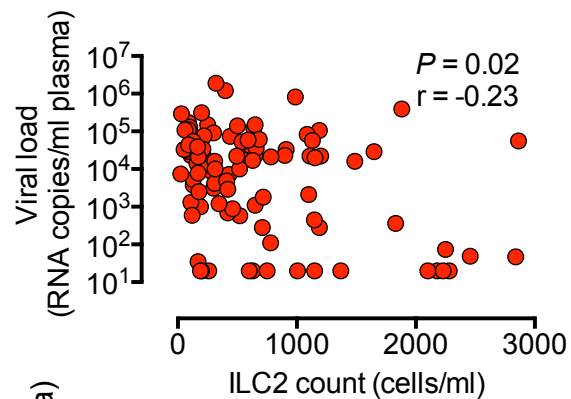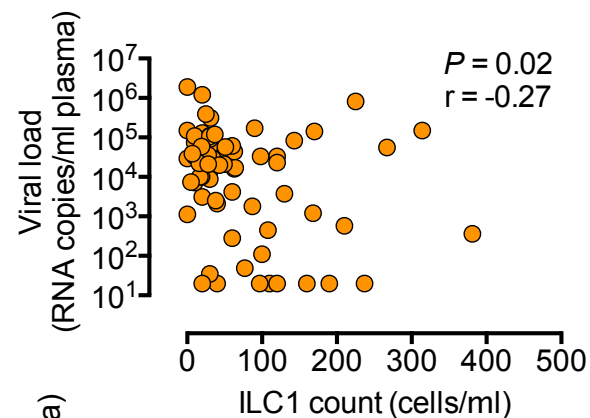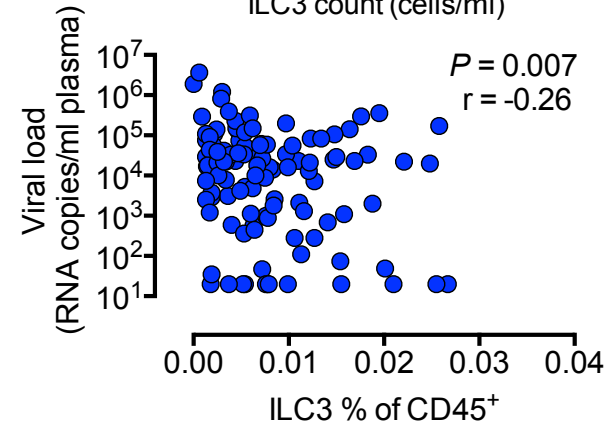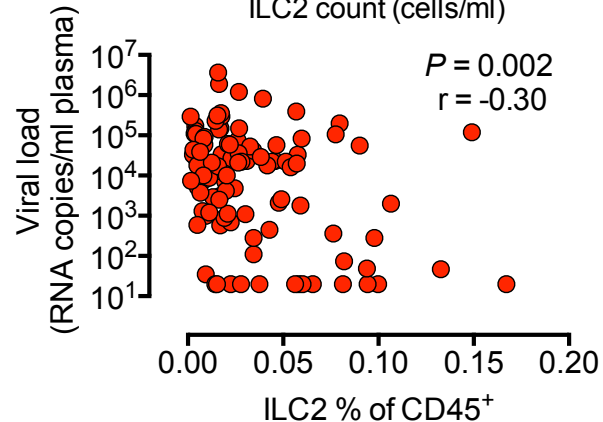
# Figure S1 related to Fig. 1
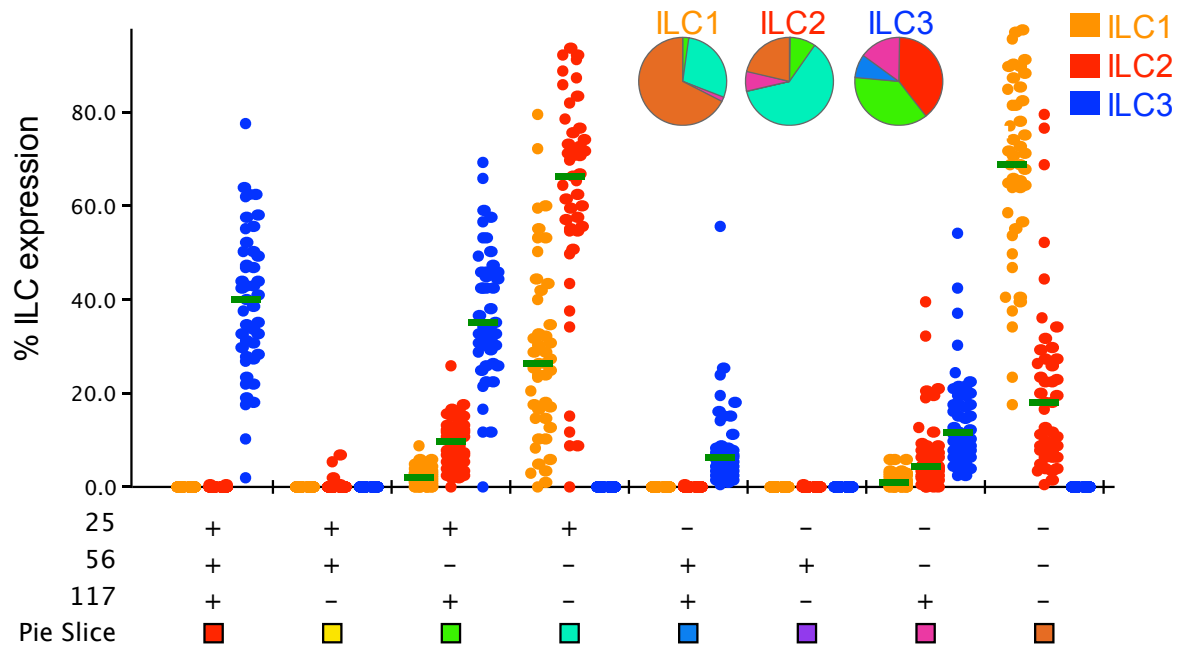
**Figure S1 related to Figure 1. Phenotype and selected transcriptional gene expression of ILC1, ILC2 and ILC3 subsets in blood. (A)** Pie charts showing distributions of 8 different ILC1, ILC2 and ILC3 subpopulations based on differential expression of CD25, CD56, CD117 and color-coded as indicated below x-axis and with y-axis showing the relative frequency of ILC1, ILC2 and ILC3 gated cells (see Figure 1A). **(B)** Data from n=43 HIV uninfected individuals. Selected canonical ILC gene expression (CD127, cKit/CD117, KLRG1 and IL1R) from 9 HIV-1 uninfected individuals obtained by RNAseq after ILC2, ILC3 and CD4$^+$ T cell sorted cells from blood.

# Figure S2 related to Fig. 1

**A**



**B**
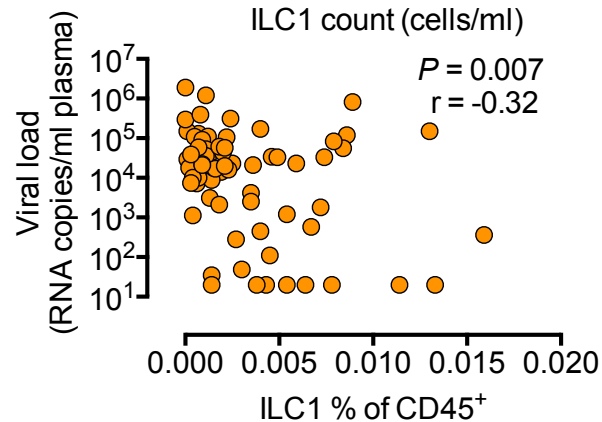


**C**
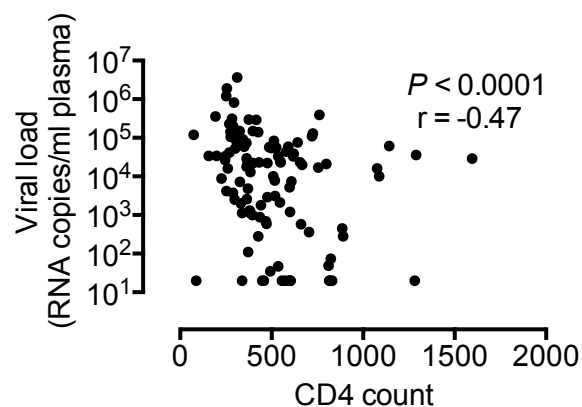
**Figure S2 related to Figure 1. ILCs are depleted in HIV-1 infection and correlates negatively with HIV RNA copies/mL plasma. (A)** Correlation of ILC counts to the viral load set-points for HIV-1 infected individuals. **(B)** Correlation of ILC frequencies expressed as % of total CD45$^+$ lymphocytes to the viral load set-points. **(C)** Correlation of absolute CD4+ T cell count and HIV-1 viral load setpoints. *P-* and R-values by correlation coefficients and shown as spearman rank *r-* and *P-*values.

# Figure S3 related to Fig. 2
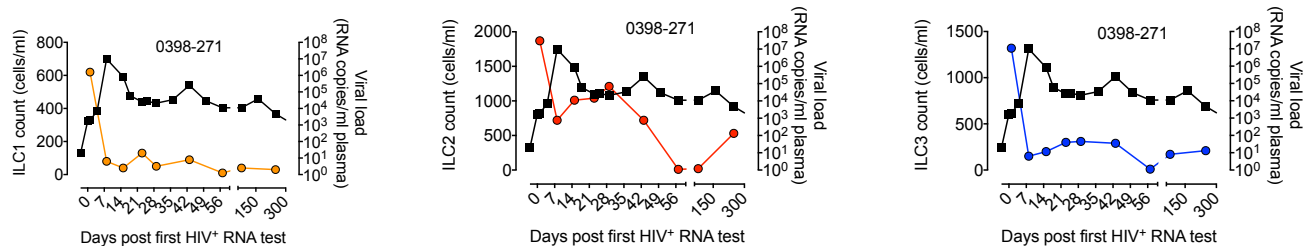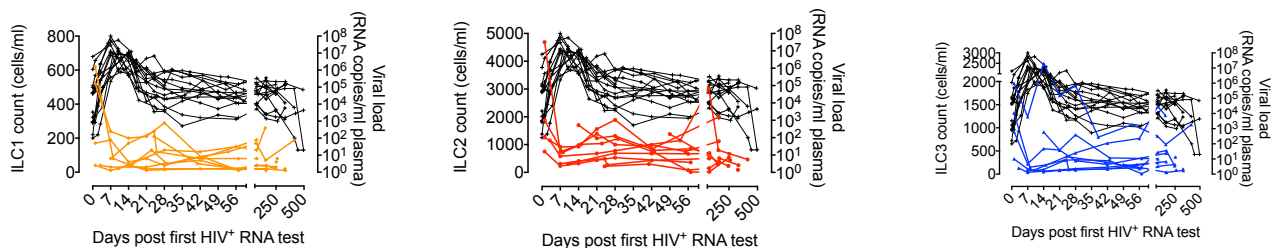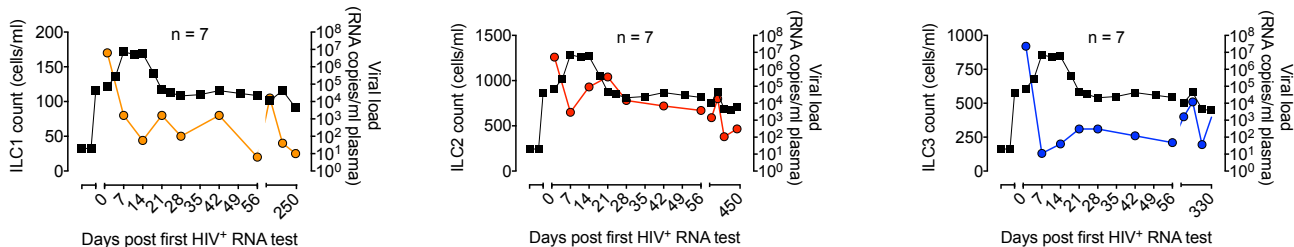
**Figure S3 related to Figure 2. ILCs are depleted during acute HIV-1 infection.**

**(A)** Data from the acutely HIV-1 infected PID 0398-271 subject followed longitudinally over 10 time-points from day 1 to day 249 from the day of first HIV$^+$ RNA test showing the absolute ILC1, ILC2 and ILC3 counts (left y-axis) with the plasma viral load shown as HIV$^+$ RNA copies/mL plasma (right y-axis, black line). **(B)** Data as in (A), but for the entire cohort (n=7). **(C)** Absolute ILC counts (cells/ml of blood) presented as cumulative data with median values shown for the entire cohort (n=7).

# Figure S4 related to Fig. 2
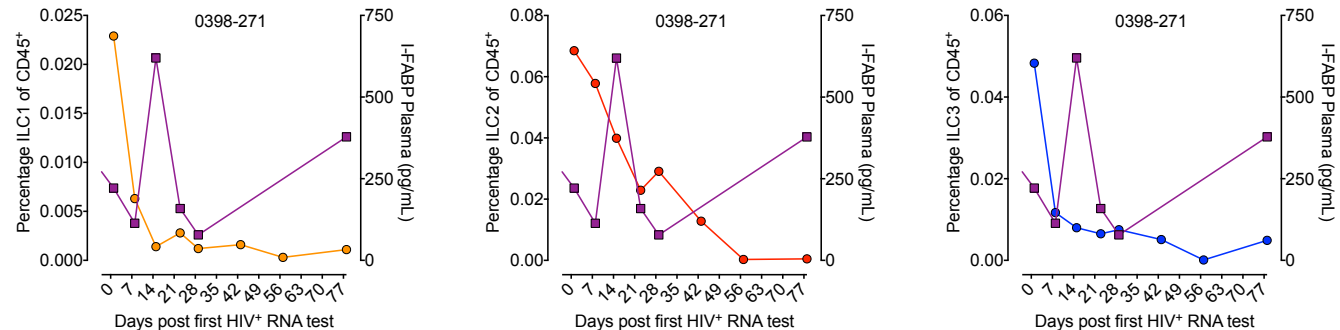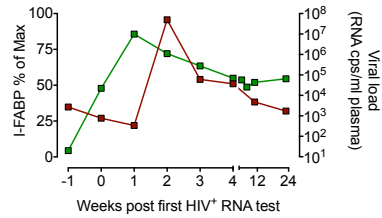
**Figure S4 related to Figure 2. Depletion of ILCs in early acute HIV-1 infection is associated with elevated levels of gut epithelial break down. (A)** Absolute Plasma Intestinal Fatty Acid Binding protein (I-FABP) levels for the acutely HIV-1 infected PID 0398-271 subject shown longitudinally over 8 time-points from day 1 to day 77 from the day of first $HIV^+$ RNA test with percentages of ILC1, ILC2 and ILC3 of total CD45+ lymphocytes (left y-axis) and with absolute plasma I-FABP levels (right y-axis, purple. **(B)** Cumulative median Pla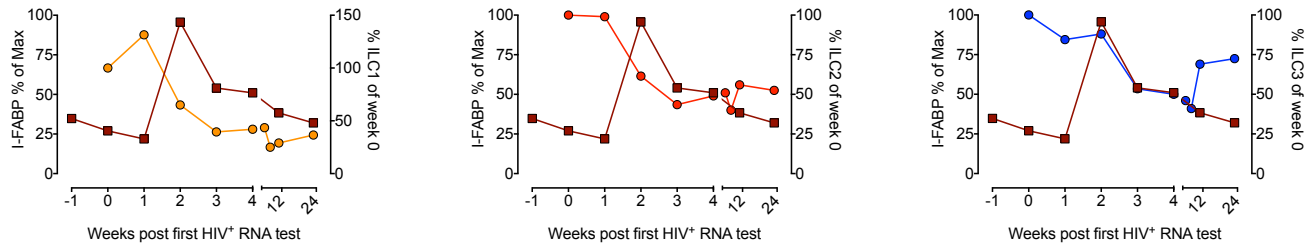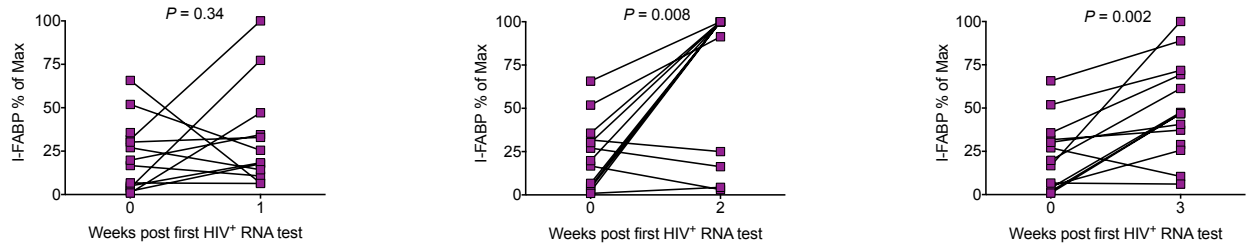sma I-FABP levels (% of the maximum value) for all acutely infected subjects, measured at the last time-point before HIV-1 infection and 7 subsequent time-points over acute infection into chronic infection (week 24) (left y-axis), plotted against the median value for ILCs over the same time period, plotted as % of week 0 ILC frequency for each ILC subset (right y-axis). **(C)** (I-FABP) levels for each subject expressed as percentage of maximum value detected during acute infection compared for week 0, 2 and 3 for matched acutely infected subjects (n=14). **(D)** Data as in (a) but showing the median values of the maximum I-FABP levels from n=14 acutely infected subjects in comparison to HIV-1 RNA copies/ml plasma. *P* - values by paired t-test.

# Figure S5 related to Fig. 5
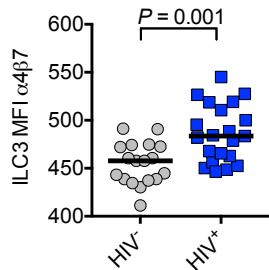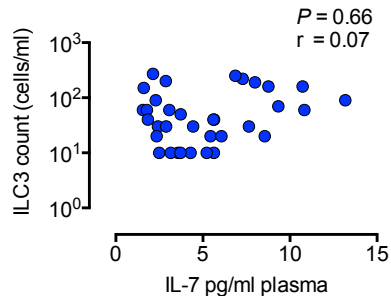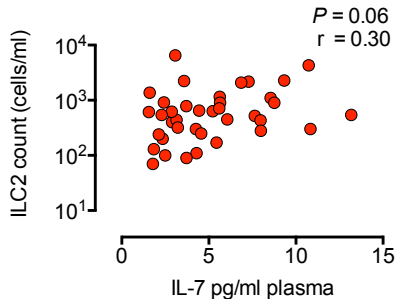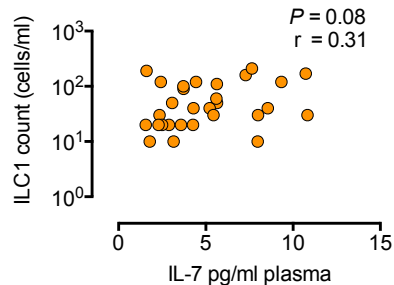
**Figure S5 related to Figure 5. Relative expression levels of α4β7 and IL-7 plasma levels in correlation to ILC counts. (A)** Relative expression levels of α4β7 expression on ILC3 gated blood cells from HIV-1 uninfected and infected subjects. **(B)** Correlation of plasma IL-7 levels (pg/ml plasma) and absolute ILC counts from chronic HIV-1 infected subjects *P* value by Mann-Withney U-test and *P* and *R*-values by spearman rank correlation.

# Figure S6 related to Fig. 6

## A
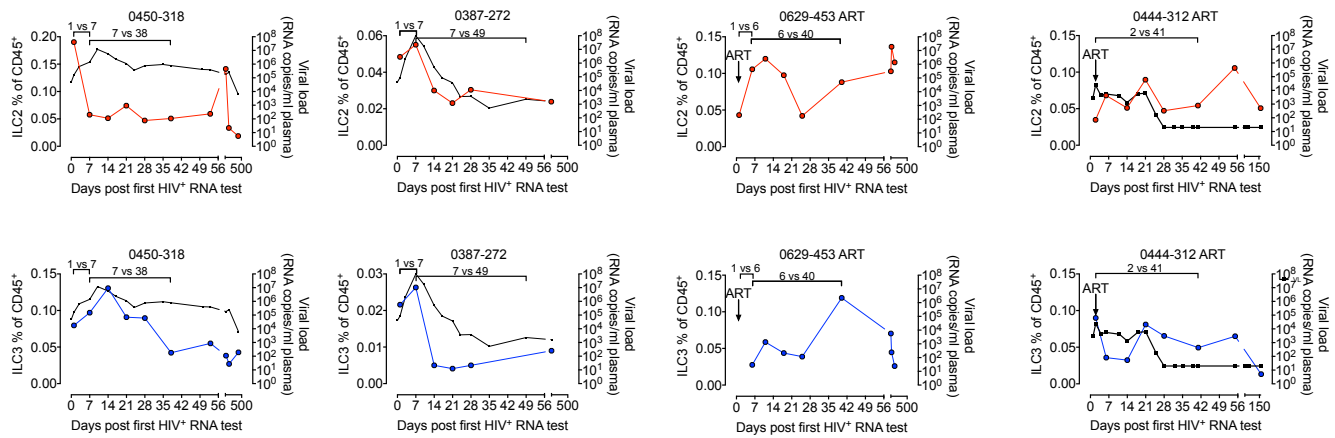


## B
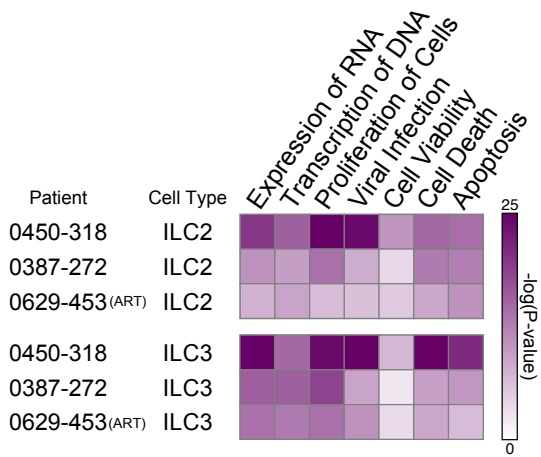


## C
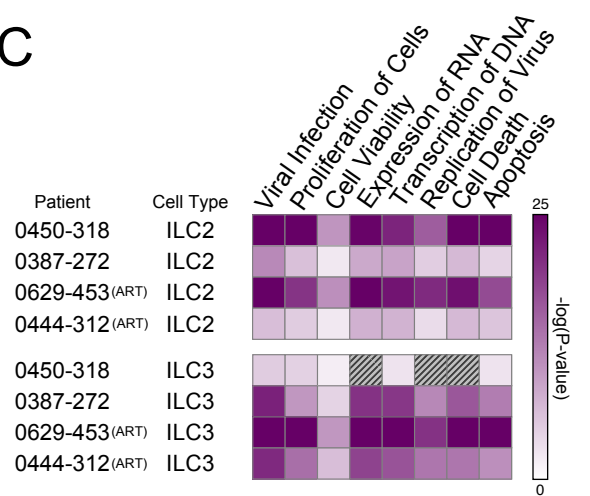
**Figure S6 related to Figure 6. RNA-Seq of ILCs in early acute HIV infection reveals dynamic changes in gene expression. (A)** Viral load (black), ILC2 (red), and ILC3 (blue) cell population data for each patient. Horizontal lines denote time points in days used for detection, peak viremia, and 6 weeks after detection, respectively. **(B)** Heat map of *P*-values for functionally enriched gene sets differentially expressed between initial viral detection and peak viremia (see Figure 6B) and **(C)** between peak viremia and approximately 6 weeks after detection (see Figure 6C).

# Figure S7 related to Fig. 7
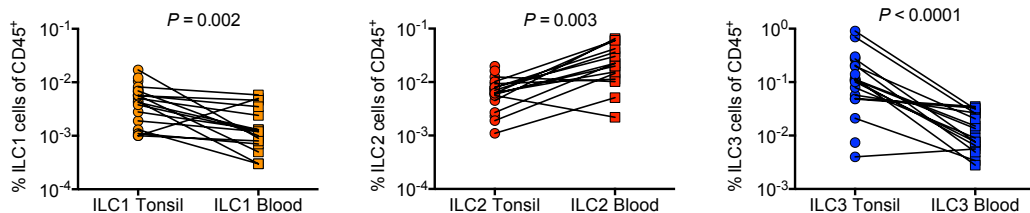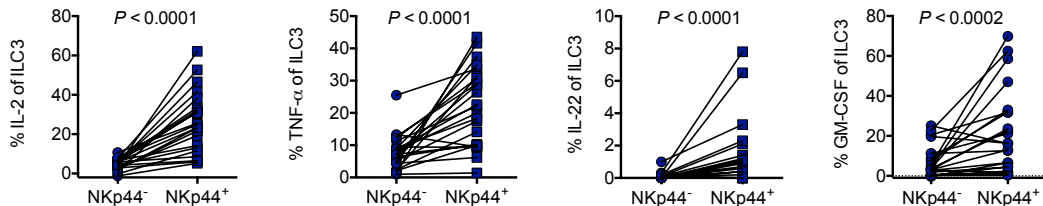
**Figure S7 related to Figure 7. ILC frequencies in blood and tonsil and intracellular cytokine production comparing NKp44⁻ and NKp44⁺ ILC3 subsets.** (A) Comparison of ILC1, ILC2 and ILC3 frequencies expressed as % of CD45⁺ lymphocyte cells obtained from matched tonsil and blood gated cells. (B) Intracellular cytokine production of IL-2, TNF-α, IL-22 and GM-CSF comparing NKp44⁻ and NKp44⁺ ILC3 gated tonsil cells. *P*-values by paired t-test.

**Immunity**

**Supplementary methods**

**IRREVERSIBLE DEPLETION OF INNATE LYMPHOID CELLS IN EARLY ACUTE HIV-1 INFECTION IN THE ABSENCE OF VIRAL SUPPRESSION**
**Kløverpris H et al**

**Study subjects:**

All subjects were treatment naïve for antiretroviral therapy unless otherwise indicated. The cohorts were followed from 4 independent studies; i) 'iThimba' Cohort from McCord Hospital, Durban, ii) the 'CAPRISA002' cohort iii) the 'GATEWAY' cohort (n=48) at Prince Mshiyeni Memorial Hospital, Umlazi. Durban and iv) the Females Rising through Empowerment, Support, and Health 'FRESH' from Umlazi. Durban(Ndhlovu et al., 2015). The median viral load was 14,000 copies/ml (interquartile range 640 - 79,302copies/ml). Absolute CD4+ T-cell counts were 397 per μl of blood (interquartile range 297 - 548 cells per μL). Samples from 'iThimba' and 'CAPRISA002' were processed from frozen PBMCs, whereas samples from the 'GATEWAY' and 'FRESH' cohorts were processed from fresh blood samples. All PBMCs were purified using standard ficoll separation.

Twenty-four individuals from the CAPRISA002 cohort were tested 3-monthly for presence of early detectable HIV-1 specific p24 antibodies and followed over 9 years with average treatment initiation starting at median values of 213 weeks after infection and sampled again at a median of 2 years into effective ART treatment at week 308. Sampling for each subject is shown for the last sample available before ART initiation (median week 213 after HIV-1 infection) and again 2 years into treatment (median week 308 after HIV-1 infection) The percentage change in absolute $CD4^{+}$ T cell counts comparing sampling at week 213 and week 308 was 90% and

therefore the median absolute reconstitution after 2 years of ART is almost a doubling of CD4$^{+}$ T cell counts compared to week 213 at ART initiation.

Tissue samples were obtained from surgical resection following tonsillectomy at King Edward Hospital, Durban, South Africa and from Stanger Hospital, KwaDukuza, KwaZulu-Natal, South Africa with a total of 12 HIV-1 infected and 12 HIV-1 uninfected individuals enrolled.

All subjects provided informed consent and each study was approved by the respective institutional review boards including the Biomedical Research Ethics Committee of the University of KwaZulu-Natal for all the studies.


**Flow cytometry**

Cells were surface stained with near-infrared live/dead cell viability staining kit (Invitrogen) and monoclonal antibodies: αCD45-V500 Horizon clone HI30 (BD Biosciences), αCD56 Brilliant Violet 711 clone HCD56 (BioLegend), αCD94 PerCP-Cy5.5 clone HP-3D9 (BD Biosciences), αCRTH2 PE-CF594 clone BM16, αCD127 PE-Cy7 clone R34.34 (Beckman Coulter), CD161 Brilliant Violet 605 clone HP-3G10 (BioLegend) and lineage markers conjugated to FITC or AlexaFlour488: αCD4 clone OKT4 (BioLegend), αCD11c clone 3.9 (BioLegend), αCD14 clone HCD14 (BD Biosciences), αCD19 clone 6D5 (BD Biosciences), αCD34 clone 561 (BioLegend), αFcER1 clone AER-37 (eBioscience), αBDCA2 clone 201A (BioLegend), αTCRαβ clone IP26 (BioLegend), αTCRγδ clone B1 (BioLegend) and intracellularly stained after Fix/Perm kit by (eBioscience) with αCD3 Brilliant Violet 785 OKT3 (BioLegend) and αCD3-FITC clone HIT3A (BD Biosciences), αGATA3-eFlour660 clone TWAJ (eBioscience), αT-bet Brilliant Violet 421 clone 4B10 (BioLegend), αEomes-PerCP-eFlour710 clone WD1928 (eBiosciences), αRORγt

clone Q21-559 (BD Biosciences), α-aryl hydrocarbon receptor (AHR) clone F3399 (eBioscience) and αHelios clone 22F6 (BioLegend). The eBioscience Fixation/Permeabilization kit from eBioscience were used for intracellular staining of transcritption factors and blocked with 20% goat serum for 20 mins prior to antibody staining.

For phenotype identification, cells were surface stained with near-infrared live/dead cell viability staining kit (Invitrogen) and monoclonal antibodies: αCD45-V500 Horizon clone HI30  (BD Biosciences), αCD56 Brilliant Violet 711 clone HCD56 (BioLegend), αCD94 PerCP-Cy5.5 clone HP-3D9 (BD Biosciences), αCRTH2 Alexa-Flour647 clone BM16 (BD Biosciences), αCD127 PE-Cy7 clone R34.34 (Beckman Coulter), αCD161 Brilliant Violet 605 clone HP-3G10 (BioLegend), αCD117-Brilliant Violet 421 clone 104D2 (BioLegend), αCD16 Brilliant Violet 650 clone 3G8 (BioLegend), αCD25 Brilliant Violet 785 clone BC96 (BioLegend), αCD69 Brilliant Ultra Violet 395 clone FN50 (Brilliant Horizon), αCCR6 Brilliant Ultra Violet 496 clone 11A9 (Brilliant Horizon), αNKp44 PE and PECy5 clone Z231 (Beckman Coulter), α4-β7 monoclonal antibody (cat#11718) (NIH) conjugated in house to PacificBlue (Life Technologies) and lineage markers conjugated to FITC or AlexaFlour488: αCD4 clone OKT4 (BioLegend), αCD11c clone 3.9 (BioLegend), αCD14 clone HCD14 (BD Biosciences), αCD19 clone 6D5 (BD Biosciences), αCD34 clone 561 (BioLegend), αFcER1 clone AER-37 (eBioscience), αBDCA2 clone 201A (BioLegend), αTCRαβ clone IP26 (BioLegend), αTCRγδ clone B1 (BioLegend),  and intracellularly stained after Fix/Perm kit by BD Biosciences αCD3-PE-CF594 clone UCHT1 (BD biosciences) and αCD3-FITC clone HIT3A (BD Biosciences). Intracellular cytokine production from all ILC subsets was

measured by PBMC stimulation for 6 hours with PMA/ionomycin or medium, with Golgiplug present. The following antibodies were used for intracellular cytokine stimulation (ICS); αTNFα Alexa700 Mab11 (BD Biosceinces), αIL13 V450 JES10-5A2 (BD Biosciences), IFNg Brilliant Violet 785 4S.B3 (BioLegend), αIL22 PerCP-eFlour710 22URTI (eBioscience), αIL17A Brilliant Violet 605 clone BL168 (BioLegend), αIL2 PE clone MQ1-7H12 (BD Bioscience).

**tSNE analysis on flow cytometry**

The R package 'Rtsne' available on CRAN (provided by Jesse Krijthe, github.com/jkrijthe/Rtsne) was used to perform the Barnes Hut implementation of tSNE on flow cytometry data. FlowJo software was used to export events of interest (in fcs format) for each sample analyzed in a manner similar to that described(Becher et al., 2014). After using the Bioconductor 'flowCore' R package to import .fcs file data and the Logicle transform(Parks et al., 2006) to scale the data similarly to that displayed in FlowJo (Top of scale = 500,000, width = 10), a similar number of events from each sample analyzed in parallel were merged and the relevant fluorescent parameters were used as input for tSNE. The results were unmerged into and appended to the corresponding input .fcs files for subsequent analysis in FlowJo. Clusters of cells identified by tSNE plots were gated manually and then further evaluated for median fluorescence intensities of each phenotypic marker using heatplots.

**RNAseq**

Briefly, 10 μL of mixed lysate from each sample was transferred to a skirted 96 well plate. Genetic material was pulled down and purified by mixing the lysate in each

well with 2.2x volumes of Agencourt RNAClean XP SPRI beads (Beckman Coulter) and washing thrice with 100 µL of 80% ethanol. After drying, the SPRI beads were re-suspended in 4 µL of pre-reverse transcription (RT) mix, incubated for 3 min at 72°C, and placed on ice. Next, Smart-Seq2 Whole Transcriptome Amplification (WTA) was performed: 7 µL of RT mix was added to each well and RT was carried out; then, 14 µL of PCR mix was added to each well and PCR was performed. After WTA, 0.8x volumes of Agencourt AMPure XP SPRI beads (Beckman Coulter) were used to clean up the cDNA product, which was then quantified using a Qubit dsDNA HS Assay Kit (Life Technologies). Library size and quality, meanwhile, were measured by Bioanalyzer using a High Sensitivity DNA Analysis Kit (Agilent Technologies).

**Gene Expression Data Analysis**

Sequencing data from the NextSeq500 was demultiplexed and aligned against hg19 using TopHat(Li and Dewey, 2011; Trapnell et al., 2014). Expression values, in counts, were generated in RSEM for every sample. Genes expressed in less than one third of the samples were discarded from future analyses as were samples with fewer than 2500 total counts or 1000 detected genes (genes with counts>0). A simple PCA of all samples was also conducted in order to identify any samples of low technical quality for removal.

EdgeR was used to calculate differential expression between CD4[+] T cells, ILC2 cells, and ILC3 cells isolated from 9 HIV negative FRESH cohort patients. First, expression counts were normalized and gene-specific dispersions were estimated. Global fitting was used to conduct likelihood ratio tests between pairs of cell types.

The top differentially expressed genes (P < 0.01, FDR < 0.01) between each of the following comparisons can be found in Table S1-S3: CD4 vs. ILC2, CD4 vs. ILC3, ILC2 vs. ILC3Additionally, to succinctly demonstrate differences between cell types over subsets of genes previously linked to ILCs, we adapted a previously described scoring system (Shalek et al, Nature, 2014). Here, we used human orthologs of genes found to be differentially expressed between various ILC populations in mice(Robinette et al., 2015). Orthologs were identified biomaRt (http://bioconductor.org/packages/release/bioc/html/biomaRt.html). Briefly, to generate each cell's Innate Lymphoid Cell Score, we transformed normalized expression counts for each gene into "induction values" by dividing by average expression and summing the induction values for each gene in our list for each sample. Finally, scores across all samples were linearly scaled to fit between (0, 1).

Differential expression analysis was conducted using edgeR, available on Bioconductor, via generalized linear modeling(Robinson and Oshlack, 2010). Lacking biological replicates in patient samples, dispersion estimates were calculated from 9 HIV uninfected patient samples. Tag-wise dispersion estimates were generated following trended dispersion estimates calculated using a log-linear trend. For each patient, Likelihood Ratio Tests (LRTs) were conducted for each cell type between both peak viremia and detection, and peak viremia and the time point closest to 6 weeks after infection (Figure 6 Figure S6A). These LRTs generated a log fold change (logFC) and P-value for each gene. All genes with P<0.01 were then uploaded into Ingenuity Pathway Analysis (IPA, Qiagen) with their corresponding logFCs. Core analyses were conducted on each cell type for each patient for each LRT test; the significant functionally enriched gene sets, upstream regulators, and identified

signaling pathways were then compared across patients and cell types. A list of all gene sets identified by IPA for each core analysis for the detection vs. peak viremia and peak viremia vs. 6 weeks comparisons is provided in Table S4 and S5, respectively.