Complement receptor 2 and IL-8 production identifies in adults and neonates naïve T cells recently arising from the thymus

Marcin L Pekalski[1], Arcadio Rubio García[1], Ricardo C Ferreira[1], Daniel B Rainbow[1], Deborah J Smyth[1], Meghavi Mashar[1], Jane Brady[1], Natalia Savinykh[1], Xaquin Castro Dopico[1], Sumiyya Mahmood[1], Simon Duley[1], Helen E Stevens[1], Neil M Walker[1], Antony J Cutler[1], Frank Waldron-Lynch[1], David B Dunger[2], Claire Shannon-Lowe[3], Alasdair J Coles[4], Joanne L Jones[4], Chris Wallace[1,5], John A Todd*[1], Linda S Wicker*[1]

[1]JDRF/Wellcome Trust Diabetes and Inflammation Laboratory
Department of Medical Genetics
NIHR Cambridge Biomedical Research Centre
Cambridge Institute for Medical Research
Cambridge Biomedical Campus, University of Cambridge
Cambridge, UK

[2]Department of Paediatrics, MRL Wellcome Trust-MRC Institute of Metabolic Science, NIHR Cambridge Comprehensive Biomedical Research Centre, University of Cambridge, Cambridge, UK

[3]Institute for Immunology and Immunotherapy and Centre for Human Virology, The University of Birmingham, Birmingham, UK

[4]Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK

[5]Department of Medicine, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK, and, MRC Biostatistics Unit, Cambridge Institute of Public Health, Cambridge Biomedical Campus, Cambridge, UK

Address correspondence to M.L.P. (marcin.pekalski@cimr.cam.ac.uk) or L.S.W. (linda.wicker@cimr.cam.ac.uk)
*These authors contributed equally.

The maintenance of a diverse, naïve T cell repertoire arising from the thymus (recent thymic emigrants: RTEs) is critical for health[1]. Recent studies have reported a unique naïve $CD4^+$ T cell subset in neonates and early childhood characterized by IL-8 production[2,3]. Here we demonstrate that IL-8 production is a characteristic feature of RTEs in adults, children and neonates and that a hallmark of these cells is the expression of complement receptor 2 (CR2) and the preferential production of IL-8 after activation. Although decreasing in number with age due to thymic involution and homeostatic expansion of naïve $CD4^+$ T cell in humans[4], $CR2^+$ RTEs persist into old age, have the highest levels of T cell receptor excision circles (TRECs), co-express complement receptor 1 (CR1), TLR1 and produce IL-8 upon TCR stimulation. We have observed these phenotypes in the vast majority of cord blood naïve $CD4^+$ T cells and in newly-generated naïve T cells appearing during reconstitution of the immune system in adults depleted of T cells through treatment for multiple sclerosis. A memory subset of $CR2^+$ $CD4^+$ T cells expressing high levels of CR1 and producing IL-8 following activation was also discovered, consistent with the hypothesis that RTE-specific gene expression confers a functional competence retained by particular memory T cells possibly because of their complement-dependent reactivity to pathogens. We suggest that assessing CR2 expression on naïve $CD4^+$ T cells will give a measure of thymic function during aging of the immune system and in a number of clinical situations including bone marrow transplantation[5], HIV infection[6], and immune reconstitution following immune depletion[7] or chemotherapy[8].

Thymic involution decreases naïve $CD4^+$ T cell production with age in humans and is compensated for by the homeostatic expansion of naïve cells that have emigrated from the thymus earlier in life[1,4]. Naïve T cells that have undergone decades of homeostatic expansion show reduced T cell receptor diversity that potentially negatively impacts host defence[9]. CD31 (PECAM-1) expression identifies cells that have divided more often in the periphery ($CD31^-$) from those that have not ($CD31^+$), although $CD31^+$ T cells still divide with age as evidenced by the dilution of TRECs[10] and, therefore, better markers of undivided naïve T cells are clearly needed[11]. As we and others have reported[12,13], CD25 is an additional marker of naïve T cells that have expanded in the periphery. $CD31^+$ $CD25^-$ naïve $CD4^+$ T cells predominate at birth and contain the highest content of TRECs as compared to their $CD31^-$ and $CD25^+$ counterparts[12].

Here, we assessed the heterogeneity of four naïve $CD4^+$ T cell subsets defined by CD31 and CD25 expression in neonates and adults in a population study of 391 donors (**Fig. 1a, Supplementary Fig. 1a**). $CD31^+$ $CD25^-$ naïve $CD4^+$ T cells decreased with age and this decrease was compensated for by

the homeostatic expansion of three subsets of naïve T cells: $CD31^+$ $CD25^+$, $CD31^-$ $CD25^-$ and $CD31^-$ $CD25^+$. As expected[14], the proportion of both naïve $CD4^+$ and $CD8^+$ T cells declined with age (**Supplementary Fig. 1b**). To define markers associated with the least expanded naïve subset, we performed statistically powered, genome-wide RNA analysis of FACS-purified naïve $CD4^+$ T cells from 20 adults sorted into four subsets based on CD31 and CD25 expression (**Supplementary Fig. 1c**). Principal component analysis of differentially expressed genes amongst the four subsets showed a clear separation between the groups, particularly between $CD31^+$ and their $CD31^-$ counterparts (**Supplementary Fig. 1d**). Genes upregulated in $CD31^+$ $CD25^-$ naïve cells as compared to the $CD31^-$ $CD25^-$ subset included two genes not normally associated with T cells: *AOAH*, which encodes the enzyme acyloxyacyl hydrolase that inactivates LPS and is highly expressed in innate immune cells[15], and *CR2*, which encodes a cell surface protein that binds C3d and other complement components[16] and is also a receptor for EBV in humans[17] (**Fig. 1b**) (**Supplementary Tables 1A-E)**. CR2 is expressed by follicular dendritic cells and B cells, and in the latter case, lowers their activation threshold when complement and antigen are both present[18]. CR2 expression has also been reported on $CD8^+$ and $CD4^+$ foetal T cells[19] and thymocytes[20]. Genes upregulated in expanded $CD31^-$ $CD25^-$ cells as compared to $CD31^+$ $CD25^-$ cells are consistent with the occurrence of activation and differentiation events during the homeostatic proliferation of this naïve subset and include a gene encoding an interferon-induced intracellular DNA receptor *PYHIN1*, *CTLA4* and *IL2RB*, as well as transcription factors such as *IRF4*, *PRDM1* (encoding BLIMP-1) and *MAF*. Expression of two genes, *TOX* (a transcription factor reported to regulate T-cell development in the thymus[21]) and *CACHD1*, an uncharacterized gene that may encode a protein that regulates voltage-dependent calcium channels, was downregulated when $CD31^+$ $CD25^-$ cells either lost expression of CD31 (**Fig. 1b**) or upregulated CD25 (**Supplementary Fig. 1e**).

CR2 expression analysed at the protein level verified the microarray results with the $CD31^+$ $CD25^-$ naïve $CD4^+$ T cell subset having the highest proportion of cells positive for CR2, with the proportion decreasing with age (**Fig. 1c, Supplementary Fig. 2a**). The reduction of $CR2^+$ naïve cells by age was more evident when considered out of total $CD4^+$ T cells (**Supplementary Fig. 2b**). The $CR2^+$ fraction of the $CD31^+$ $CD25^-$ naïve $CD4^+$ subset has the highest CR2 density. A similar pattern of decreasing levels of CR2 with age was also observed on naïve $CD8^+$ T cells (**Supplementary Fig. 2c**). The proportion and density of CR2 expression was greatly reduced in the three naïve $CD4^+$ populations previously shown to have a reduced TREC content when compared to $CD31^+$ $CD25^-$ naïve $CD4^+$ cells[12].

To determine whether CR2 is a molecular marker of $CD31^+$ $CD25^-$ naïve $CD4^+$ T cells that have proliferated the least in the periphery since emigrating from the thymus, we sorted $CR2^{hi}$, $CR2^{low}$ and $CR2^-$ $CD31^+$ $CD25^-$ naïve $CD4^+$ T cells from four adult donors (**Fig. 1d**) and $CR2^+$ and $CR2^-$ cells from three additional adult donors (**Supplementary Fig. 2d**) and assessed TREC levels. Sorted $CD31^-$ naïve $CD4^+$ T cells were used as a comparator (**Fig. 1d**) since the loss of CD31 expression from naïve cells is associated with a substantial loss of TRECs[4,10,12]. $CR2^+$ cells had more TRECs than $CR2^-$ cells in all cases. Where cell numbers were sufficient to separate the $CR2^{hi}$ and $CR2^{low}$ naïve $CD4^+$ T cells, the $CR2^{low}$ cells had undergone one or two more cell divisions than the $CR2^{hi}$ cells indicating that high CR2 expression on naïve $CD4^+$ T cells identifies cells that have divided the least since leaving the thymus (**Fig. 1d**). These observations explain why naïve $CD4^+$ T cells isolated only according to CD31 expression show an age-dependent loss of TRECs[4,10].

To further test the hypothesis that CR2 expression on naïve T cells defines RTEs throughout life rather than being specific to cells generated during the neonatal period, we monitored newly generated naïve $CD4^+$ T cells in eight multiple sclerosis patients depleted of T and B lymphocytes using a monoclonal antibody specific for CD52 (alemtuzumab)[7] (**Fig. 2, Supplementary Fig. 3a**). Twelve months after depletion, all eight patients had more naïve $CD4^+$ (**Fig. 2a**) and $CD8^+$ (**Supplementary Fig. 3b**) T cells expressing CR2 as compared to baseline demonstrating that *de novo* RTEs produced in the adult thymus are also defined by CR2 expression. Interim time points were available from most patients (**Fig. 2b, Supplementary Fig. 3c**) showing that when the first few naïve $CD4^+$ T cells were detected after depletion (3-9 months post-treatment), they were essentially all $CR2^+$ with a density of CR2 equalling that seen in cord blood (**Supplementary Fig. 2a**). This observation was independent of whether a patient had good or poor naïve $CD4^+$ T cell reconstitution overall.

A potential utility of CR2 is, therefore, to assess the functionality of the human thymus. Along these lines we compared the frequency of $CR2^+$ cells within the $CD31^+$ $CD25^-$ naïve $CD4^+$ T cell subset prior to lymphocyte depletion with the ability of the thymus to reconstitute the naïve $CD4^+$ T cell compartment, as a proportion of total $CD4^+$ T cells. The two patients that failed to reconstitute their naïve T cell pool to pre-treatment levels by 12 months had the lowest levels of $CR2^+$ T cells within the $CD31^+$ $CD25^-$ naïve $CD4^+$ T cell subset (18% and 17%) prior to treatment (**Fig. 2c, Supplementary Fig. 3c**). In contrast patients who reconstituted their naïve $CD4^+$ T cell pool to or above baseline by 12 months had on average 37.5% (range 29-54%) $CR2^+$ cells in the $CD31^+$ $CD25^-$ naïve $CD4^+$ T cell subset prior to treatment.

To evaluate the potential function of $CR2^+$ naïve $CD4^+$ cells, we compared RNA isolated from sorted $CR2^+$ and $CR2^-$ $CD31^+$ $CD25^-$ naïve $CD4^+$ T cells *ex vivo* and after activation (**Fig. 3a**, **Supplementary Tables 2-4**). This strategy revealed a transcriptional signature that reinforced the uniqueness of the $CR2^+$ RTEs. *Complement receptor 1* (*CR1*) and *TLR1* were more highly expressed in $CR2^+$ cells, and co-expression at the protein level for CR1 and CR2 was observed in $CD31^+$ $CD25^-$ naïve $CD4^+$ T cells from healthy controls (**Fig. 3b**) and MS patients reconstituting their naïve cells (**Supplementary Fig. 4a)**. Co-expression of CR1 and CR2 was also observed on naïve CD8 T cells (**Supplementary Fig. 4b)**.

Following activation, *IL8* was more highly expressed in $CR2^+$ cells whereas *IL2*, *IL21, LIF* and *IFNG* were more highly expressed in $CR2^-$ cells. *TNF, LTA* and *IL23A* were highly upregulated with activation but there was no difference between the subsets. *IL8* upregulation was of particular interest since it has been termed a phenotype of neonatal naïve cells[2]. We therefore measured IL-8 and IL-2 protein production from sorted $CR2^+$ versus $CR2^-$ cells after activation and verified the RNA results showing that IL-8 is preferentially produced by the $CR2^+$ subset whereas the opposite is the case for IL-2 (**Fig. 3c**). Because CR2 rapidly disappears from the surface of cells activated *in vitro*, it prevents the analysis of cytokine production using this marker after stimulation. Therefore, we examined IL-8 production in $CD4^+$ T cells stratified by CD45RA expression from cord blood, MS patients reconstituting their naïve T cell compartment (<1 year after treatment) and MS patients in whom naïve T cells had undergone homeostatic expansion (>10 years after treatment) (**Fig. 3d**). Notably all MS patients with high levels of RTEs that were tested (N=4) had a high proportion of their naïve $CD4^+$ T cells producing IL-8 at similar levels as that of cord blood naïve $CD4^+$ T cells. On the other hand, patients with naïve T cells that had expanded (N=3) produced much less IL-8. IL-8 production by naïve $CD4^+$ T cells positively correlated with their CR2 expression confirming the ability of $CR2^+$ RTE to produce IL-8. Our results are compatible with those from van den Broek *et al*. showing reduced IL-8 production in children following neonatal thymectomy[3].

When analysing IL-8 production by $CD4^+$ T cells we noted that in healthy controls (**Fig.3c**) and in MS patients who were more than 10 years past lymphocyte depletion (**Fig.3d**), a fraction of activated $CD45RA^-$ memory $CD4^+$ T cells produced IL-8. We therefore hypothesized that memory T cells can be generated from IL-8 producing $CR2^+$ naïve T cells. CR2 expression was observed on a proportion of central and effector memory $CD4^+$ T cells and Tregs expressed the lowest levels of CR2 (**Fig. 4a, Supplementary Fig. 5a**). CR2 expression by memory cells was correlated with CR2 expression by naïve T cells (**Fig. 4b**) and was age dependent (**Fig. 4c**) suggesting that with age the $CR2^+$ memory

cells either lose CR2 expression or contract due to competition with other CR2$^-$ memory cells. CR2$^+$ memory CD8$^+$ T cells were also observed and were similarly age dependent (**Supplementary Fig. 5b**). As seen with the equivalent CD4$^+$ naïve T cell subsets, RNA analysis of sorted CD4$^+$ CR2$^+$ and CR2$^-$ central memory T cells (**Supplementary Table 6**) showed CR1 to be the most differentially expressed gene (**Fig. 4d**), a phenotype confirmed at the protein level (**Fig. 4e**). CR2$^+$ memory cells produced higher levels of IL-8 after activation as compared to CR2$^-$ memory cells, and unlike naïve CR2$^+$ cells (**Fig. 3c**), all memory cells producing IL-8 produced IL-2 (**Fig. 4f**). Amongst the differentially expressed genes between the CR2$^-$ and CR2$^+$ memory cells, a gene of particular note is *complement factor H* (*CFH*) (**Fig. 4d**), which is upregulated in both CR2$^-$ and CR2$^+$ memory cells as compared to naïve cells but has 2.7-fold higher levels in CR2$^+$ memory cells compared to CR2$^-$ memory cells. Shared expression between CR2$^+$ naïve and CR2$^+$ memory T cells *ex vivo* are highlighted in **Supplemental Fig. 6** and include a potentially relevant transcription factor for CR2$^+$ CD4$^+$ cells, *ZNF462,* which is expressed 13.0-fold higher on CR2$^+$ versus CR2$^-$ naïve cells and 7.2-fold higher on CR2$^+$ versus CR2$^-$ memory cells. Four genes, *ADAM23, ARHGAP32, DST* and *PLXNA4*, shared by the two CR2$^+$ subsets may enhance migratory properties that augment host surveillance[22].

Our results demonstrate that relatively recent development in the thymus, not absolute age, confers a unique innate phenotype to naïve CD4$^+$ T cells that includes expression of complement receptors, TLR1 and a group of genes, including those encoding the transcription factors HELIOS (encoded by *IKZF2*) and ZNF462, that appear to program the RTEs to preferentially secrete IL-8 and have reduced production of other cytokines following activation (**Supplementary Table 5**). We propose that the presence of CR1 and CR2 in naïve CD4$^+$ as well as CD8$^+$ T cells contribute to host defence (**Supplementary Fig. 6**), possibly in a manner similar to that provided by these same molecules on B and follicular dendritic cells[18]. Aspects of the innate signature of RTEs are retained by a subset of memory T cells that express both CR1 and CR2 and secrete IL-8 upon activation, suggesting there is selection because of these functional attributes. Gibbons and colleagues[2] have characterized IL-8 as a "proinflammatory immunoprotective cytokine of neonatal T cells" that recruits neutrophils and co-stimulates $\gamma\delta$ T cells to produce IFN-γ, whereas we propose that this description should be expanded to include all CR2-expressing naïve CD4$^+$ RTEs and a subset of memory CD4$^+$ T cells. Our findings complement those of Liszewski *et al*. who have shown that human T cells express and process C3 to components that can bind CR2[23]. Furthermore, a complement–regulatory protein, CD46, influences cytokine production from human CD4$^+$ T cells[24].

As humans age, we have shown that with homeostatic expansion naïve T cells differentiate into naïve cells that have reduced expression of CR1 and CR2 and no longer produce IL-8 with activation. These attributes could contribute to the reduced host defence seen in some older individuals[1], although there is marked heterogeneity at the rate of loss of RTEs amongst people. Further analyses of T cells that secrete IL-8 and the functional effects of EBV, which binds to CR2[17], on $CR2^+$ T cells are ongoing. We suggest that loss of CR2 expression on naïve cells with homeostatic expansion contributes to the increasing severity of EBV infection with age[25]. We also propose that the expression of CR2 on $CD31^+CD25^-$ naïve $CD4^+$ T cells defines cells that have divided the least in the periphery since emigrating from the thymus and that the abundance of this subset reflects the functional age of the thymus. The ability of CR2 to be a surrogate marker of TREC levels in naïve cells will be useful in a number of clinical settings where thymic function is assessed[5-8].

ACKNOWLEDGMENTS:

This research was supported by the Cambridge NIHR BRC Cell Phenotyping Hub. In particular, we wish to thank Chris Bowman and Anna Petrunkina Harrison for their advice and support in cell sorting. We thank Sarah Dawson, Pamela Clarke, Meeta Maisuria-Armer and Gillian Coleman for their help in processing blood samples and Jane Kennet and Katerina Anselmiova for coordinating and obtaining blood samples from the Investigating Genes and Phenotypes Associated with Type 1 Diabetes study. We thank Karen May for obtaining blood samples from MS patients and Thaleia Kalatha for determining sample availability from patients. We thank Emma Jones for the use of the NanoString Instrument and Sarah Howlett for editorial review of the manuscript. We gratefully acknowledge the participation of all NIHR Cambridge BioResource volunteers, and thank the NIHR Cambridge BioResource centre and staff for their contribution. We thank the National Institute for Health Research and NHS Blood and Transplant. We thank the NIHR/Wellcome Trust Clinical Research Facility. This work was funded by the JDRF (9-2011-253), the Wellcome Trust (091157) and the National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre. CW was supported by a Wellcome Trust grant (089989). The research leading to these results has received funding from the European Union's 7th Framework Programme (FP7/2007-2013) under grant agreement no.241447 (NAIMIT). The study was supported by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement 633964 (ImmunoAgeing). The Cambridge Institute for Medical Research (CIMR) is in receipt of a Wellcome Trust Strategic Award (100140).

AUTHOR CONTRIBUTIONS:

MLP, JAT, LSW co-designed the study, evaluated the results and co-wrote the paper. ARG and CW contributed to the design of the study, evaluated the results and edited the manuscript. CW performed statistical analysis of microarray data. ARG performed statistical analysis of NanoString and RNA-seq data. MLP, RCF, DR, DS, MM, JB, NS, XCD, SM, SD, AJC and CS-L performed experiments. HS coordinated sample collection and processing. NW coordinated sample and data management. AJColes and JLJ co-designed and co-supervised the alemtuzumab study and edited the manuscript. DBD, LSW and JAT co-designed the Diabetes—Genes, Autoimmunity, and Prevention study. FW-L helped organise the provision of the blood samples in the Investigating Genes and Phenotypes Associated with Type 1 Diabetes study. JAT led the Investigating Genes and Phenotypes Associated with Type 1 Diabetes and Genes and Mechanisms of Type 1 Diabetes in the Cambridge BioResource studies.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

METHODS

Human samples. Donors of peripheral blood volunteered for one of three observational studies: Genes and Mechanisms in Type 1 Diabetes in the Cambridge BioResource (N=371, 114 males, 257 females, four of the female donors self-reported autoimmune disease—two with autoimmune thyroid disease, one with vitiligo and one with coeliac disease, donors recruited via the NIHR Cambridge BioResource), which was approved by NRES Committee East of England - Norfolk (ref: 05/Q0106/20); Diabetes—Genes, Autoimmunity, and Prevention, a study of newly diagnosed children with T1D and siblings of children with T1D (7 males aged 5 to 16, 8 females aged 1-14; all donors were autoantibody-negative siblings of patients with T1D who did not have T1D themselves); and Investigating Genes and Phenotypes Associated with Type 1 Diabetes (6 cord blood samples; 6 adults, none with self-reported autoimmunity). Diabetes—Genes, Autoimmunity, and Prevention was originally approved by the National Research Ethics Committee London – Hampstead and is now held under the ethics of Investigating Genes and Phenotypes Associated with Type 1 Diabetes, which was approved by NRES Committee East of England - Cambridge Central (ref: 08/H0308/153). NIHR Cambridge BioResource donors were collected with the prior approval of the National Health Service Cambridgeshire Research Ethics Committee. MS patients (6 females and 2 males, aged 27-49) were from the placebo arm of the trial “Keratinocyte Growth Factor - promoting thymic reconstitution and preventing autoimmunity after alemtuzumab (Campath-1H) treatment of multiple sclerosis”

(REC reference: 12/LO/0393, EudraCT number: 2011-005606-30). Additional MS patients treated with alemtuzumab greater than 10 years before analysis consented to long-term follow-up (CAMSAFE REC 11/33/0007).

Whole blood and PBMC immunostaining. Blood samples were directly immunophenotyped within 5 hours following donation. Samples were blocked for 10 min with mouse IgG (20 µg/ml), stained for 40 min at room temperature with appropriate antibodies and then lysed with freshly prepared 1X BD FACS Lysing Solution (BD Biosciences). After lysis of red blood cells, samples were washed with BD CellWASH (BD Biosciences). Finally, the samples were fixed with freshly prepared 1X BD CellFIX (BD Biosciences). The samples were stored at 4 °C in the dark until analysis using a BD Fortessa flow cytometer. PBMC samples, prepared as previously described[26], were blocked for 10 min, stained for 1 hour at 4°C, washed twice and fixed as described for peripheral blood immunophenotyping except for intracellular staining when surface-stained cells after the wash-step were placed in FOXP3 Fix/perm buffer (eBioscience). Phenotyping panels are detailed in **Supplemental Table 7**. CD25 detection sensitivity was increased[26] by simultaneous application of two anti-CD25 monoclonal antibodies labelled with the same fluorochrome (clones 2A3 and M-A251, BD Biosciences). Antibody concentrations used were based on the manufacturer's instructions as well as on optimization studies. Appropriate isotype controls and fluorescence-minus-one conditions were used during the development of staining panels.

Flow cytometry and data analysis. Immunostained samples were analysed on a BD LSRFortessa cell analyzer and data were visualized using Flowjo (TreeStar).

Cryopreserved PBMC. PBMC isolation, cryopreservation and thawing were performed as previously described[26]. Briefly PBMC isolation was carried out using Lympholyte (CEDARLANE). PBMCs were cryopreserved in heat-inactivated, filtered human AB serum (Sigma-Aldrich) and 10% DMSO (Hybri-MAX, Sigma-Aldrich) at a final concentration of 10×10^6/ml and were stored in liquid nitrogen. Cells were thawed in a 37 °C water bath for 2 min. PBMCs were subsequently washed by adding the cells to 10 ml of cold (4 °C) X-VIVO (Lonza) containing 10% AB serum per 10×10^6 cells, in a drop-wise fashion. PBMCs were then washed again with 10 ml of cold (4 °C) X-VIVO containing 1% AB serum per 10×10^6 cells.

T cell subset purification by cell sorting and DNA isolation. $CD4^+$ T cells (RosetteSep Human $CD4^+$ T Cell Enrichment Cocktail, STEMCELL Technologies) were washed and immediately incubated with

antibodies to surface molecules (**Supplementary Table 7**) for 40 minutes at 4 °C, washed and followed by sorting on a FACSAria Fusion flow cytometer cell sorter) into X-VIVO medium (Lonza) containing 5% human AB serum (Sigma). In order to isolate DNA, sorted cell subsets were checked for purity and DNA was isolated using a DNA extraction reagent (QIAGEN).

TREC assay. TREC assay was performed as described previously[12]. Briefly, a quantitative PCR assay was purchased from Sigma-Genosys for the signal joint TCR excision circle (sjTREC) that arises through an intermediate rearrangement in the TCRD/TCRA locus in developing TCR$\alpha\beta^+$ T lymphocytes. An assay for the gene encoding albumin was used to normalise the data.

sjTREC.F TCGTGAGAACGGTGAATGAAG

sjTREC.R CCATGCTGACACCTCTGGTT

sjTREC.P FAM-CACGGTGATGCATAGGCACCTGC-TAMRA

Alb.F GCTGTCATCTCTTGTGGGCTGT

Alb.R ACTCATGGGAGCTGCTGGTTC

Alb.P FAM-CCTGTCATGCCCACACAAATCTCTCC-TAMRA

For each sample, 24 ng of DNA was incubated in duplicate with both primers (700 nM), probe (150 nM) and 12.5 µl TaqMan mastermix (Applied Biosystems) and processed using the Applied Biosystems™ 7900HT Fast Real-Time PCR System. sjTREC data were calculated with application of the ΔΔCt method; sjTREC content for each subset was graphed as the proportion of sjTREC content present in the most naïve population sorted: CD31$^+$ CR2^{high} naïve CD4$^+$ T cells in **Fig. 1d** and CD31$^+$ CR2$^+$ naïve CD4$^+$ T cells in **Supplementary Fig. 2d**.

T cell activation. FACS-purified T cell subsets or total CD4$^+$ T cells (RosetteSep) were stimulated with either anti-CD3/CD28 beads (Life Technologies) at 3 cells per bead overnight or cell stimulation reagent (PMA and Ionomycin, eBioscience) in the presence of protein transport inhibitors (eBioscience) for six hours at 37 °C in 96-well U-bottom plates. IL-8$^+$/IL-2$^+$ T cells were identified with a staining panel shown in **Supplementary Table 7**.

Microarray gene expression analysis. Total RNA was prepared from cell subsets isolated by sorting using TRIzol reagent (Life Technologies). Single-stranded cDNA was synthesised from 200 ng of total RNA using the Ambion WT Expression kit (Ambion) according to the manufacturer's instructions. Labelled cDNA (GeneChip Terminal Labelling and Hybridization Kit, Affymetrix) was hybridized to a 96 Titan Affymetrix Human Gene 1.1 ST array.

Power calculations to determine the sample size required were performed using the method of Tibshirani[27], using a reference dataset from the Affymetrix GeneST array (deposited in ArrayExpress (http://www.ebi.ac.uk/arrayexpress/, accession number E-MTAB-4852) using the TibsPower package [http://github.com/chr1swallace/TibsPower] in R. We chose 20 pairs to have a false discovery rate (FDR) close to zero whilst detecting a 5-fold change in gene expression in 20 genes with a false negative rate of 5% or a 2-fold change in gene expression in 20 genes with a false negative rate of 40%, at a significance threshold of 10^{-6}.

Microarray gene expression log2 intensities were normalised using vsn2[28]. Analysis of differential expression (log2 intensities) was conducted pairwise between each cell subtype using paired t tests with limma[29]. P values were adjusted using the Benjamini-Hochberg algorithm. Illustrative principal component analysis was performed on the union of the most differentially expressed genes in each pairwise comparison, the list of which is available as a **Supplementary Table 1E**. Data are deposited with ArrayExpress, accession number E-MTAB-4853.

NanoString and RNA-seq: sample preparation and data analysis. See **Supplementary Fig. 7** for a description of the NanoString and RNA-seq experimental design. $CR2^+$ and $CR2^-$ naïve and memory cell subsets isolated by sorting $CD4^+$ T cells (RosetteSep) from four donors were pelleted directly or following activation and lysed in QIAGEN RLT buffer and frozen at -80 °C. To extract RNA, lysates were warmed to room temperature and vortexed. The RNA was extracted using Zymo Research Quick-RNA MicroPrep kit following the manufacturer's recommended protocol including on-column DNA digestion. RNA was eluted in 6 µl of RNase-free water. NanoString RNA expression analysis was performed using the Human Immunology v2 XT kit, 5 µl of RNA (5 ng/µl) was used per hybridisation and set up following the recommended XT protocol. Hybridisation times for all samples were between 16 and 20 hours. A NanoString Flex instrument was used and the Prep Station was run in high sensitivity mode and 555 fields of view were collected by the Digital Analyser. For RNA-seq analysis, 10 µl of RNA (8 ng/µl) were processed by AROS Applied Biotechnology using the Illumina TruSeq Access method that captures the coding transcriptome after library prep.

Raw NanoString expression measurements were normalized with application of NanoString software (nSolver 2.5). Subsequently, a paired differential expression analysis was carried out using DESeq2 v1.12.3[30], with preset size factors equal to 1 for all samples. Analyses were performed using a FDR of 0.05%. A missing FDR is reported for genes that were found to contain an expression outlier by

DESeq2 Cook's distance-based flagging of p-values. NanoString data are deposited with ArrayExpress, accession number E-MTAB-4854.

RNA sequencing yielded on average 35.9 million paired-end reads per library. Maximum likelihood transcript read count estimates for each sample were obtained with Kallisto v0.42.5[31], using Ensembl Release 82[32] as a reference transcriptome. Gene expression estimates were derived by aggregating all their constituent transcript read counts, which were then employed to perform a paired differential expression analysis using limma v3.28.5[29]. Analyses were performed using a FDR of 0.05%. A missing FDR is reported for genes that did not contain at least 2 counts per million (CPM) in at least 2 samples. Data from RNA-seq are deposited with European Nucleotide Archive, http://www.ebi.ac.uk/ena, accession number EGAS00001001870.

Statistical analysis of flow cytometry data interrogating T-cell subsets. Statistical analyses of the percentage of cells expressing CR2 and CR1 were performed and presented using Prism 5 software (Graphpad.com). Comparisons between cell subsets were performed using a paired Student's t-test. $P < 0.05$ was considered significant, error bars show the SD of the samples at each test condition. We included samples from 389 individuals in the analysis of flow cytometry data versus age. Regression was done using a non-parametric method, LOESS[33]. The grey zones define a 95% confidence interval for each regression line. Statistical tests were implemented using Prism software and R software (http://www.R-project.org).

1. Goronzy, J.J., Fang, F., Cavanagh, M.M., Qi, Q. & Weyand, C.M. Naive T cell maintenance and function in human aging. *J Immunol* **194**, 4073-4080 (2015).
2. Gibbons, D.*, et al.* Interleukin-8 (CXCL8) production is a signatory T cell effector function of human newborn infants. *Nat Med* **20**, 1206-1210 (2014).
3. van den Broek, T.*, et al.* Neonatal thymectomy reveals differentiation and plasticity within human naive T cells. *J Clin Invest* **126**, 1126-1136 (2016).
4. den Braber, I.*, et al.* Maintenance of peripheral naive T cells is sustained by thymus output in mice but not humans. *Immunity* **36**, 288-297 (2012).
5. Weinberg, K.*, et al.* Factors affecting thymic function after allogeneic hematopoietic stem cell transplantation. *Blood* **97**, 1458-1466 (2001).
6. Dion, M.L.*, et al.* HIV infection rapidly induces and maintains a substantial suppression of thymocyte proliferation. *Immunity* **21**, 757-768 (2004).
7. Jones, J.L.*, et al.* Human autoimmunity after lymphocyte depletion is caused by homeostatic T-cell proliferation. *Proc Natl Acad Sci U S A* **110**, 20200-20205 (2013).
8. Hakim, F.T.*, et al.* Age-dependent incidence, time course, and consequences of thymic renewal in adults. *J Clin Invest* **115**, 930-939 (2005).
9. van der Geest, K.S.*, et al.* Quantifying Distribution of Flow Cytometric TCR-Vbeta Usage with Economic Statistics. *PLoS One* **10**, e0125373 (2015).

10. Junge, S.*, et al.* Correlation between recent thymic emigrants and CD31+ (PECAM-1) CD4+ T cells in normal individuals during aging and in lymphopenic children. *Eur J Immunol* **37**, 3270-3280 (2007).
11. Zhang, S.L. & Bhandoola, A. Losing TREC with age. *Immunity* **36**, 163-165 (2012).
12. Pekalski, M.L.*, et al.* Postthymic expansion in human CD4 naive T cells defined by expression of functional high-affinity IL-2 receptors. *J Immunol* **190**, 2554-2566 (2013).
13. van der Geest, K.S.*, et al.* Low-affinity TCR engagement drives IL-2-dependent post-thymic maintenance of naive CD4+ T cells in aged humans. *Aging Cell* **14**, 744-753 (2015).
14. Wertheimer, A.M.*, et al.* Aging and cytomegalovirus infection differentially and jointly affect distinct circulating T cell subsets in humans. *J Immunol* **192**, 2143-2155 (2014).
15. Janelsins, B.M., Lu, M. & Datta, S.K. Altered inactivation of commensal LPS due to acyloxyacyl hydrolase deficiency in colonic dendritic cells impairs mucosal Th17 immunity. *Proc Natl Acad Sci U S A* **111**, 373-378 (2014).
16. Noris, M. & Remuzzi, G. Overview of complement activation and regulation. *Semin Nephrol* **33**, 479-492 (2013).
17. Young, K.A., Herbert, A.P., Barlow, P.N., Holers, V.M. & Hannan, J.P. Molecular basis of the interaction between complement receptor type 2 (CR2/CD21) and Epstein-Barr virus glycoprotein gp350. *J Virol* **82**, 11217-11227 (2008).
18. Carroll, M.C. & Isenman, D.E. Regulation of humoral immunity by complement. *Immunity* **37**, 199-207 (2012).
19. Thornton, C.A., Holloway, J.A. & Warner, J.O. Expression of CD21 and CD23 during human fetal development. *Pediatr Res* **52**, 245-250 (2002).
20. Watry, D.*, et al.* Infection of human thymocytes by Epstein-Barr virus. *J Exp Med* **173**, 971-980 (1991).
21. Wilkinson, B.*, et al.* TOX: an HMG box protein implicated in the regulation of thymocyte selection. *Nat Immunol* **3**, 272-280 (2002).
22. Masopust, D. & Schenkel, J.M. The integration of T cell migration, differentiation and function. *Nat Rev Immunol* **13**, 309-320 (2013).
23. Liszewski, M.K.*, et al.* Intracellular complement activation sustains T cell homeostasis and mediates effector differentiation. *Immunity* **39**, 1143-1157 (2013).
24. Cardone, J.*, et al.* Complement regulator CD46 temporally regulates cytokine production by conventional and unconventional T cells. *Nat Immunol* **11**, 862-871 (2010).
25. Cohen, J.I. Epstein-Barr virus infection. *N Engl J Med* **343**, 481-492 (2000).
26. Dendrou, C.A.*, et al.* Cell-specific protein phenotypes for the autoimmune locus *IL2RA* using a genotype-selectable human bioresource. *Nat Genet* **41**, 1011-1015 (2009).
27. Tibshirani, R. A simple method for assessing sample sizes in microarray experiments. *BMC Bioinformatics* **7**, 106 (2006).
28. Huber, W., von Heydebreck, A., Sueltmann, H., Poustka, A. & Vingron, M. Parameter estimation for the calibration and variance stabilization of microarray data. *Stat Appl Genet Mol Biol* **2**, Article3 (2003).
29. Ritchie, M.E.*, et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
30. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
31. Bray, N.L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**, 525-527 (2016).
32. Cunningham, F.*, et al.* Ensembl 2015. *Nucleic Acids Res* **43**, D662-669 (2015).
33. Cleveland, W.S. Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association* **74**, 829-836 (1979).

List of supplementary tables:

Supplementary Table 1: CD25_CD31 naïve T cells (requires downloading)
Microarray gene expression analysis of differentially expressed genes amongst four naïve $CD4^+$ T cell subsets defined by CD25 and CD31 expression.

Supplementary Table 2: CR2+ vs CR2- Naive_exvivo (requires downloading)
RNA expression analysis of differentially expressed genes in *ex vivo* $CR2^+$ versus $CR2^-$ naïve $CD4^+$ T cells. The file contains RNA data and analyses from NanoString and RNA-seq. See the first tab on the excel sheet for a description of the columns within the file.

Supplementary Table 3: CR2+ vs CR2- Naive_activated (requires downloading)
RNA expression analysis of differentially expressed genes in activated $CR2^+$ versus $CR2^-$ naïve $CD4^+$ T cells. The file contains RNA data and analyses from NanoString and RNA-seq. See the first tab on the excel sheet for a description of the columns within the file.

Supplementary Table 4: RNAseq_selected_examples_Naive and Memory_exvivo
Selected results of RNA-seq expression analysis experiment performed on *ex vivo* $CR2^+$ and $CR2^-$ naïve and memory $CD4^+$ T cells (full data in Supplementary Tables 2 and 6). .

Supplementary Table 5: NanoString_selected_examples_Naive_Activated
Selected results of NanoString RNA expression experiment performed on $CR2^+$ and $CR2^-$ naïve $CD4^+$ T cells after activation (full data in Supplementary Table 3).

Supplementary Table 6: CR2+ vs CR2- Memory_exvivo (requires downloading)
RNA expression analysis of differentially expressed genes in *ex vivo* $CR2^+$ versus $CR2^-$ memory $CD4^+$ T cells. The file contains RNA data and analyses from NanoString and RNA-seq. See the first tab on the excel sheet for a description of the columns within the file.

Supplementary Table 7: Antibody panels used for immunophenotyping and FACS-sorting

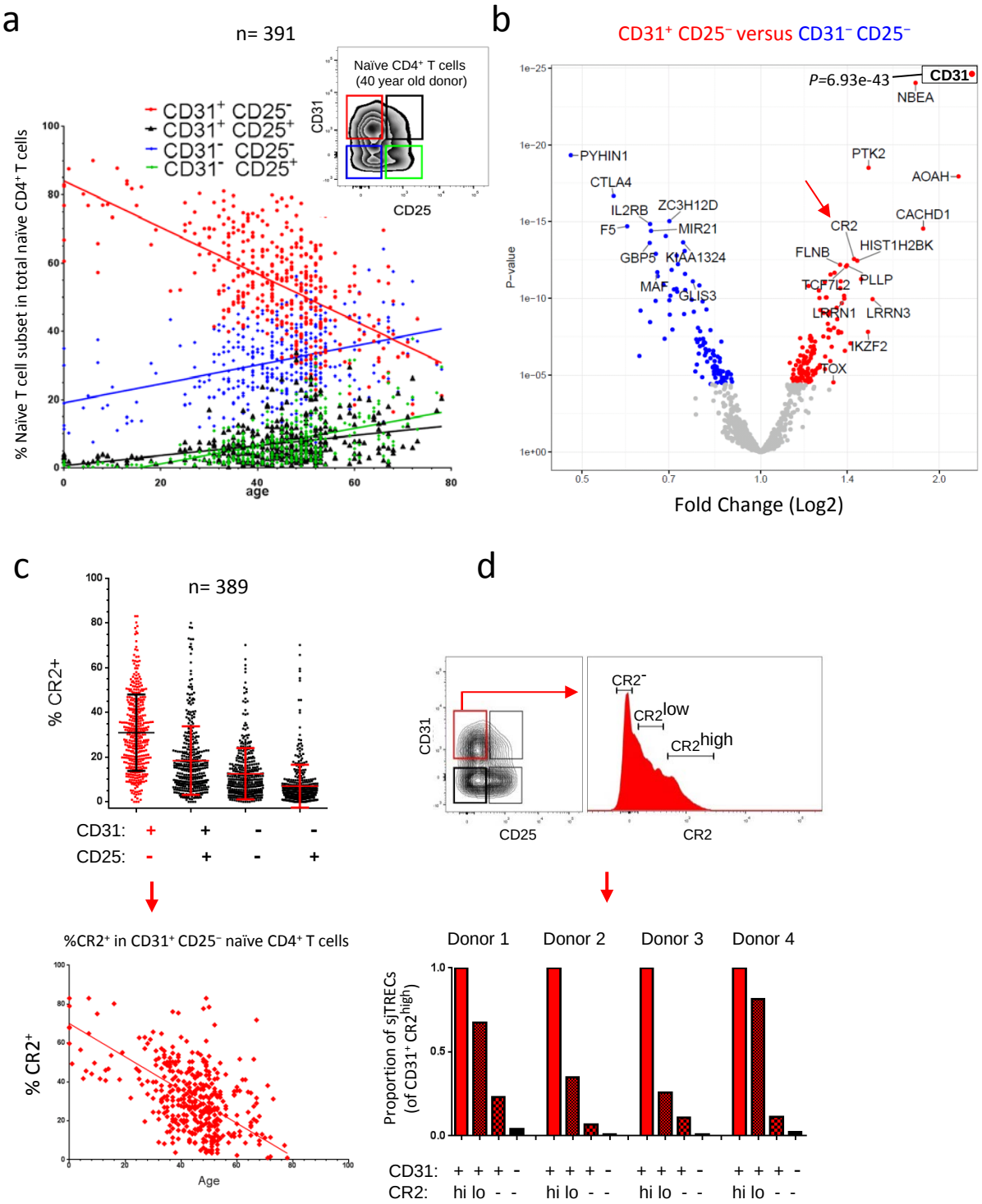
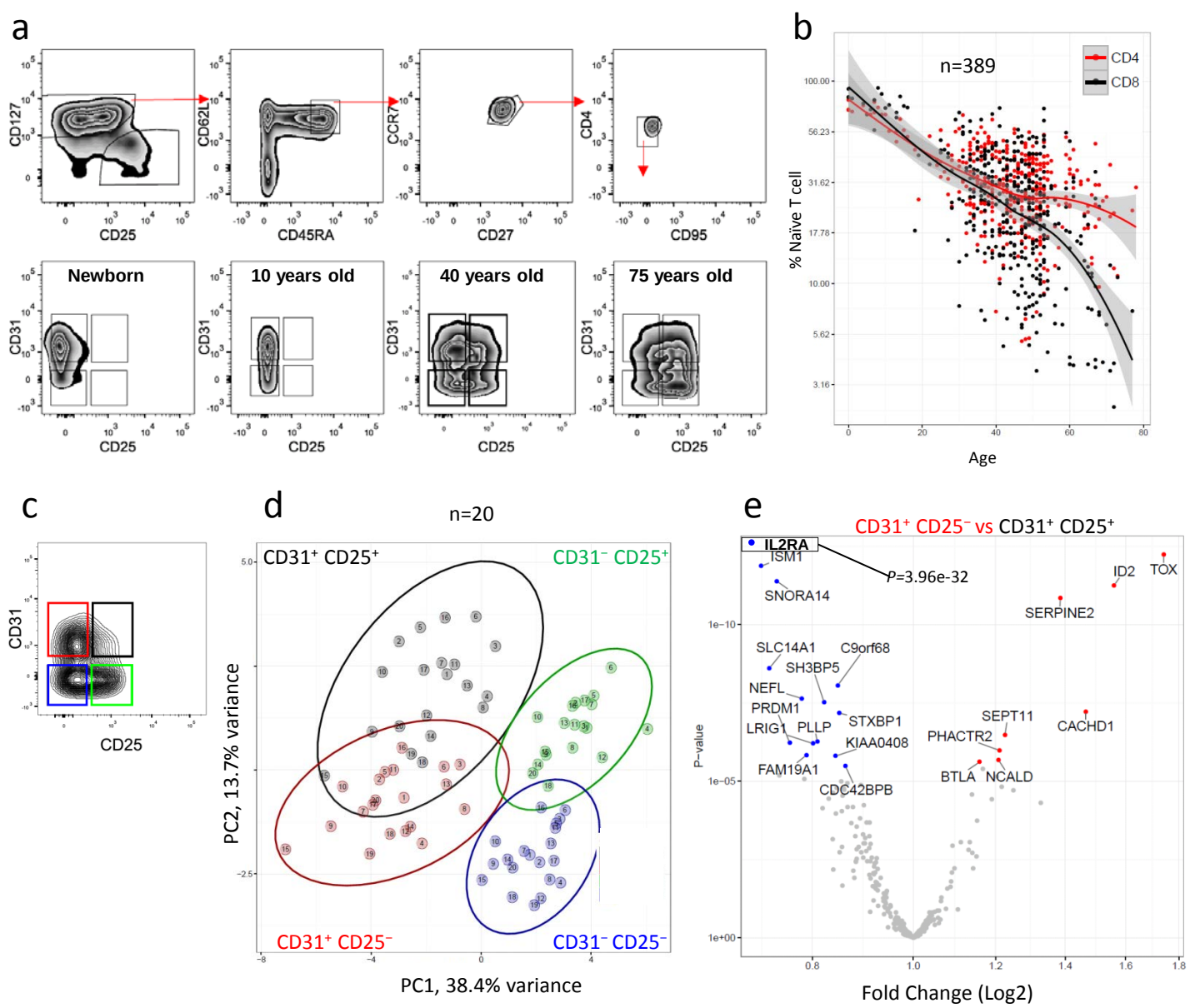
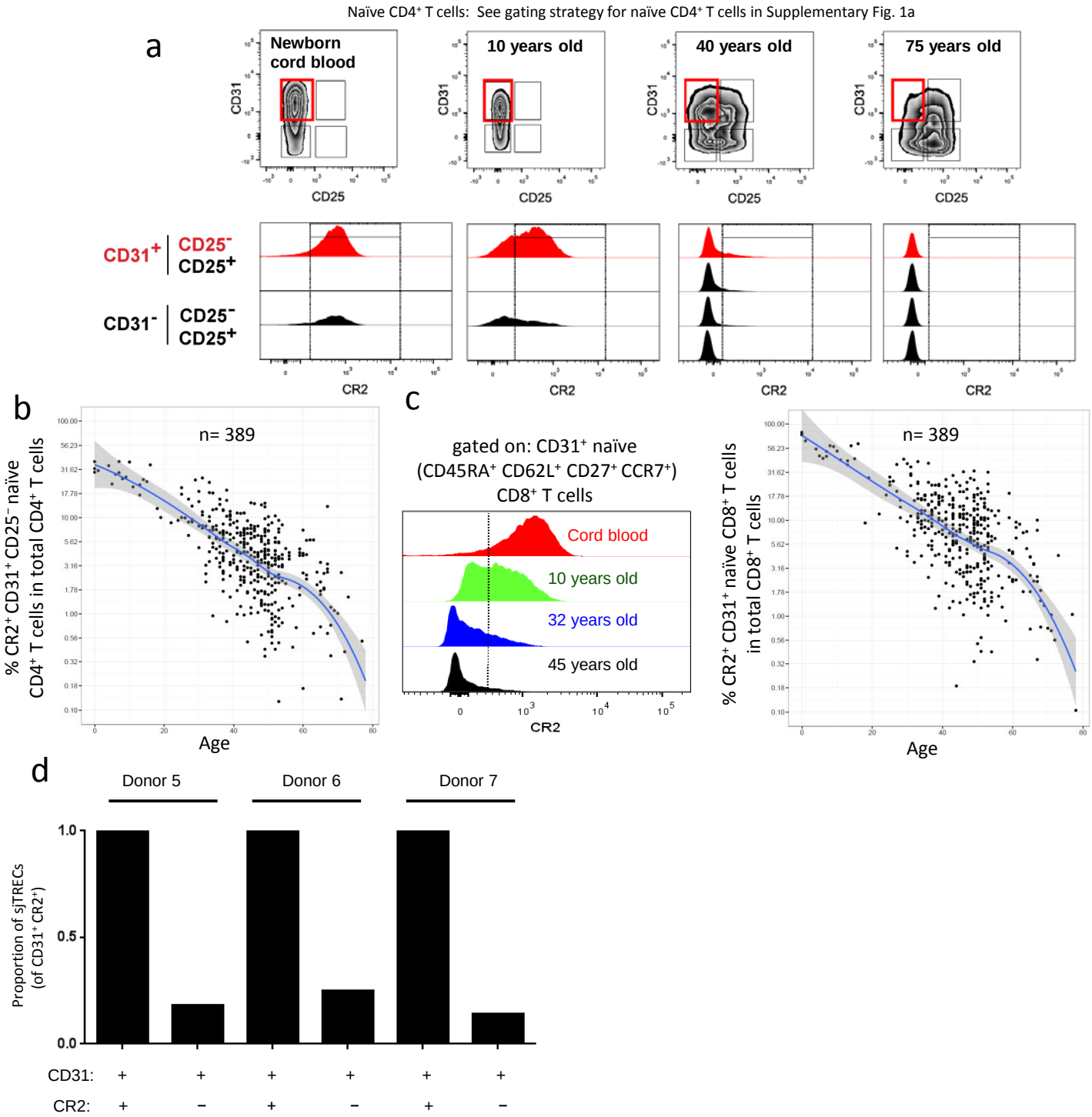


Figure 1 CR2 marks the most naïve CD4⁺ T cell subset. (**a**) The proportion of four subsets of naïve CD4⁺ T cells defined by CD31 and CD25 as a function of age (color coding shown above graph, full gating strategy in **Supplementary Fig. 1a**). (**b**) Volcano plot representing differences in gene expression between CD31⁺ CD25⁻ and CD31⁻ CD25⁻ naïve CD4⁺ T cells; genes with increased expression in CD31⁺ CD25⁻ (red) versus those increased in CD31⁻ CD25⁻ (blue) naïve CD4⁺ T cells. (**c**) Complement receptor 2 (CR2) expression by human naïve CD4⁺ T cells. Frequency (mean ± s.d.m.) of each naïve CD4⁺ T cell subset that is CR2⁺ (gating shown in **Supplementary Figure 2a**). Frequency of CD31⁺ CD25⁻ naïve CD4⁺ T cells that are CR2⁺ as a function of age. (**d**) Representative sorting strategy for CD31⁺ CD25⁻ naïve CD4⁺ T cells identified as CR2⁻, CR2^{low} and CR2^{high} that were then assessed for sjTRECs.



Supplementary Figure 1 Gene expression changes in human naïve CD4⁺ T cells that have undergone homeostatic expansion. (**a**) Gating strategy defining human naïve CD4⁺ T cells; naïve T cells were further stratified by their surface expression of CD31 and CD25. Representative examples of naïve CD4⁺ T cells from four donors of different ages from a total number of 389 analyzed. (**b**) Frequency of naïve CD4⁺ and CD8⁺ naïve T cells out of total CD4⁺ and CD8⁺ T cells, respectively, as a function of age. (**c**) FACS sorting strategy of naïve CD4⁺ T cells stratified by expression of CD31 and CD25. (**d**) Principal component analysis (PCA) based on differentially expressed genes amongst four naïve CD4⁺ T cell subsets stratified by surface expression of CD31 and CD25 (subsets sorted from 20 donors). (**e**) Volcano plot representing differences in gene expression between CD31⁺ CD25⁻ and CD31⁺ CD25⁺ naïve CD4⁺ T cells; genes increased in CD31⁺ CD25⁻ (red) versus those increased in CD31⁺ CD25⁺ (blue) naïve CD4⁺ T cells.



Supplementary Figure 2 CR2 expression by human CD4⁺ and CD8⁺ naïve T cells. (**a**) Representative examples (from a total of 389 donors) of CR2 profiles from four naïve T cell subsets stratified by CD31 and CD25 expression. (**b**) The frequency of CD31⁺ CD25⁻ naïve CD4⁺ T cells out of total CD4⁺ T cells that are CR2⁺ as a function of age. (**c**) Representative examples (from a total of 389 donors) of CR2 expression on naïve CD8⁺ T cells and the frequency of CD31⁺ naïve CD8⁺ T cells out of total CD8⁺ T cells that are CR2⁺ as a function of age (see **Supplementary Fig. 5b** for gating strategy of naïve CD8⁺ T cells). (**d**) sjTREC assessment of naïve CD4⁺ T cells identified as CR2⁺ and CR2⁻ out of the CD31⁺ CD25⁻ naïve CD4⁺ T cell subset.

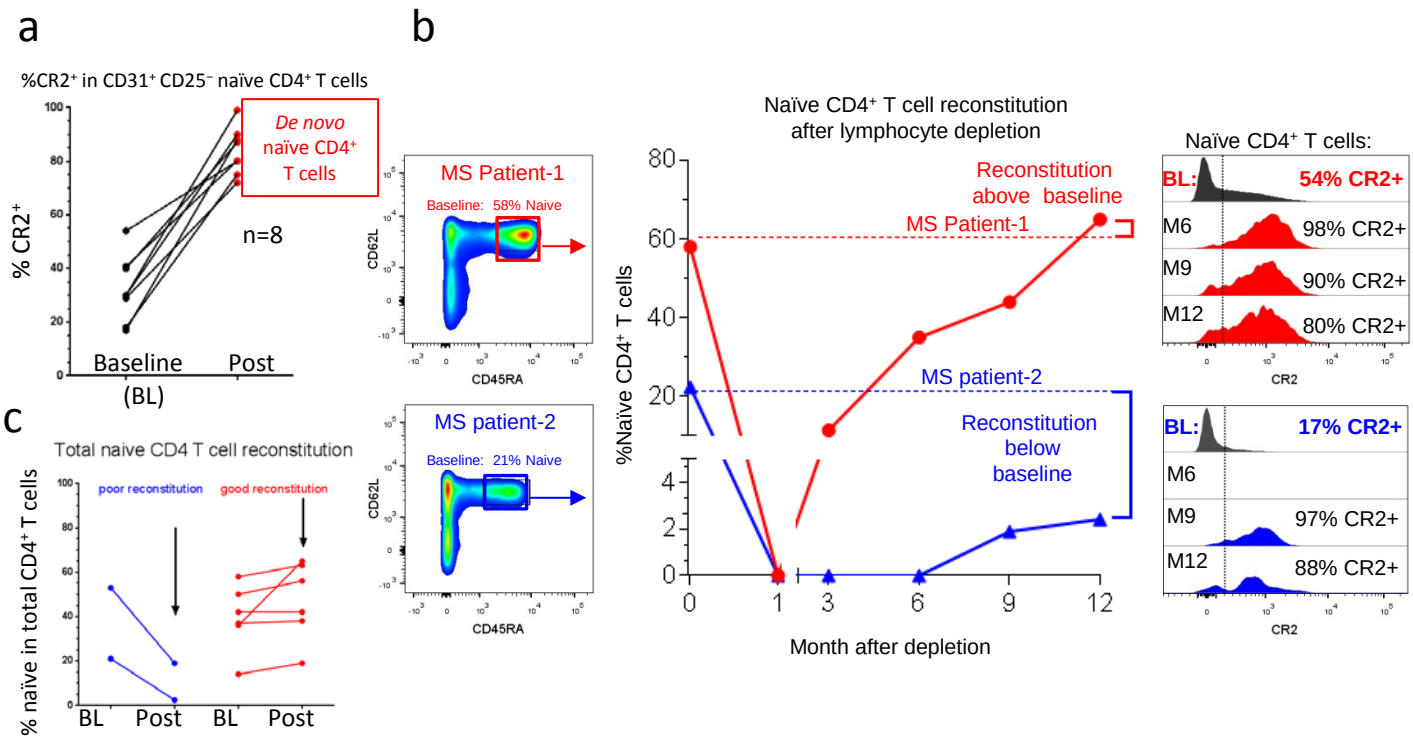
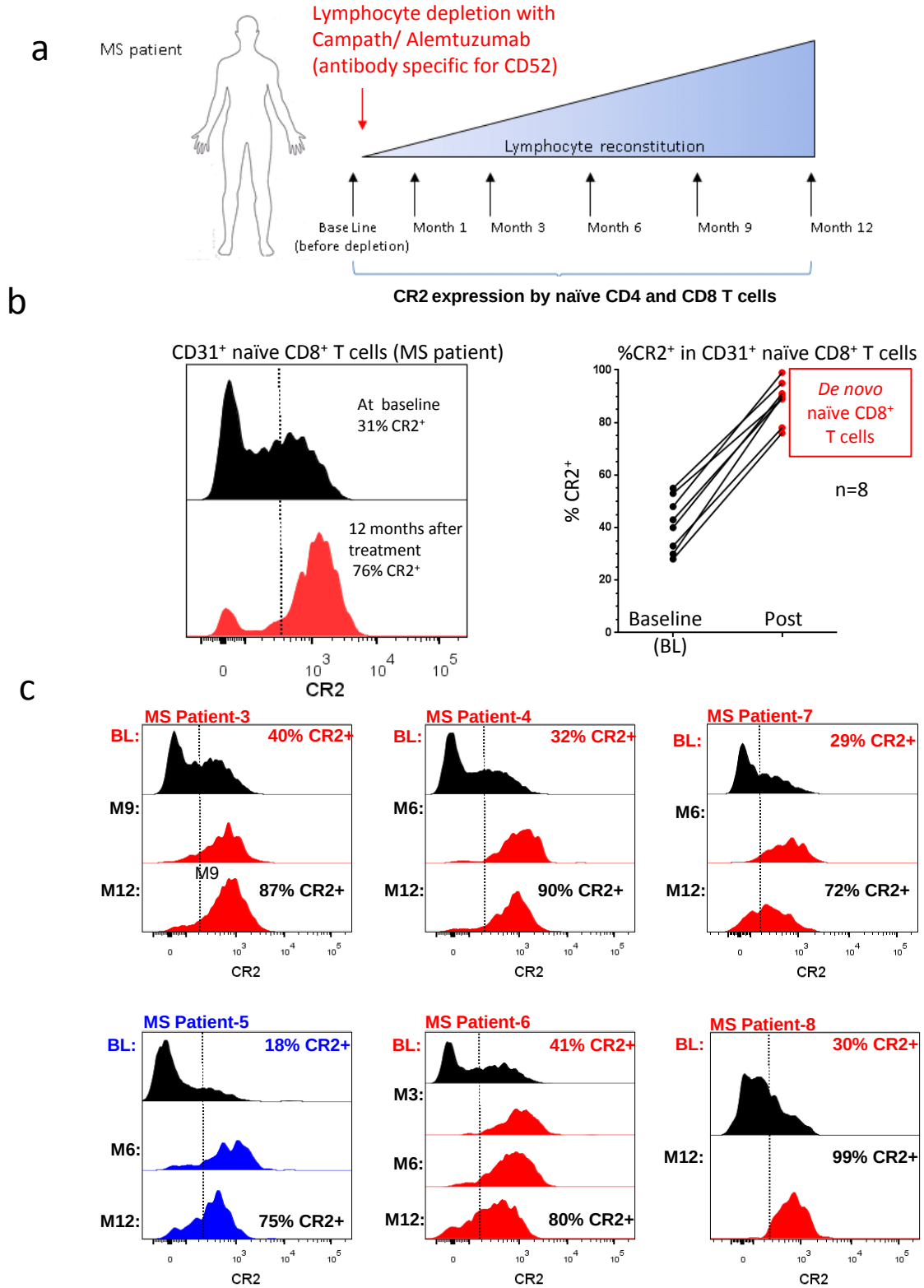


Figure 2 Complement receptor 2 (CR2) expression by human naïve CD4⁺ T cells during *de novo* reconstitution (a) Frequency of CD31⁺ CD25⁻ naïve CD4⁺ T cells expressing CR2 in MS patients before and 12 months after lymphocyte depletion with anti-CD52, Campath). (b) Representative examples of naïve CD4⁺ T cell reconstitution in two patients; frequency and expression level of CR2 on CD31⁺ CD25⁻ naïve CD4⁺ T cells before and during reconstitution. Six additional patients are shown in **Supplementary Figure 3b.** (c) Naïve CD4⁺ T cell frequency out of total CD4⁺ T cells before and 12 months after lymphocyte depletion.



Supplementary Figure 3 CR2 is highly expressed on naïve T cells that leave the adult thymus. (**a**) Treatment and sampling time points of MS patients. (**b**) Representative example and summary analysis (in eight MS patients) of CR2 expression on CD8$^+$ naïve T cells before (at baseline) and at 12 months after reconstitution. (**c**) CR2 expression profiles of CD31$^+$ CD25$^-$ naïve CD4$^+$ T cells from MS patients before and at various time points after lymphocyte depletion. Frequencies of CR2$^+$ cells in the CD31$^+$ CD25$^-$ naïve CD4$^+$ T cell subset are shown in the upper right hand corner of the baseline and month 12 histograms. Profiles depicted in red denote patients with good total naïve T cell reconstitution at 12 months; profiles depicted in blue represent patients with poor naïve T cell reconstitution at 12 months (see **Fig. 2c**).

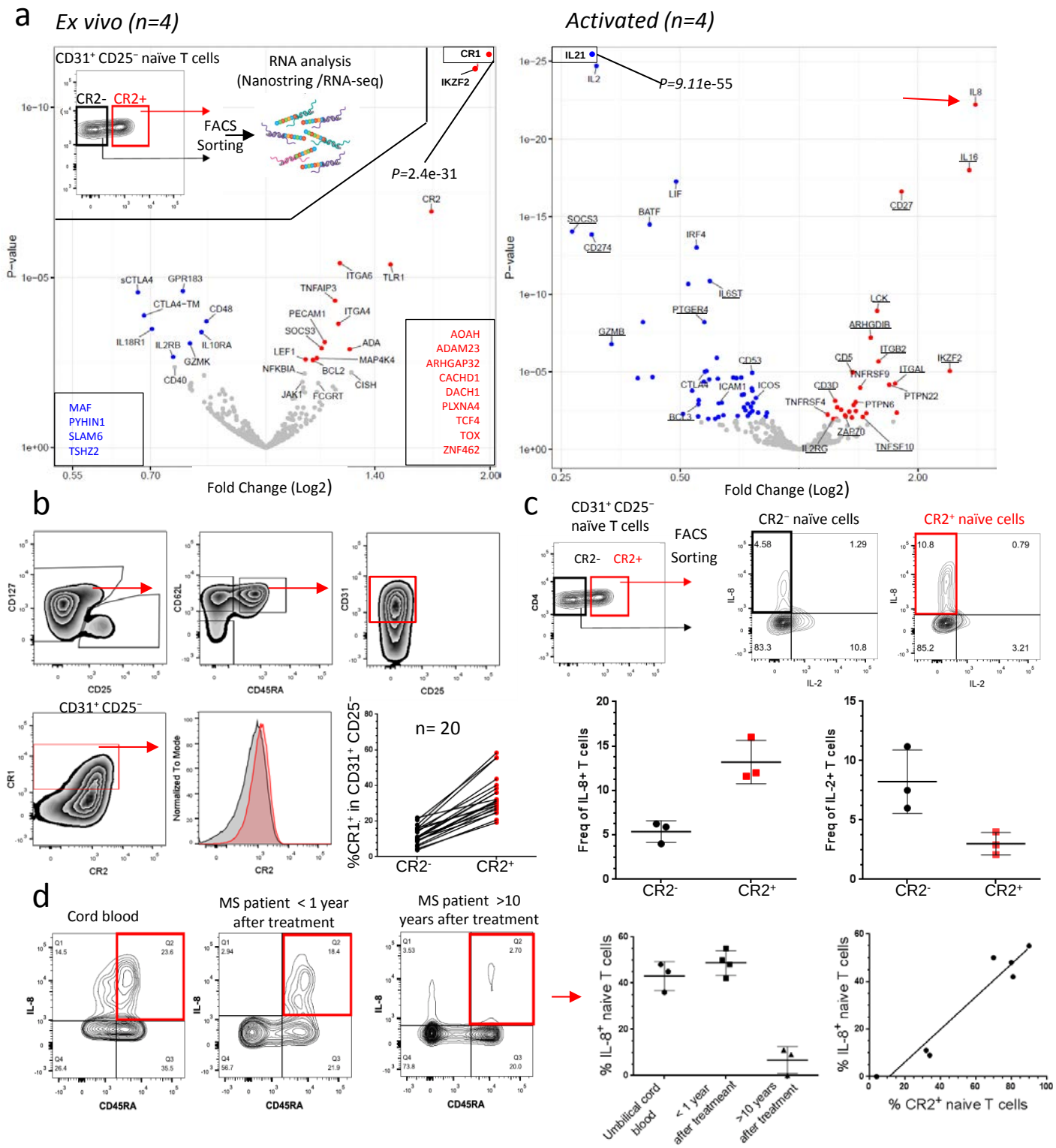
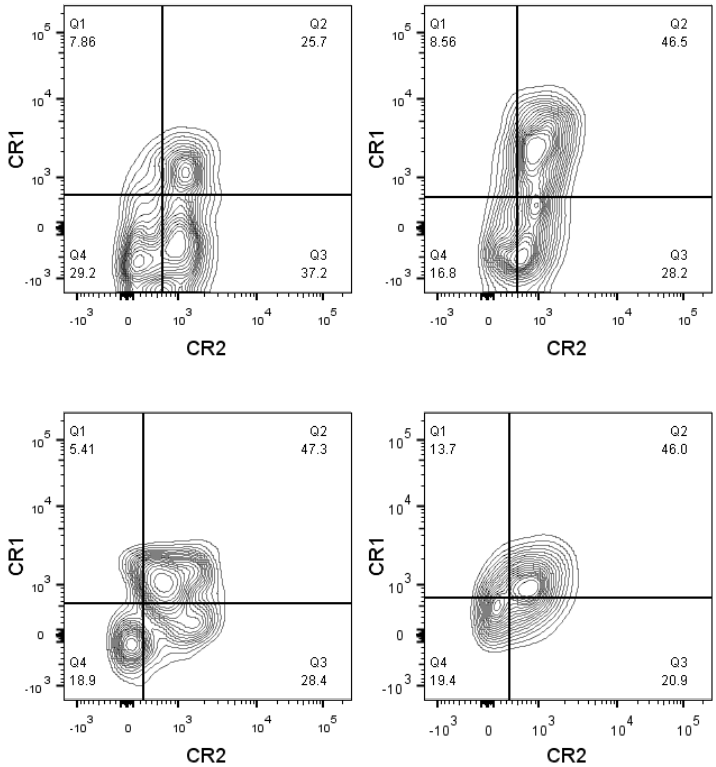


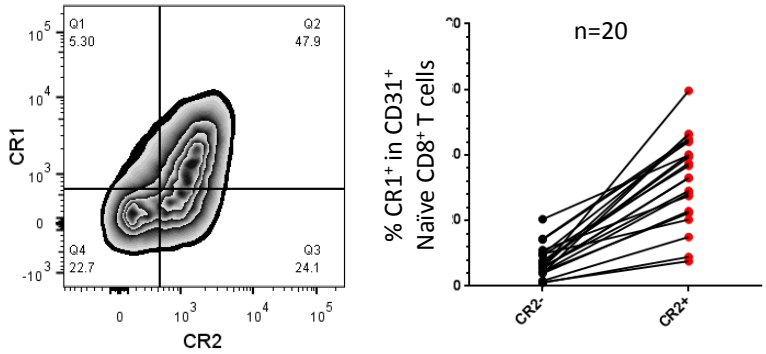
Figure 3 CR2⁺ naïve CD4⁺ T cells have a unique molecular signature. (**a**) Volcano plots representing changes in gene expression of CR2⁺ versus CR2⁻ naïve CD4⁺ T cells (gating strategy shown as insert) *ex vivo* and after activation. Genes underlined have reduced expression after activation. Genes in boxes are from RNA-seq expression data. (**b**) *Ex vivo* CR1 protein expression by CR2⁺ versus CR2⁻ cells within the CD31⁺ CD25⁻ naïve CD4⁺ T cell subset (N=20, age range 0 to 17). (**c**) Representative histograms and compiled frequencies of IL-8 and IL-2 production following activation of CR2⁺ and CR2⁻ cells sorted from CD31⁺ CD25⁻ naïve CD4⁺ T cells. (**d**) Representative histograms and compiled frequencies of IL-8 production from naïve (CD45RA⁺) T cells in cord blood and two adult MS patients. Correlation of the frequencies of IL-8 production from naïve CD4⁺ T cells with CR2⁺ cells in the CD31⁺ CD25⁻ naïve CD4⁺ T cell subset.

a Naïve CD4⁺ T cells in MS patients during *de novo* T cell reconstitution

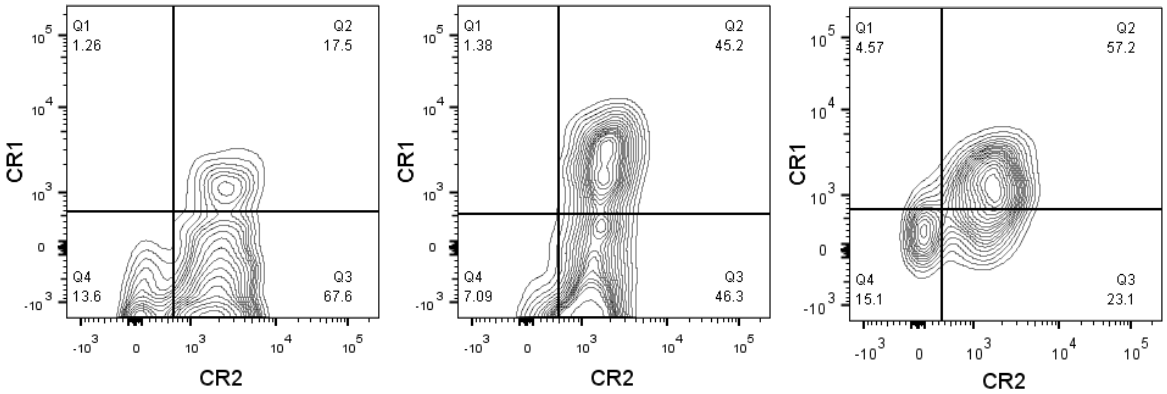


Supplementary Figure 4 CR2 and CR1 are co-expressed on naïve T cells. (**a**) CR1 and CR2 expression on CD31⁺ CD25⁻ naïve CD4⁺ T cells present in alemtuzumab-treated MS patients during reconstitution. (**b**) Representative example and summary analysis (n=20) of cell-surface CR1 expression on CR2⁺ and CR2⁻ CD31⁺ naïve CD8⁺ T cells. CR2 and CR1 expression on CD31⁺ naïve CD8⁺ T cells present in alemtuzumab-treated MS patients during reconstitution.

b Naïve CD8⁺ T cells from healthy donors



Naïve CD8⁺ T cells in MS patients during *de novo* T cell reconstitution



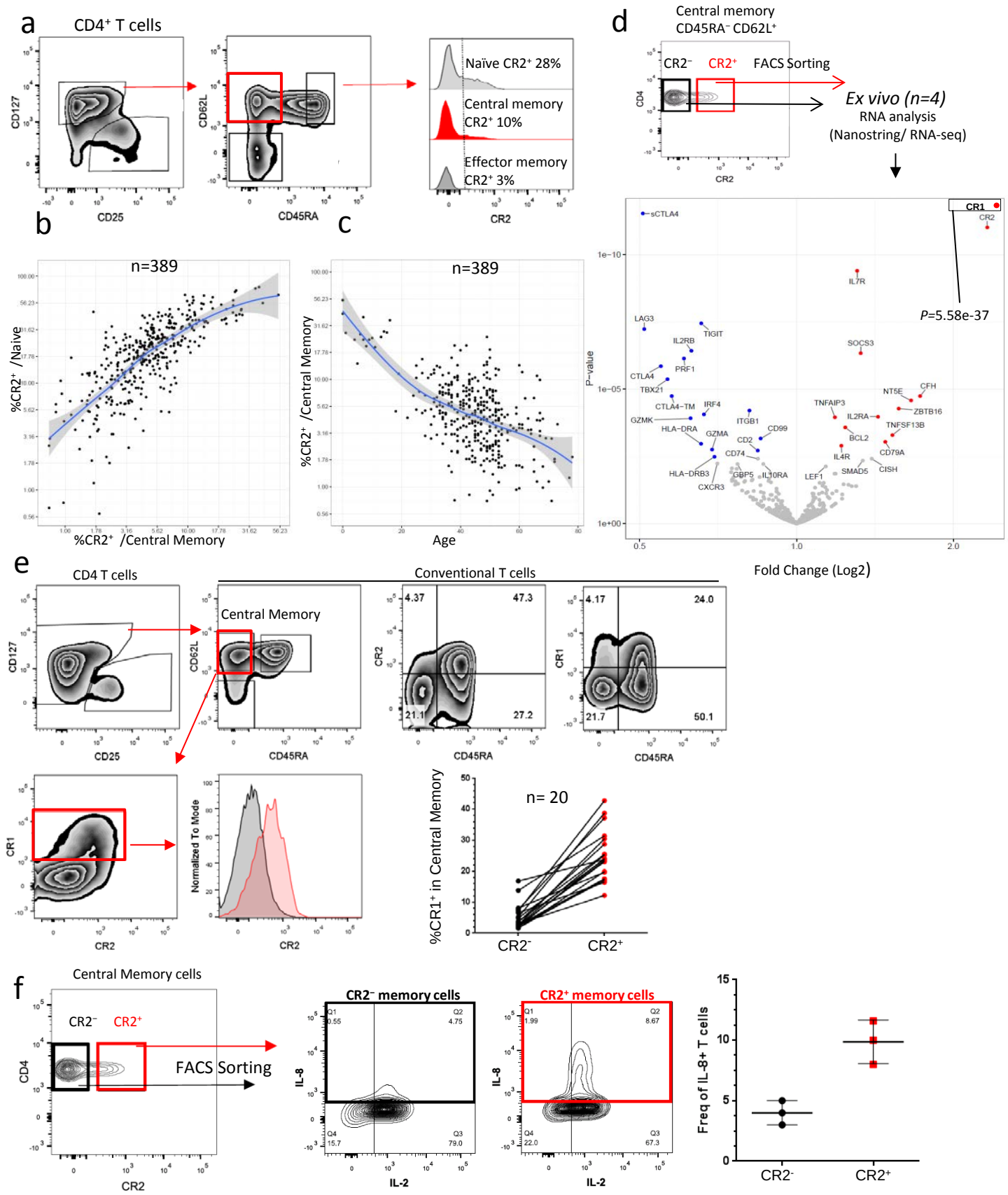
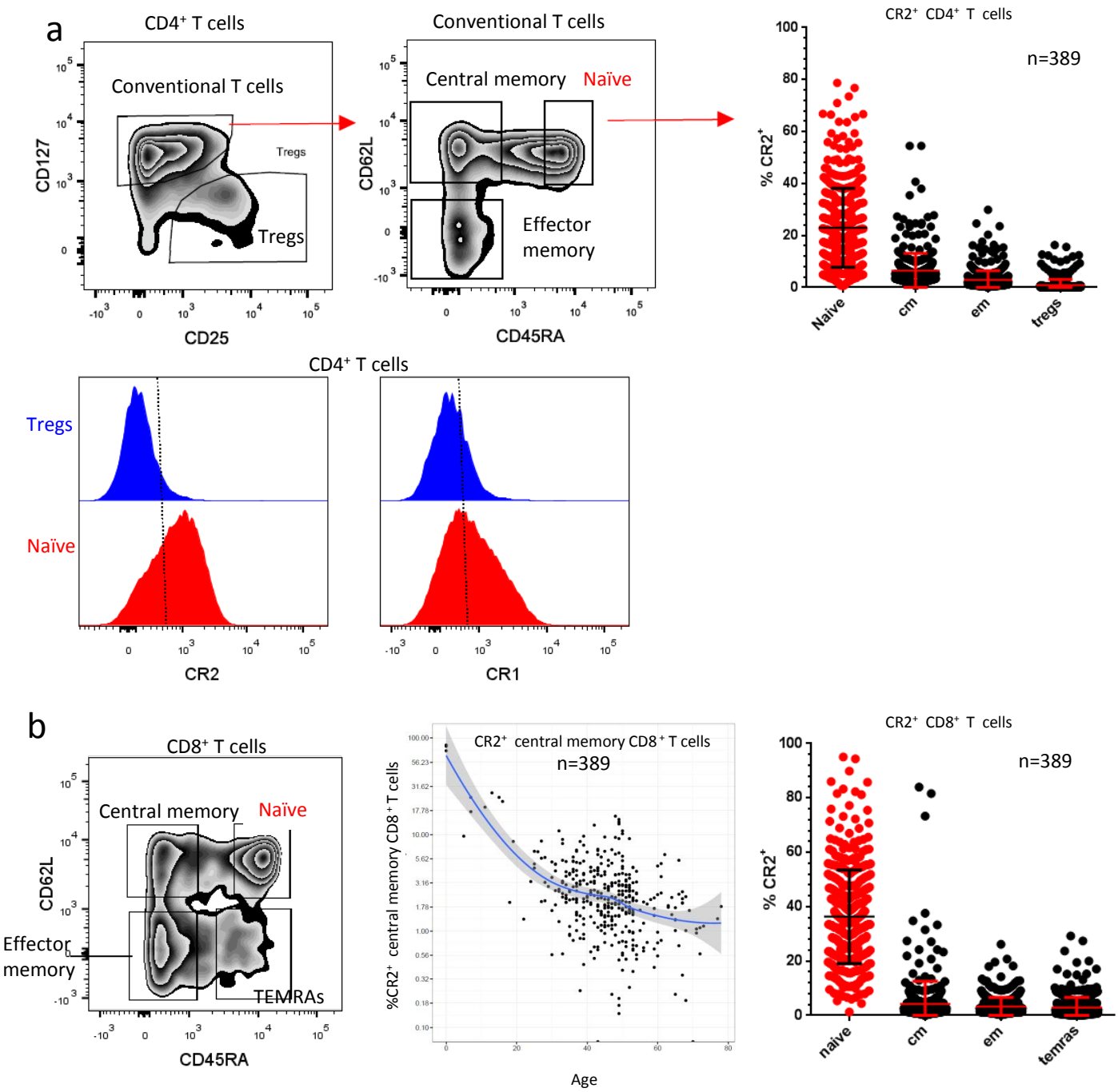
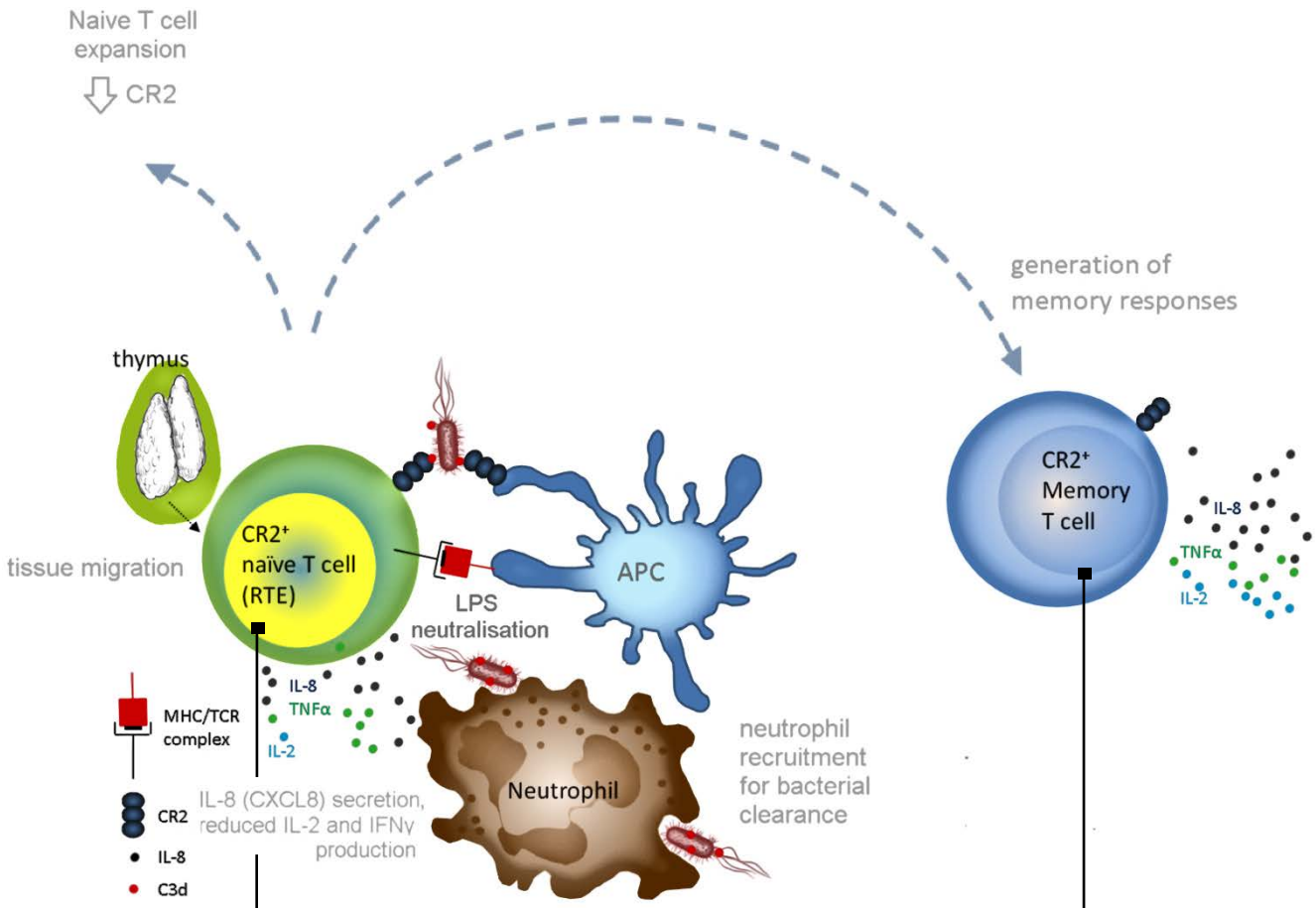


Figure 4 CR2 and CR1 are co-expressed on a subset of memory T cells. (**a**) Representative gating strategy and CR2 expression on memory subsets. Correlation of CR2 expression on memory T cells with CR2 expression on naïve T cells (**b**) and with age (**c**). (**d**) Representative gating of memory cells sorted by CR2 expression and gene expression analysis results. Color coding as described in **Fig. 1**. (**e**) FACS analysis of CR1 and CR2 co-expression. (**f**) Representative example and compiled data of IL-8 and IL-2 production from sorted CR2$^+$ and CR2$^-$ memory CD4$^+$ T cells.



Supplementary Figure 5 CR2 is expressed by subsets of CD4+ and CD8+ memory T cells. (**a**) Representative gating of Tregs and naïve, central memory and effector memory CD4+ T cells and the distribution of CR2+ cells within each of these subsets. Representative examples of CR1 and CR2 expression on naïve CD4+ T cells and CD4+ Tregs. (**b**) Representative gating of naïve and memory CD8+ T cell subsets. CR2 expression on central memory CD8+ T cells as a function of age. Distribution of CR2+ cells within naïve and memory CD8+ T cell subsets.



CR2⁺ naïve T cells *ex vivo*

AOAH, ADA, CACHD1, DACH1, FCGRT, ITGA4, ITGA6, TCF4, TLR1, LRRN3, TOX, IKZF2 (HELIOS)

CR2⁺ memory T cells *ex vivo*

ANK1, TNFSF13B (BAFF), CFH, CD79A, ZBTB16, IL2RA (CD25), NT5E (CD73), IL7R

Shared expression between CR2⁺ naïve and CR2⁺ memory T cells *ex vivo*

CR1, CR2	Complement receptors, regulate lymphocyte activation, recognition of microbial products
ADAM23	Metalloproteinase, binds to αVβ3 integrin, regulates cell-cell and cell-matrix interactions.
ARHGAP32	Rho GTPase-activating protein 32 neuron-associated GTPase-activating protein, regulates cell morphology.
DST	Dystonin, Bullous pemphigoid antigen; plakin protein family of adhesion junction plaque proteins, potentially aids in migration through tissues.
PLXNA4	Plexin A4 binds to neuropilin 1 (Nrp1) and neuropilin 2 (Nrp2), regulation of cell migration.
TNFAIP3	zinc finger protein and ubiqitin-editing enzyme, attenuates TNF signalling, inhibits NF-kappa B activation as well as inhibits TNF-mediated apoptosis.
BCL2	Anti-apoptotic protein.
CISH	CISH controls T cell receptor (TCR) signalling, and suppressor of cytokine signalling (SOCS), involved in anti-bacterial responses.
SOCS3	Suppressor of cytokine signalling 3, inhibition of JAK/STAT activation, Regulator of Infection and Inflammation.
ZNF462	Key transcription factor directing CR2⁺ lineage.

CTLA4, GBP5, GZMK, PYHIN1, SLAM6 ↓

Supplementary Figure 6 Gene expression in CR2⁺ T cells *ex vivo* and following activation. Gene names shown in the first two boxes are some of the differentially expressed genes having higher expression in either CR2⁺ naïve or CR2⁺ memory CD4⁺ T cells versus their CR2⁻ counterparts. Genes that share their expression patterns in both CR2⁺ naïve and memory cells versus their CR2⁻ counterparts are listed in the larger boxes (genes up-regulated in red, genes down-regulated in blue).

αβTCR⁺ CD4⁺ T cells (4 donors)

FACS-sorting

Naïve

Memory

CR2⁺

CR2⁻

CR2⁺

CR2⁻

Activated

Activated

Ex vivo

Ex vivo

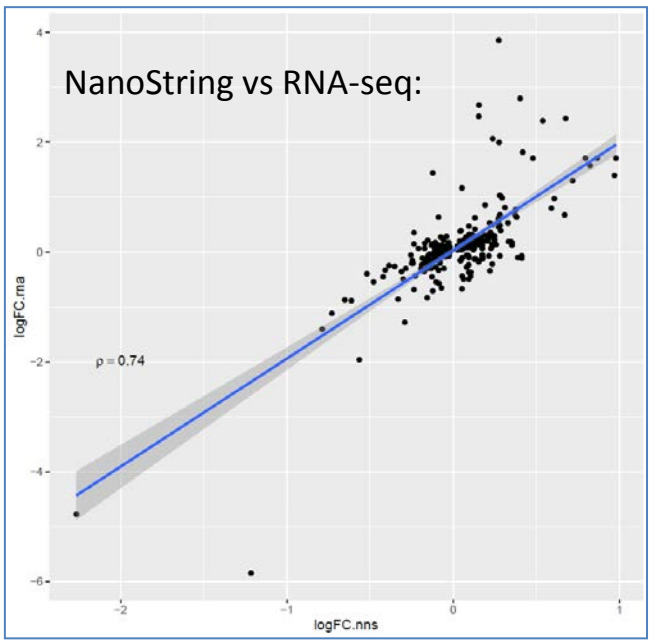
Ex vivo

Ex vivo

RNA expression analysis

RNA-seq

NanoString



Supplementary Figure 7 Design of the RNA analysis experiment performed on naïve CR2⁺ and CR2⁻ and memory CR2⁺ and CR2⁻ CD4⁺ T cells. Two RNA analysis methods, RNA-seq and NanoString (Human Immunology gene panel), were applied on RNA purified from FACS-sorted T-cell subsets from four donors. Fold change (Log2) differences obtained from RNA-seq and NanoString were highly correlated (example shown at the bottom right represents fold change between CR2⁺ and CR2⁻ memory CD4⁺ T cells on both platforms).

Supplementary Table 4 Selected results of RNA analysis experiment (RNA-seq) performed on CR2^{+} and CR2^{-} naïve and memory T cells *ex vivo.*

		Naïve *ex vivo*			Memory *ex vivo*		
Gene ID		**CR2+**	**CR2-**		**CR2+**	**CR2-**	
CR1							
	Donor1	271.382*	27.106	FC**	1008.486	39.988	FC
	Donor2	162.204	12.753	10.204	825.464	22.576	27.027
	Donor3	180.323	9.47	Adj P value***	1374.156	37.313	Adj P value
	Donor4	367.677	56.653	0.001	619.707	37.049	0.00001
CR2							
	Donor1	53.892	7.447	FC	53.824	0.881	FC
	Donor2	41.293	3.413	10.204	33.583	0.231	58.824
	Donor3	46.893	2.554	Adj P value	55.151	1.041	Adj P value
	Donor4	49.316	5.371	0.001	21.741	0.52	0.001
ZNF462							
	Donor1	6.921	0.44	FC	31.936	4.507	FC
	Donor2	3.14	0.237	12.987	17.958	0.971	7.194
	Donor3	13.643	0.887	Adj P value	24.472	4.212	Adj P value
	Donor4	10.583	0.967	0.005	23.426	2.914	0.005
CACHD1							
	Donor1	46.55	31.59	FC	1.997	0.102	FC
	Donor2	40.215	20.102	1.757	1.395	0.37	n/a
	Donor3	33.138	13.265	Adj P value	0.423	0.098	Adj P value
	Donor4	45.642	29.65	0.037	2.467	1.041	n/a
ADAM23							
	Donor1	11.703	2.482	FC	94.556	20.536	FC
	Donor2	2.344	0.711	6.803	45.053	9.016	5.051
	Donor3	21.67	1.561	Adj P value	90.442	14.808	Adj P value
	Donor4	14.583	2.578	0.02	36.686	6.765	0.00001
PLXNA4							
	Donor1	12.262	2.242	FC	83.232	20.638	FC
	Donor2	3.562	0	6.024	20.666	1.618	4.926
	Donor3	21.494	3.866	Adj P value	59.546	7.258	Adj P value
	Donor4	25.493	3.867	0.02	36.034	11.188	0.014
AOAH							
	Donor1	19.157	9.009	FC	7.017	7.388	FC
	Donor2	64.166	22.709	2.155	14.26	9.385	1.185
	Donor3	80.594	25.749	Adj P value	12.827	13.116	Adj P value
	Donor4	129.859	88.895	0.046	46.509	32.628	0.512
CFH							
	Donor1	0.037	0.08	FC	47.235	18.842	FC
	Donor2	0.094	0.759	n/a	41.539	23.903	2.653
	Donor3	0	0	Adj P value	45.677	10.805	Adj P value
	Donor4	0	0	n/a	58.334	19.254	0.016
DST							
	Donor1	11.404	5.966	FC	49.574	20.875	FC
	Donor2	15.467	7.823	1.770	30.896	8.692	2.653
	Donor3	1.653	0.816	Adj P value	33.11	10.089	Adj P value
	Donor4	32.451	20.196	0.038	61.639	25.655	0.006
ARHGAP32 (RICS)	Donor1	37.381	24.183	FC	50.943	19.79	FC
	Donor2	52.589	19.39	2.105	59.209	21.083	2.427
	Donor3	49.637	18.443	Adj P value	31.417	11.391	Adj P value
	Donor4	55.383	28.629	0.02	35.429	19.462	0.007
PYHIN1							
	Donor1	5.516	12.772	FC	30.121	47.477	FC
	Donor2	8.718	17.589	0.444	45.466	68.102	0.703
	Donor3	3.694	9.151	Adj P value	43.3	51.65	Adj P value
	Donor4	3.061	6.07	0.02	33.66	44.076	0.039

*Gene counts per million reads **FC= Fold change of CR2+ vs CR2- ***Adj P value= Adjusted P value

Supplementary Table 5 Selected results of RNA analysis experiment (NanoString) performed on CR2⁺ and CR2⁻ naïve T cells after activation.

Gene ID		Naïve activated		
		CR2⁺	CR2⁻	
IL8				
	Donor1	2529*	889.06	FC**
	Donor2	68.93	24.01	2.797
	Donor3	139.7	52.12	Adj P value***
	Donor4	187.1	49.8	7.36E-21
IL2				
	Donor1	3937	20724.1	FC
	Donor2	9108	31985.9	0.307
	Donor3	16729	40603.6	Adj P value
	Donor4	11808	40883.7	3.52E-23
IL21				
	Donor1	3.22	56.12	FC
	Donor2	25.36	280.62	0.107
	Donor3	34.18	385.16	Adj P value
	Donor4	26.44	341.48	3.19E-52
IFNG				
	Donor1	9.03	62.53	FC
	Donor2	8.45	16.54	0.426
	Donor3	13.5	36.02	Adj P value
	Donor4	36.45	83.95	3.11E-04
LIF				
	Donor1	417.8	768.8	FC
	Donor2	318.6	777.85	0.489
	Donor3	304.7	622.44	Adj P value
	Donor4	301.3	631.73	3.70E-16
TNF				
	Donor1	945.3	1129.56	FC
	Donor2	641.2	762.38	0.898
	Donor3	1487	1352.93	Adj P value
	Donor4	1580	1913.69	4.06E-01
IL23A				
	Donor1	578.4	535.52	FC
	Donor2	248.4	248.08	1.123
	Donor3	463.3	316.94	Adj P value
	Donor4	441.1	429.69	4.68E-01
LTA				
	Donor1	1115	1012.51	FC
	Donor2	1182	1187.05	1.137
	Donor3	1857	1361.83	Adj P value
	Donor4	2217	1964.91	2.66E-01

*Counts
**FC= Fold change of CR2+ vs CR2-
***Adj P value= Adjusted P value

Supplementary Table 7 Antibody panels used for immunophenotyping and FACS-sorting.

Panel	Antigen	Fluorochrome	Clone ID	Manufacturer
CR2 Immunophenotyping	CD3	BV510	OKT3	BioLegend
	CD4	BUV398	SK3	BD Biosciences
	CD8	APC-Cy7	RPA-T8	BioLegend
	CD56	BV711	NCAM	BD Biosciences
	CD127	PE-Cy7	eBioRDR5	eBioscience
	CD25*	APC	M-A251+2A3	BD Biosciences
	CD45RA	BV785	HI100	BioLegend
	CD62L	BV605	DREG56	BD Biosciences
	CCR7	BV421	GO43H7	BioLegend
	CD27	AlexaFluor700	M-T271	BioLegend
	CD95	PerCP eFluor710	DX2	BioLegend
	CD31	FITC	WM59	eBioscience
	CR2 (CD21)*	PE	BU32+Ly4	BioLeg+BD Bio
CR1/CR2 Immunophenotyping	CD3	BV510	OKT3	BioLegend
	CD4	BUV398	SK3	BD Biosciences
	CD8	APC-Cy7	RPA-T8	BioLegend
	CD56	BV711	NCAM	BD Biosciences
	CD127	PE-Cy7	eBioRDR5	eBioscience
	CD25*	APC	M-A251+2A3	BD Biosciences
	CD45RA	BV785	HI100	BioLegend
	CCR7	BV605	GO43H7	BioLegend
	CD27	AlexaFluor700	M-T271	BioLegend
	CD95	PerCP eFluor710	DX2	eBioscience
	CD31	FITC	WM59	eBioscience
	CR2 (CD21)*	PE	BU32+Ly4	BioLeg+BD Bio
	CR1 (CD35)	BV421	E11	BD Biosciences
FACS sorting of four naïve CD4+ subsets defined by CD25 and CD31 expression using CD4+ TCRab+ cells enriched by negative selection (RosetteSep)	TCRab+	FITC	IP26	BioLegend
	CD4	AF700	RPA-T4	BioLegend
	CD25*	APC	M-A251+2A3	BD Biosciences
	CD127	PE-Cy7	eBioRDR5	eBioscience
	CD45RA	BV785	HI100	BioLegend
	CD31	PE	WM59	BioLegend
	CD62L	BV605	RPA-T8	BioLegend
	CCR7	Pacific Blue	GO43H7	BioLegend
	Viability Dye	eFluor780		eBioscience
FACS sorting of naïve and memory subsets defined by CR2 expression using CD4+ TCRab+ cells enriched by negative selection (RosetteSep)	CD4	AF700	RPA-T4	BioLegend
	CD25*	APC	M-A251+2A3	BD Biosciences
	CD127	PE-Cy7	eBioRDR5	eBioscience
	CD45RA	BV785	HI100	BioLegend
	CD62L	BV605	RPA-T8	BioLegend
	CD31	BV421	WM59	BioLegend
	CR2 (CD21)	PE	BU32	BioLegend
	Viability Dye	eFluor780		eBioscience
Cytokine production (IL-8 and IL-2) following activation	CD4	AF700	RPA-T4	BioLegend
	CD25*	APC	M-A251+2A3	BD Biosciences
	CD127	PE-Cy7	eBioRDR5	eBioscience
	CD45RA	BV785	HI100	BioLegend
	IL-2	BV510	MQ1	BioLegend
	IL-8	FITC	E8N1	BioLegend
	Viability Dye	eFluor780		eBioscience

*Two clones each of anti-CD25 and anti-CR2 were used to enhance detection.