

Supplementary Materials:

Materials and methods

Participants and samples

Our analysis encompasses two sub-cohorts recruited consecutively; sub-cohort 1 (74 samples) including young and older individuals, and sub-cohort 2 (168 samples) including older individuals only.

Exclusion criteria were as follows; known immunodeficiency (including HIV-infection), organ transplantation, use of immunosuppressive or immune-modulating drugs within the last year (excluding acetylsalicylic acid 100mg/day), cancer or treatment for cancer within the previous 5 years, insulin dependent diabetes, moderate or advanced renal failure, liver disease, endocrine disorders (except corrected thyroid dysfunction), manifest autoimmune disease, dementia/mental incompetence, alcohol/other drug abuse, acute infection or illness in the last 4 weeks, raised body temperature ($>37.5^{\circ}\text{C}$), moderate or severe heart failure (NYHA III or IV), inability to lie flat. Inclusion and exclusion criteria were checked and medications recorded at the time of recruitment. A more detailed medical history was available for sub-cohort 2 only. Demographics and medications for both sub-cohorts are listed in supplementary Table S1, relevant medical history items for sub-cohort 2 are shown in supplementary Table S2. Venous blood (approximately 30 ml) was collected into sodium-heparin plasma tubes (BD, Oxford, UK).

Whole blood antibody staining for enumerating CD28^{null} CD4 T-cells

Fresh, sodium-heparin-anticoagulated whole blood (100 μL) was incubated with monoclonal, fluorescence-labeled staining antibodies as follows. Sub-cohort 1 (74 samples); CD45-APC, CD3-Violet(V)500 (BD), CD4-PB, CD8-APC-H7, CD27-PE (all BD), and CD28 AF700 (BD). Sub-cohort 2; anti-CD45-AlexaFluor(AF)700, anti-CD3-Pacific Blue (PB), anti-CD4-Brilliant Violet (BV)510, anti-CD8 α -allophycocyanin(APC)-H7, anti-CD27-APC, anti-CD28-phycoerythrin(PE)/Cyanine(Cy)7 (all BioLegend, Cambridge, UK). Cells were incubated either for 30 min at 4°C in the dark (sub-cohort 1) or for 20 minutes at RT in the dark (sub-cohort 2). followed by addition of BD lysis buffer (BD) and incubation for 10 min according to the manufacturer's instructions. Samples were then centrifuged (5

min at 350g), resuspended in PBS and spun again (5 min at 350g). Finally, cells were fixed (1% paraformaldehyde in water) and washed. Pellets were resuspended in 300 μ L of staining buffer (BD) prior to acquisition within 1-4h. Careful analysis confirmed that both protocols provided equivalent results for CD28^{null} CD4 T-cell frequencies.

Additional phenotypic characterization of CD28^{null} CD4 T-cells

After thawing cells, 1x10⁶ PBMCs were placed in a tube, washed with PBS and stained with the following surface antibodies: anti-CD3-BV605, anti-CD4-BV510, anti-CD8 α -APC/H7, anti-CD28-PE/Cy7, anti-CD27-APC, and anti-CD49d-PE (all BioLegend, Cambridge, UK), as well as Yellow live-dead stain (Invitrogen, Paisley, UK), for 20 minutes at RT in the dark. Subsequently cells were washed with PBS 1x and centrifuged (8' at 400g). Finally cells were fixed with 0.5% Paraformaldehyde (in deionized water), washed and resuspended in staining buffer (BD). Samples were acquired on an LSR II cytometer within the following hour.

CMV-peptides

Twenty-five μ g per peptide (PepMixTM, JPT Peptide Technologies, Berlin, Germany), each covering one complete protein, were dissolved in 100 μ L of DMSO (Sigma-Aldrich) (0.25 μ g/ μ L). Two μ L of the 0.25 μ g/ μ L CMV-peptide solutions were used in T-cell stimulation assays (details further down) to provide a final concentration of 1 μ g/mL for each peptide. Peptides pools for the following CMV proteins were used (in n = number of individuals); UL85 (pp65) (n = 3), UL86 (n = 2), UL55 (n = 1) and UL86 / UL55 (n = 1). SEB (Sigma) dissolved in DMSO was used at 1 μ g/ml (final concentration, positive control), 2 μ L DMSO alone (i.e. the peptide solvent) was used as negative control.

CMV-reactivity of CD28^{null} CD4 T-cells

CMV-specific CD4 T-cells from 7 older CMV+ individuals were activated by incubating PBMC overnight (37°C, humidified 5% CO₂ atmosphere) with overlapping peptide pools representing the CMV proteins, UL55, UL83, or UL86 (PepMix, JPT Peptide Technologies, Berlin, Germany). For each tube, 2x10⁶/200 μ L of PBMC suspension was incubated with 50 μ L stimulation solution containing 1 μ L of antigen/positive control or 2 μ L of negative control preparation plus 49 or 48 μ L, respectively, of complete RPMI media (final incubation volume 250 μ L). CMV-antigen

preparation consisted of 2 μL of the selected CMV peptide pool dissolved in dimethyl sulfoxide (DMSO, Sigma- Aldrich, Gillingham, UK) at a final peptide concentration of 0.25 $\mu\text{g}/\mu\text{L}$, the positive control was 1 μL of staphylococcal enterotoxin B (SEB, Sigma) in DMSO at 0.25 $\mu\text{g}/\mu\text{L}$, the negative control preparation was 2 μL of DMSO in complete media. After 2 hours in a standard incubator (37°C, humidified 5% CO₂ atmosphere) 5 μg of Brefeldin A (BFA) in 1 μg of DMSO (Sigma, Gillingham, UK) was added in 249 μL of complete RPMI giving a final BFA concentration of 10 $\mu\text{g}/\text{ml}$ in a final volume of 500 μL for the remaining incubation time of 14 hours. At the end of the incubation time, 100 μL of 20mM EDTA (Sigma) was added to each tube followed by incubation for 10 minutes at 37°C in a water bath. Samples were then washed with wash buffer consisting of PBS with 0.5% w/v BSA (Sigma) and 0.1% w/v sodium azide (Sigma) and centrifuged at 400g for 8 minutes at RT. Cells were washed once more with PBS prior to staining. Samples were then stained for 30 minutes at 4°C in the dark with the following surface antibodies: anti-CD3-Pacific Blue, anti-CD4-BV510, anti-CD8a-APC/H7, anti-CD49d-PE and anti-CD28-PE/Cy7 (all Biolegend), and a Yellow live-dead stain (Invitrogen, Paisley, UK). Following the surface stain, 1ml of 1x BD 'FACS Lysing Solution' (BD) was added to each tube followed by incubation for 10 min in the dark. Cells were then washed by adding 3 mL PBS followed by centrifugation (400g for 8 minutes at RT). Subsequently BD 'Permeabilization solution 2' was added to each tube and samples were incubated for 10 minutes in the dark. After an additional wash step cells were stained intracellularly with anti-IL-2 APC, anti-TNF α -AF 700 and anti-IFN- γ FITC (all Biolegend) for 30 minutes at 4°C in the dark. Finally, cells were washed again with PBS and re-fixed in a 0.5% paraformaldehyde (in water) and stored until analysis at 4°C in the dark.

Quality control for flow-cytometry

Comparable day-to-day performance of the LSR II was ascertained by running CS&T calibration beads (BD) on a daily basis. In addition, 8-peak Rainbow beads were used prior to every run in order to adjust PMTs in such a way that with respect to each detector, every peak was always in the same channel.

HLA-typing and sequence analysis

Two-field resolution HLA class II typing was performed by Sanger sequencing after group-specific amplification (Protrans, Hockenheim, Germany) using 48 and 96 capillary-sequencers (3730 and 3730xl DNA Analyzer, Thermo Fisher Scientific, Waltham, MA USA). The group-specific amplification ensured hemizygous, unambiguous sequences of the entire exon 2 in all samples without the need of assigning heterozygous positions. Allele assignment was done using the HLA-HiType software package (inno-train, Kronberg, Germany) and the most recent IMGT/HLA Database Release.

HLA-sequence information was obtained from www.hla.alleles.org and processed in Python (v3.6.1) making use of the following libraries; biopython 1.70 for sequence alignment and handling, pandas 0.20.1 for data and spreadsheet handling, and NumPy 1.13.1 for array and matrix operations.

Statistical Analysis

Histograms and Q/Q plots were used to determine data distribution for cell populations of interest. The Kolmogorov-Smirnov and Shapiro-Wilk tests were used to test for normality.

Non-parametric tests (Mann-Whitney) were used to compare groups. T-cell frequencies were log-transformed where appropriate for normalizing distribution or improving data presentation. P-values ≤ 0.05 were considered significant for single endpoints. Multiple end-point correction (Bonferroni) was applied when appropriate ($0.05/n$, where n is the number of endpoints).

Receiver-Operator characteristic (ROC) analysis was used in order to test the ability of parameters to discriminate groups, usually CMV- and CMV+ individuals. In the entire study cohort (n=242), hierarchical multiple linear and binary logistic regression models were generated in order to test the effect of sex, age, CMV status or the presence/absence of HLA-allele groups on the frequencies (multiple linear regression) of CD28^{null} CD4 T-cells or the presence or absence of an expansion (binary logistic regression). Sex was always introduced at the 1st level (block 1) of hierarchical regression models; age at the 2nd level (block 2), and CMV infection status at the 3rd level (block 3). Variable selection for HLA allele groups (linear regression) was based on a Spearman Rho correlation matrix. Alleles/allele groups correlating with another allele/allele group with $R > 0.3$ were included but those correlating with $R > 0.8$ were

excluded. Additional data reduction was achieved by principle component analysis (SPSS). Differences in regards to the frequencies of individual amino acids at certain positions in exon 2 of DRB1 or DQB1 between individuals with or without expanded CD28null CD4 T-cell populations were tested using Pearson's CHI square test.

Supplementary Figures

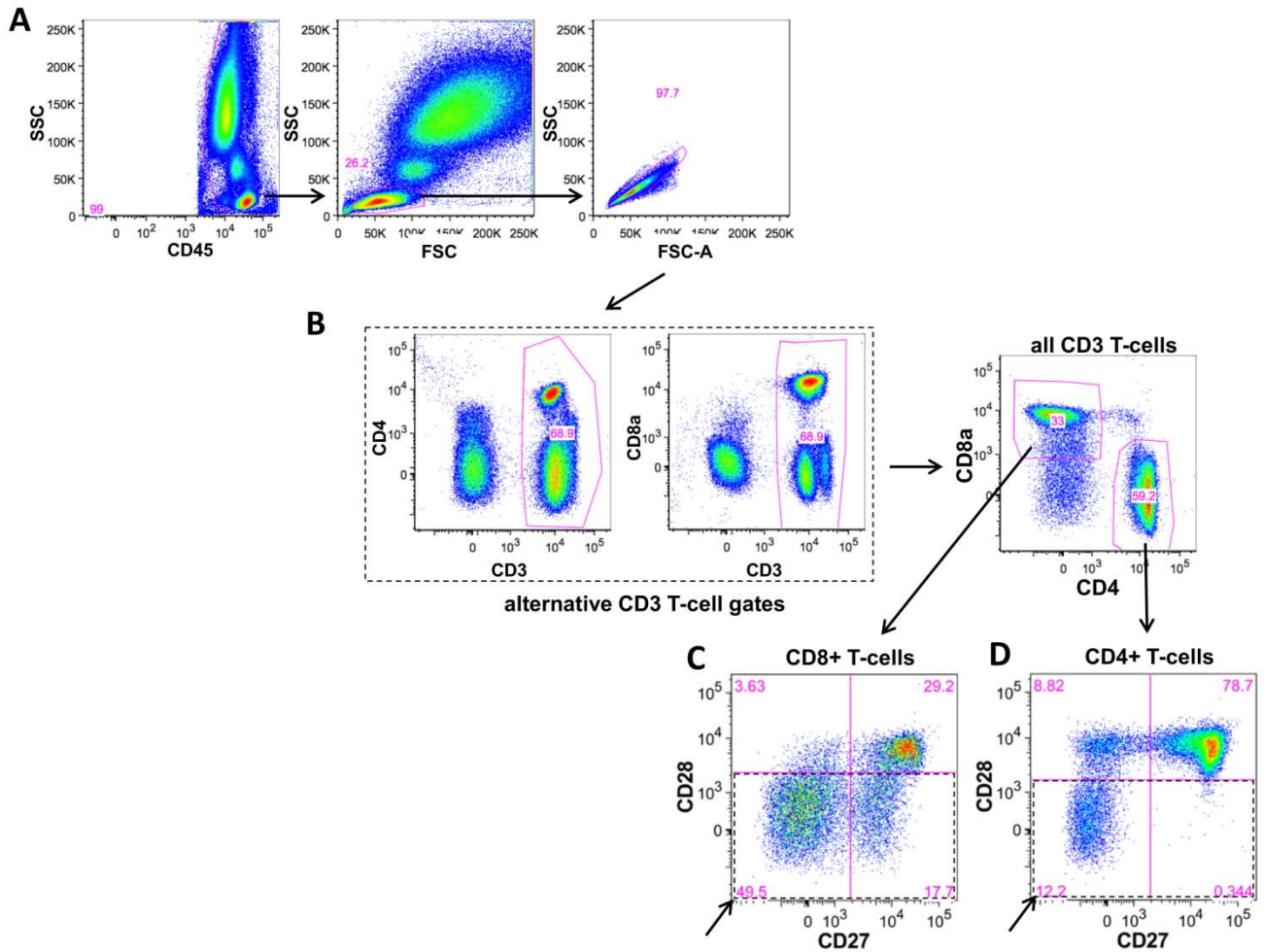


Fig. S1. Gating strategy for the analysis of CD28^{null} CD4 and CD28^{null} CD8 T cells. (A) A Leucocyte gate was created in a CD45/SSC plot and subsequently lymphocytes were gated in a SSC/FSC plot. Singlets were gated in a FSC-A/FSC-H plot, doublets were excluded. (B) From this gate two CD3 T-cell gates were created in CD3/CD4 and CD3/CD8 fluorescence dot-plots respectively, allowing to include activated T-cells downregulating CD3, CD4, or CD8 on the surface. The two CD3 T-cell gates were then combined logically ('OR') generating the CD3 T-cell gate from which CD4 T-cells were then selected (C). For the analysis of CD28^{null} CD8 T-cells, CD8 T-cells were subsequently visualized in a CD27/CD28 fluorescence dot-plot to allow determining the size of the CD28^{null} CD8 T-cell populations (dot line) and its CD27⁺ and CD27⁻ subsets. (D) For the analysis of CD28^{null} CD4 T-cells, CD4 T-cells were subsequently visualized in a CD27/CD28 fluorescence dot-plot to allow determining the size of the CD28^{null} CD4 T-cell populations (dot line) and its CD27⁺ and CD27⁻ subsets. Arrows indicate CD28^{null} CD8 and CD28^{null} CD4 T cells. Representative plots from a CMV+ individual are shown.

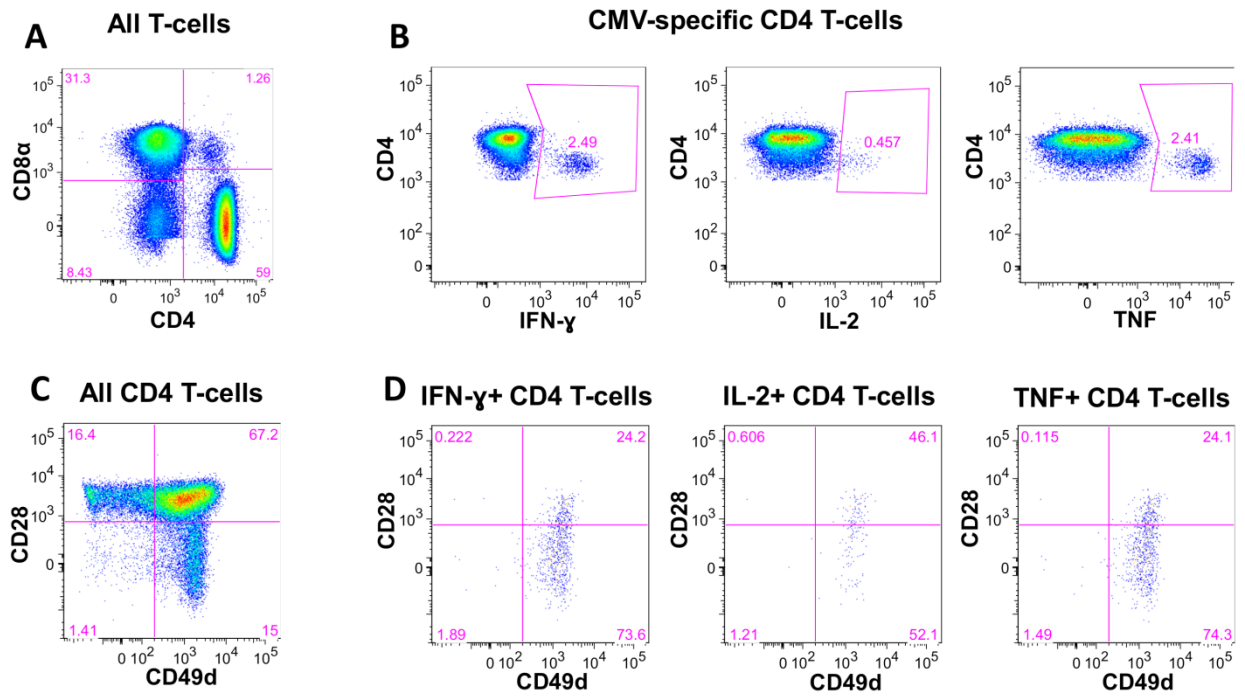


Fig. S2. Gating strategy for enumerating and phenotyping CMV-activated CD4 T-cells.

(A) A T-cell gate was created in analogy to the gating strategy shown in Supplementary Fig S1. (B) CD4 T-cells were then visualized in CD4 versus IFN- γ , TNF, or IL-2 fluorescence dot-plot in order to identify activated CD4 T-cells. (C) CD4 T-cells were next shown in a CD28/CD49d fluorescence dot-plot to draw limits for positive and negative populations. (D) Activated CD4 T-cell populations were visualized in CD28/CD49d fluorescence dot-plots and the limits for CD28 and CD49d positive and negative populations (quadrants) were overlaid on these plots to determine the distribution of activated CD4 T-cells with respect to CD28 and CD49d expression

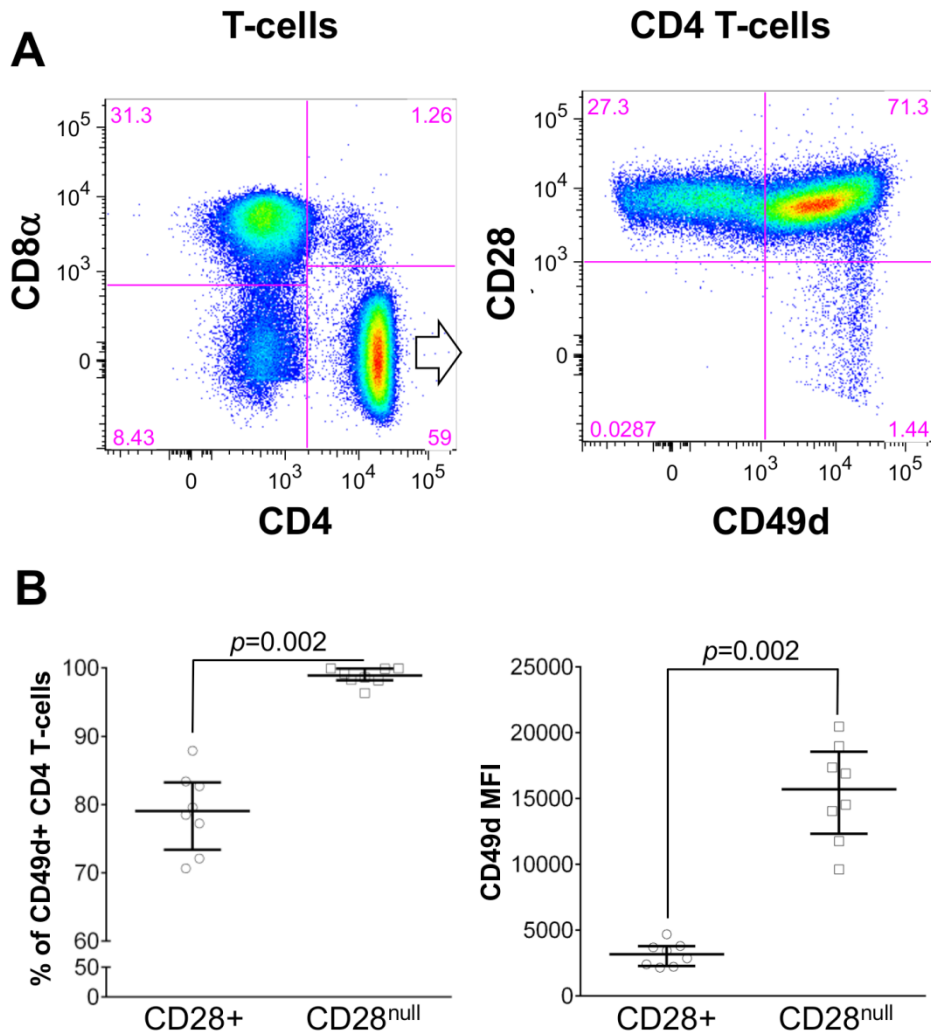


Fig. S3. CD28^{null} CD4 T-cell phenotype.

(A) Cells were gated in analogy to the strategy shown in Supplementary Fig S1 CD4 T-cells were then visualized in CD49d versus CD28 fluorescence dot-plots and expression limits were set for each marker determining positive and negative subsets. (B) Scatter plots show the distribution of CD49d+ CD4 T-cells into CD28^{null} and CD28+ subsets (left) as well as the expression level (mean fluorescence intensity, MFI) of CD49d in both subsets. The data shown is from 8 individuals of which 7 are the same as shown in Figure 3. Error bars show median and interquartile range.

Supplementary Tables

Table S1. Participant demographics and medications

Parameter	Young		p-value	Older		p-value
Total n in group	27		n.a. ^a	215		
Age range (mean ± STD)	19 – 32 (23.5±3.5)		n.a.	60 – 94 (71.7±7.3)		n.a.
Sub-groups	CMV–	CMV+		CMV–	CMV+	
Total n in group	13	14	n.a.	93	122	n.a.
Age range (mean ± STD)	20-25 (21.6±1.9)	19-32 (25.3±3.8)	n.s. ^b	60-92 (71.5±7.4)	60-94 (71.8±7.3)	n.s.
Antihypertensive medications (excl. diuretics and beta-blockers)	n.a.	n.a.	n.a.	28	40	n.s.
Number	n.a.	n.a.	n.a.	1-3	1-3	
Diuretics	n.a.	n.a.	n.a.	3	18	0.005
ACE inhibitors	n.a.	n.a.	n.a.	3	4	n.s.
Beta-blockers	n.a.	n.a.	n.a.	12	16	n.s.
Ca-Antagonists	n.a.	n.a.	n.a.	13	14	n.s.
ATIIR antagonists	n.a.	n.a.	n.a.	5	15	n.s.
Anti-hypertensives other	n.a.	n.a.	n.a.	3	6	n.s.
Oral anti-hyperglycemics	n.a.	n.a.	n.a.	2	5	n.s.
Statins	n.a.	n.a.	n.a.	20	29	n.s.

^a not applicable

^b not available

Table S2. Participant medical history (Sub-cohort 1)

Parameter^a	CMV-	CMV+	p-value
Age range (mean ± STD)	60 – 94 (71.7±7.3)		n.a. ^b
Total n per sub-group	73	95	n.a.
Age range (mean ± STD)	61-92 (71.2±7.4)	63-94 (72.2±7.1)	n.s. ^c
Previous stroke	0	1	n.s.
Previous TIA	1	4	n.s.
Hypertension	26	35	n.s.
Atrial Fibrillation	7	7	n.s.
Cardiac Prosthesis/Stent	2	7	n.s.
Heart Failure	2	2	n.s.
Angina pectoris	5	4	n.s.
COPD	1	2	n.s.
Malignancy/Cancer in past	4	8	n.s.
Osteoarthritis	17	26	n.s.
Diabetes	3	5	n.s.
Oral anticoagulation	6	14	n.s.
Smoker	5	6	n.s.
If yes: pack/year history	19	36	n.s.
Ex-smoker	35	46	n.s.
If yes: pack/year history	16	18	n.s.

^a medical history based on interview

^b not applicable

^c not statistically significant

Table S3. HLA-DRB1*03:01 protects from CD28^{null} CD4 T-cell expansion^a

Variable	Parameter estimate	Standard Error	p-value	OR	95% CI for OR	Pseudo R² (model fit)^b
sex	-1.060	.611	.082	.346	.105 - 1.146	30.3- 40.4%
Age	.054	.024	.022	1.056	1.008 - 1.106	
DRB1*03:01	-2.124	1.021	.003	.108	.025 - .474	
constant	-2.571	1.687	.127	.076	n.a.	

^aA binary logistic regression model was constructed based on sex (step 1), age (step 2), and HLA-DQB1*02:01 as well as HLA-DRB1*03:01 (step 3). The table shows step 3 of the model. It provided correct classification for 73.4% of cases.

^bThe model fit is shown as the range between Cox & Snell and Nagelkerke pseudo R².

Table S4. Distribution of HLA-DRB1 exon 2 positions 11,71, and 74 ('haplotypes') in individuals with or without CD28^{null} CD4+ T-cell expansions

HLA-DRB1 haplotype	DRB1 position			No expansion	Expansion	associated RA severity	Associated HLA-DRB1 alleles (present in cohort)
	Pos. 11	Pos. 71	Pos. 74	% of individuals in cohort			
'VKA'	Val	Lys	Ala	15.6%	25%	<i>severe</i>	<u>*04:01P</u>
						<i>severe</i>	<u>*04:04P,</u> <u>*04:05P,</u> <u>*04:08P,</u> <u>*10:01P</u>
'VRA'	Val	Arg	Ala	6.3%	12.5%		<u>*11:03P,</u> <u>*13:01P,</u> <u>*13:02P</u>
'SEA'	Ser	Glu	Ala	18.8%	18.8%	<i>mild</i>	<u>*03:01P</u>
'SKR'	Ser	Lys	Arg	43.8%	9.4%	<i>mild</i>	<i>n.a.</i>
Age (years)	n.a.	n.a.	n.a.	71.7 (21.4)	72.4 (8.0)		<i>n.a.</i>

Table S5. The SKR ‘haplotype’ protects from CD28^{null} CD4 T-cell expansion^a

Variable^b	Parameter estimate	Standard Error	p-value	OR	95% CI for OR	Pseudo R² (model fit)^c
Sex (m)	-1.058	.625	.091	.347	.102 - 1.182	
Age	.057	.024	.019	1.059	1.009 - 1.111	
VKA	.272	.769	.724	1.313	.291 – 5.928	37.1 – 42.2%
VRA	1.294	1.223	.290	3.648	.332 – 40.087	
SKR	-2.233	.762	.003	.107	.024 - .477	
constant	-2.915	1.748	.095	.054	n.a.	

^aA binary logistic regression model was constructed based on the HLA-DRB1 haplotypes present in our cohort and most strongly associated with the severity of rheumatoid arthritis. The model provided correct classification in 75% of cases.

^bSEA was not included in the model as its frequency was the same in both groups.

^cThe model fit is shown as the range between Cox & Snell and Nagelkerke pseudo R².

Table S6. Participant lipid profiles (Sub-cohort 1)

Parameter	CMV- (n=72)	CMV+ (n=88)	p-value
Total cholesterol	5.4 ± 1.0	5.3 ± 1.1	n.s. ^a
HDL-cholesterol	1.8 ± .6	1.7 ± .5	n.s.
LDL-cholesterol	3.1 ± .9	3.0 ± 1.0	n.s.
Non-HDL cholesterol	3.7 ± .9	3.6 ± 1.0	n.s.
LDL/HDL-cholesterol ratio	2.0 ± .8	1.9 ± .8	n.s.
Triglycerides	1.3 ± .9	1.2 ± .6	n.s.

^a not statistically significant

Table S7. Demographics of individuals used for CMV-specific CD28^{null} CD4 T-cell characterization

Individual	Age	Sex	CMV-proteins selected	% of CD28^{null} CD4 T-cells
1	67	Female	UL83	1.27
2	82	Female	UL55	1.78
3	62	Male	UL86	20.04
4	82	Male	UL83	4.89
5	80	Male	UL55 / UL85 ^a	8.65
6	68	Male	UL85	21.78
7	66	Female	UL83	10.82

^aThis individual had big responses for UL55 and UL85 and both were used in separate stimulations.