CD29 marks superior cytotoxic human T cells

Benoît P. Nicolet[1], Aurélie Guislain[1], Floris P.J. van Alphen[2], Raquel Gomez-Eerland[3], Ton N. M. Schumacher[3], Maartje van den Biggelaar[2,4], and Monika C. Wolkers[1,*]

[1] Department of Hematopoiesis, Sanquin Research and Landsteiner Laboratory, Oncode institute, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands.

[2] Department of Research Facilities, Sanquin Research, Amsterdam, The Netherlands.

[3] Division of Molecular Oncology and Immunology, Oncode Institute, The Netherlands Cancer Institute (NKI), Amsterdam, The Netherlands.

[4] Department of Molecular and Cellular Haemostasis, Sanquin Research, Amsterdam, the Netherlands

* Corresponding author: M. Wolkers, Sanquin Research, Department of Hematopoiesis, Plesmanlaan 125, 1066 CX Amsterdam, The Netherlands. Tel: +31-20-5127003; Fax: +31-20-5123474; Email: m.wolkers@sanquin.nl.

One Sentence Summary: Transcriptomics and proteomics on fixed human $CD8^+$ T cells with a differential cytokine production profile identified T cells with high cytotoxic potential.

ABSTRACT

Cytotoxic $CD8^+$ T cells can effectively kill target cells by producing cytokines, chemokines and granzymes. Expression of these effector molecules is however highly divergent, and tools that identify and pre-select potent killer cells are lacking. Human $CD8^+$ T cells can be divided into IFN-γ and IL-2 producing cells. Unbiased transcriptomics and proteomics analysis on cytokine-producing fixed $CD8^+$ T cells revealed that IL-2^+ cells produce helper cytokines, and that IFN-γ^+ cells produce cytotoxic molecules. IFN-γ^+ T cells could be identified with the surface marker CD29 already prior to stimulation. CD29 also marked T cells with cytotoxic gene expression from different tissues in single-cell RNA-sequencing data. Notably, the cytotoxic features of $CD29^+$ T cells were maintained during cell culture, suggesting a stable phenotype. Pre-selecting CD29-expressing MART1 TCR-engineered T cells potentiated the killing of target cells. We therefore propose that CD29 expression can help evaluate and select for potent therapeutic T cell products.

INTRODUCTION

$CD8^+$ T cells can effectively clear cells from our body that are infected with intracellular pathogens. In addition, $CD8^+$ T cells are critical for immunosurveillance to eliminate pre-cancerous or cancerous cells. To exert their effector function, T cells produce pro-inflammatory molecules such as IFN-γ and TNF-α, granzymes and perforin, and chemokines such as CCL3 (MIP1α) and CCL4 (MIP1β)(*1–3*). Intriguingly, human T cells do not respond uniformly to activation, but rather show a heterogeneous production profile of effector molecules (*4*, *5*). This may derive from stochastic and oscillating production of effector molecules, as suggested in developmental processes (*6*). However, it may also be due to intrinsic differences in the effector response to antigen by distinct $CD8^+$ T cell types.

Interestingly, even though activated human $CD8^+$ T cells efficiently produce IFN-γ and IL-2, we and others found that they generally produce only one of the two cytokines (*5*, *7*, *8*). Only a small proportion of human $CD8^+$ T cells produces IFN-γ and IL-2 combined (*5*, *7*, *8*). Several studies proposed that human and mouse $CD8^+$ T cells can provide T cell 'help' by expressing CD154/CD40L (*7*, *9*, *10*), in addition to their well-known cytotoxic function. CD40L-expressing human $CD8^+$ T cells were shown to produce IL-2, but not IFN-γ (*7*). However, whether this cytokine production profile is stochastic, or whether IFN-γ and IL-2 producing $CD8^+$ T cells represent two different T cell subsets and/or a divergent differentiation status is yet to be determined. It is also not known whether IL-2 and IFN-γ producing $CD8^+$ T cells possess a different cytotoxic potential.

In this study, we set out to dissect the properties of human $CD8^+$ T cells that display a differential cytokine production profile. Previous studies used cytokine capture assays to isolate T cells for gene expression analysis (*11*). While this method is very accurate for low abundance responses, the risk of selecting false-positive cells by capturing cytokines from the neighbouring cell is substantial. We therefore developed a protocol that allowed for efficient isolation of RNA and protein from fixed T cells after intracellular cytokine staining. With this top-down approach, we performed an unbiased RNA-seq and Mass spectrometry analyses on IFN-γ and IL-2 producing primary human $CD8^+$ T cells.

We found that IFN-γ producing T cells exhibit a cytotoxic expression profile, which is sustained during T cell culture. The surface molecules CD29 and CD38 identified the cytotoxic IFN-γ-producing T cells, and the non-cytotoxic IL-2 producing $CD8^+$ T cells, respectively, a feature that is maintained during prolonged T cell culture. Using these two surface markers allowed us to distinguish T cells with a distinct cytokine production profile and cytotoxic potential already prior to T cell stimulation. Furthermore, the CD29 gene expression signature was a good prognostic marker for long term survival in melanoma patients. In conclusion, we here provide evidence that CD29 selects for bona fide cytotoxic $CD8^+$ T cells, a marker that could potentially be used to evaluate the quality of therapeutic T cell products and possibly to select for effective adoptive T cell products.

RESULTS

Efficient recovery of RNA and protein from T cells after intracellular cytokine staining

To determine whether and how the gene expression profile of IFN-γ producing $CD8^+$ T cells differed from IL-2 producing cells, we needed a method to separate high percentages of cytokine producers while preventing possible contaminations from false-positive cells. We therefore developed a protocol for efficient recovery of RNA and proteins from FACS-sorted fixed cells after intracellular cytokine staining. With the standard TRIZOL RNA isolation protocol, >99% of the mRNA isolated from PMA-ionomycin activated and formaldehyde-fixed T cells was lost, as determined by RT-PCR of several standard housekeeping genes, and of *IFNG* and *IL2* mRNA (fig. S1A, B). To conserve the RNA integrity, we performed the intracellular cytokine staining (ICCS) and FACS in an RNA-protecting high salt buffer that contained RNAse inhibitors (fig. 1A). This alteration in the protocol preserved the quality of ICCS, and thus the distinction of different cytokine producers (*5*). To free RNA from protein complexes, we included a proteinase K digestion step prior to RNA isolation (*12*) (fig. 1A). These adjustments yielded mRNA levels from fixed T cells comparable to "fresh" T cells that were prepared without permeabilization and formaldehyde fixation (fig. S1A, B). Importantly, the RNA purity and integrity was compatible with RNA-seq analysis, as revealed by nanodrop analysis and RNAnano Chip assay (fig. S1C).

To achieve mass spectrometry (MS) -grade protein recovery, we used the Filter Aided Sample Preparation (FASP) protocol (*13*). This protocol is compatible with MS analysis from formalin fixed material, as it contains a 95°C heating step to reverse the formaldehyde cross-linking of the ICCS procedure (*14*, *15*). Indeed, the correlation coefficient of MS analysis from fixed and fresh samples was r=0.973 (fig. S1D). Thus, RNA and protein could be efficiently recovered from

cytokine-producing, formaldehyde-fixed T cells, which allowed us to study their expression profile in depth.

Differential gene expression profile of IFN-γ and IL-2-producing CD8$^+$ T cells

To determine the gene expression profile of IFN-γ producing cells with that of IL-2 producing cells, human CD8$^+$ T cells from 3 biological replicates, each containing a pool of 40 healthy donors, were activated for 2 days with α-CD3/α-CD28. T cells were removed from the stimulus and cultured for 4 days in the presence of human recombinant IL-15. ICCS after 4h of activation with PMA-ionomycin identified four T cell populations: IFN-γ producers (IFNG), IL-2 producers (IL2), IFN-γ/IL-2-coproducing T cells (double positive, DP), and T cells that did not produce detectable levels of either cytokine (double negative, DN; fig. 1A).

Of the RNA-sequencing data, on average 19.8 million reads (~92%) of the total 21.4 million reads could be mapped to the genome of each T cell population. Of these, 2349 genes were differentially expressed in at least one population, with a cut-off of log2 fold change >0.5, p-adjusted <0.01 (table S1). According to Ensembl-BioMart gene annotation, 85% (n=1984) of the differentially expressed genes were protein coding, 5.3% were antisense, 3.9% were pseudogenes, 2.6% were long-intergenic non-coding RNA (lincRNA), and 3.2% comprised other classes of non-coding RNA (ncRNA; fig. 1B).

Pearson's correlation coefficient of differentially expressed genes revealed that each biological replicate clustered according to its cytokine production profile (fig. 1C). Interestingly, IFNG cells clustered with DP cells, and IL2 cells with DN cells (fig. 1C). This close relationship of IFNG and

DP cells and of IL2 with DN cells was also revealed by gene cluster analysis using k-means (k=5), in particular in cluster 2 and 4 (241 and 434 genes, respectively) (fig. 1D, table S1). Yet, cluster 1 (507 genes) was more specific for IL2 T cells, and for IFNG T cells cluster 3 and 5 (474 and 326 genes, respectively; fig. 1D, table S1).

We next interrogated whether gene expression modulators like transcription factors (TFs), long non-coding RNAs (lncRNAs), and RNA binding proteins (RBPs) were differentially expressed in IFNG and IL2 producers. Most of the 134 differentially expressed TFs were shared between DN and IL2 producers, and between IFNG and DP producers (fig. 1E, table S2). This included GATA3 and IRF8 for IL2/DN producers, and PRDM1 (Blimp-1), NFATc2 and NFATc3 for IFNG/DP producers (fig. S1E, table S2). Specific gene expression was detected for Eomes and RORA in IFNG producers, and for RORC (ROR-γt) in DP producers (table S2). A small cluster of TFs including IKZF2 (Helios) was specifically up-regulated in DN cells (fig. S1E). Also the 351 differentially expressed lncRNAs revealed the close kinship between DN/IL2 producers and IFNG/DP producers (fig. 1F, table S2). For instance, the IFN-γ promoting IFNG-AS1 (NEST) (*16*) was highly expressed in IFNG/DP producers (fig. S1F). Conversely, GATA3-AS1 known to re-enforce the Th2 phenotype in $CD4^+$ T cells (*17*) showed highest gene expression in IL2 producers (fig. S1F).

RBPs are critical regulators of RNA splicing, stability and translational control (*18*). Again, most of the 180 differentially expressed RBPs were similarly expressed in IFNG/DP producers and in IL2/DN producers (fig. 1G, table S2). Combined, these data revealed a differential gene expression profile in IFNG and IL2 producers, which correlates with distinct expression of gene regulators.

Differential protein signature of IL-2 and IFN-γ producing CD8$^+$ T cells

We next determined the protein expression profile of the four FACS-sorted CD8$^+$ T cell populations by MS. We identified in total 5995 proteins from the 3 biological replicates. After removing possible contaminants and filtering for proteins that were present in each biological replicate in at least one of the four populations, we identified a total of 3833 proteins. Each sample contained >3500 identified proteins (fig. 2A, table S1). Pearson's correlation of the MS analysis also revealed the kinship between IFNG and DP producers, and between IL2 and DN producers (fig. 2B), and by k-means clustering of the 81 differentially expressed proteins (9 clusters; fig. 2C). Of note, 42 (52%) of the differentially expressed proteins overlapped with the differential gene expression profile, including IL-2 and IFN-γ (fig. 2D). In conclusion, IL2 and IFNG producers display a distinct gene and protein expression signature.

IFN-γ producing CD8$^+$ T cells display a cytotoxic profile

To determine whether IFNG and IL2 producers showed a distinct gene expression also of other secreted effector molecules, we selected differentially expressed genes from fig. 1 that 1) had expression levels >100 RPM and that 2) were annotated as "secreted" in the human protein atlas (*19*). After removing nuclear components and mitochondrial and ribosomal proteins (see method section), we identified 67 differentially expressed genes encoding secreted proteins (fig. 3A). For IL2 producers, these included the prototypic Th2 helper effector molecules IL-4, IL-3, leukemia inhibitory factor (LIF) (*20*, *21*), and the Th2-inducing cytokine IL-11 and prostaglandin-endoperoxide synthase 2 (PTGS2, or COX-2) (*22*, *23*) (fig. 3A). IL-2 and PTGS2 were also found differentially expressed in the MS analysis (fig. 3B). Conversely, IFNG producers - and in most cases also DP producers - contained substantially higher gene expression levels of the prototypic

cytotoxic molecules TNF-α, Granulysin, Granzyme A, Granzyme H and Perforin (fig. 3A). Also the chemokines CCL5 (RANTES), CCL3 (MIP1α), CCL4 (MIP1β) and its alternative splicing variant CCL4L2, all of which are associated with antiviral and antibacterial activity of cytotoxic $CD8^+$ T cells (*24–26*), were highly expressed in IFNG and DP producers (fig. 3A). IFN-γ, TNF-α, Granulysin, Perforin, and CCL5 were also identified by MS analysis (fig. 3B).

We validated this prototypic cytotoxic gene expression profile of IFNG and DP cells in a new cohort of four individual donors. ICCS after reactivation for 4h with PMA-ionomycin confirmed that IFNG/DP producers produced significantly more TNF-α, Granulysin, Perforin, Granzyme A and B, and the chemokines CCL3, CCL4, and CCL5 than the DN/IL2 producers (fig. 3C). Not only the percentage, but also the production of cytokines/chemokine per cell was higher, as determined by the geometric mean fluorescence intensity (geo-MFI; fig. S2A). We therefore conclude that IFNG and DP $CD8^+$ T cells preferentially express cytotoxic effector molecules.

CD29 and CD38 expression preselects for IFN-γ and IL-2 producing T cells

We next interrogated if surface markers could be used to identify IFNG producers and IL2 producers prior to T cell activation. To this end, we filtered the MS data for surface molecules that were ≥1.5 fold higher on average in expression in one of the four T cell populations. 19 surface markers showed differential protein expression in at least one of the four T cell populations (fig. 4A). RNA-seq analysis identified 72 differentially expressed annotated proteins (fig. S2B). We tested the expression pattern of 28 of these markers by flow cytometry on PMA-ionomycin activated $CD8^+$ T cells on a new cohort of blood donors.

None of the 16 surface markers that were only identified by RNA-seq analysis was suitable to select for IFNG and IL2 producers by flow cytometry (fig. S2C). In contrast, the HLA class II histocompatibility antigen gamma chain CD74, and the β1-integrin were differentially detected in both MS and RNA-seq analysis (fig. S2C) showed significantly higher cell surface expression in IFNG/DP producers than in IL2/DN producers (fig. 4B; fig. S2D, E). Conversely, the expression of CD40L, the complement decay-accelerating factor CD55, and the cyclic ADP ribose hydrolase CD38 were co-expressed and significantly higher in IL2/ DN producers (fig. 4B, fig. S2D, F). Differential surface expression of CD29 and CD38 was also found in $CD4^+$ IFNG/DP and IL2/DN producers (fig. S2G).

In contrast to CD40L and CD74, CD29 and CD38 expression was also measured in non-activated $CD8^+$ T cells. Importantly, the expression of CD29 and CD38 allowed for preselection of IFNG/DP producers and DN/IL2 producers even prior to activation in $CD8^+$ and in $CD4^+$ T cells (fig. 4C, fig. S3A, B). This included the production of TNF-α, Granzyme A, Granulysin, Perforin, CCL3, CCL4, and CCL5 by $CD29^+$ T cells (fig. S3C). Because $CD29^+$ T cells included DP producers (fig. 4B), the overall IL-2 protein production is similar in $CD29^+$ and $CD38^+$ T cells, yet clearly distributed between $CD29^+$ DP and $CD38^+$ IL2 single producers (fig. S3D).

Intriguingly, even though the percentage of $CD29^+$ and $CD38^+$ T cells was greatly variable between donors (fig. S3E), their expression pattern and cytokine production profile was maintained throughout time (fig. 4D, E; fig. S3F, G). Specifically, reactivation of T cells that were FACS sorted based on CD29 and CD38 expression on day 7 after activation and cultured for another 16 days showed a robust expression of CD29 and CD38, with $CD29^+$ T cells being the prime IFNG producers, and $CD38^+$ T cells the IL2 producers (fig. 4D, E). $CD29^+$ FACS-sorted T cells also

expanded effectively in vitro (fig. S3H). Using different cytokines during CD3/CD28 activation, or during the resting phase did not alter the percentage of $CD29^+$ T cells (fig. S3I, J). We therefore conclude that the differences between $CD29^+$ and $CD38^+$ T cells are cell-intrinsic and can be used to distinguish IFNG/DP producers from IL2/DN producers.

CD29 is expressed on non-naïve $CD8^+$ T cells

We next investigated whether CD29 expression also correlated with a cytotoxic gene signature of $CD8^+$ T cells *ex vivo*. Because CD38 expression was not detected on peripheral blood-derived $CD8^+$ T cells (fig. 5A), we compared $CD29^+$ with $CD29^-$ T cells. Also *ex vivo*, $CD29^+$ T cells were effective cytokine producers upon PMA-Ionomycin stimulation (fig. S4A). In concordance with the described expression profile of cytotoxic molecules (*27*, *28*) and its expression in effector T cells (*29*), CD29 expression was detected primarily in $CD45RA^+CD27^-$ Effector $CD8^+$ T cells (T_{Eff}), $CD45RA^-CD27^-$ Effector Memory (T_{EM}) and $CD45RA^-CD27^+$ Memory $CD8^+$ T cells (T_M; fig. 5B). Of the $CD45RA^+CD27^+$ naïve T cells (T_N), only 6.6±4.2% expressed CD29 (fig. 5B). However, further analysis revealed that the $CD29^+$ 'naïve' T cells rapidly produce IFN-γ upon 4h PMA-Ionomycin stimulation, and express CD49d (fig. 5C, D), previously shown to identify memory T cells with a naïve phenotype (TMNPs (*30*)). This finding thus indicates that CD29 expression marks non-naïve $CD8^+$ T cells.

To determine how CD29 expression related to antigen-specific triggering, we studied Choriomeningitis virus (CMV) and Epstein-Barr virus (EBV) specific T cells. In healthy blood donors, CMV-specific T cells are primarily of the effector, and EBV-specific T cells of the central memory phenotype (*31*, *32*). We activated PBMCs with CMV and EBV peptide pools and found

that the IFN-γ producing CMV- and EBV- specific T cells are primarily $CD29^+$ (fig. 5E, F). In conclusion, CD29 is almost exclusively expressed by non-naïve T cells.

scRNA-seq analysis reveals the cytotoxic signature of $CD29^+$ $CD8^+$ T cells

To further characterize the gene expression profile of $CD29^+CD8^+$ T cells, we re-analysed single-cell RNA-seq (scRNA-seq) data of peripheral blood-derived $CD8^+$ T cells from Guo *et al* (*33*). To distinguish naïve T cells from non-naïve $CD29^+$ T cells, we used unsupervised clustering followed by differential expression analysis. We selected for high *LEF1*, *CCR7*, *SELL* gene expression and low *CCL5*, *GZMB*, *PRF1* gene expression to identify T cells with a T_N transcriptome signature (*34*, *35*) (fig. S4B, C). *CCL5* gene expression highly correlated with non-naïve T cells and was thus used to enrich for those for downstream analysis (fig. S4D-F).

ITGB1 (CD29) gene expression showed a non-uniform expression in non-naïve $CD8^+$ T cells (fig. 5G, left panel; fig. S4G). We compared the gene expression profile of non-naïve T cells with high or low *ITGB1* gene expression, and found that high *ITGB1* gene expression strongly associated with high gene expression levels of *GNLY* (Granulysin), *GZMB* (Granzyme B), and the transcription factor *ZNF683* (Hobit) that is associated with cytotoxic activity of human T cells (*36*) (fig. 5G, right panel). Conversely, in concordance with previously described gene expression profiles (*27*, *37*), $ITGB1^{low}$ $CD8^+$ T cells associated with high gene expression levels of *GZMK* (Granzyme K), and the inhibitory receptor *KLRB1* (CD161; fig. 5G, right panel, table S3).

scRNA-seq analysis of $CD8^+$ T cells from lung tissue (*33*) also revealed a non-uniform expression of the *ITGB1* transcript (fig. S4H, I left panel). Again, high *ITBG1* expression associated with a

cytotoxic gene expression profile (fig. S4I, right panel). When we compared the top 50 most upregulated genes in *ITGB1*^{high} T cells derived from blood and lung, and from lung and liver tissue (*38*), we identified a core signature of *ITGB1*^{high} CD8^{+} T cells that encompassed the cytotoxic molecules *GNLY*, *GZMB*, *GZMH*, and *FGFBP2*, and *LGALS1* (Galectin-1; fig. 5H, fig. S4J). Of note, the fractalkine receptor *CX3CR1* is also part of the *ITGB1* ^{high} signature (fig. 5H), which we confirmed with flow cytometry analysis of blood-derived CD8^{+} T cells (fig. 5A). However, in contrast to murine CD8^{+} T cells (*28*), in humans CX3CR1 expression is almost exclusively found on CD45RA^{+}CD27^{-} effector CD8^{+} T cells, and only on a minority of T_M or T_{EM} cells (fig. S5B) (*39*). In line with the restricted expression pattern of CX3CR1, CD29 expression is superior in predicting the level of IFN-γ and TNF-α production (fig. S5C-F). Thus, *ITGB1* gene expression identifies human T cells with a cytotoxic gene expression profile in blood and in tissues.

CD29 gene signature is prognostic for long term survival of melanoma patients

We next interrogated if *ITGB1* gene expression correlated with good clinical responses. We focussed on cutaneous melanoma (SKCM), because the Cox regression analysis for survival from all TCGA datasets on CD8B expression identified SKCM as the only tumor type with a clear benefit of CD8^{+} T cell infiltration (fig. S6A, p=0.000125). *ITGB1* is expressed by many cell types, including tumor cells (*40*). Therefore, we used the CD29^{+} core signature as defined in Figure 5H. Intriguingly, analysis of TCGA data of melanoma patients revealed that the CD29^{+} core signature correlated well with a positive clinical outcome (high CD29 signature: median= 10.08 years, low CD29 signature: median= 4.84 years; p=<0.001; fig 6A). In particular, using the patient cohort with high CD8^{+} T cell infiltrates levels estimated by CIBERSORT (*41*) further showed a clear benefit of the CD29 core signature for long-term survival (high CD29 signature : median= 13.5

years, low CD29 signature: median= 5.37 years; fig. 6B). Thus, the CD29 signature has a good prognostic value for melanoma patients.

MART1 TCR-engineered CD29$^+$ CD8$^+$ T cells potently kill tumor cells

We then examined whether CD29 expression is indicative of effective T cell responses against tumor cells. We retrovirally transduced CD8$^+$ T cells with the codon-optimized MART1-TCR that recognizes the HLA-A*0201 restricted 26-35 epitope of MART1(*42*). FACS-sorted CD29$^+$ and CD38$^+$ MART1 TCR-engineered T cells were then exposed for 5h to a patient-derived HLA-A201$^+$ MART1^{high} expressing melanoma tumor cell line (MART1+), or to the HLA-A201$^-$ MART1^{low}-expressing cell line (MART1-)(*43*, *44*) (fig. 6C). MART1 TCR-engineered CD29$^+$ T cells produced substantially higher levels of IFN-γ, TNF-α, Granulysin, and CD107α in response to MART1 antigen-expressing tumor cells than CD38$^+$ T cells (fig. S6B). CD29$^+$ T cells were the major source of IL-2$^+$IFN-γ$^+$ double producing T cells, and CD38$^+$ T cells were enriched for IL2 single producers (fig. 6C, fig. S6B).

Lastly, we determined whether MART1 TCR-engineered CD29$^+$ T cells were also more potent in killing tumor cells than CD38$^+$ T cells. CFSE-labelled MART1+ and MART1- tumor cells were co-cultured for 20h with FACS-sorted MART1 TCR-engineered total CD8$^+$ T cells, or with FACS sorted CD29$^+$ or CD38$^+$ T cells. We found that MART1+ tumor cells, but not MART1- cells, were effectively killed by MART1 TCR-engineered T cells (fig. 6D). Notably, in line with their cytotoxic gene and protein expression profile, CD29$^+$ T cells showed superior killing when compared to CD38$^+$ T cells (fig. 6D, E). The superior killing was evident for different effector:target ratio's (fig. 6C).

Intriguingly, when we compared the killing capacity of the TCR engineered T cell product with the percentage of CD29-expressing cells present in the total T cell product, we found a clear correlation of these two parameters (fig. 6F). Thus, CD29 expression identifies T cells with the highest cytotoxic potential in response to tumor cells.

DISCUSSION

In this study, we present a method to perform RNA-seq and mass spectrometry analysis from formaldehyde-fixed, permeabilized primary $CD8^+$ T cells. This method uses the endogenous protein expression measured by intracellular cytokine staining for selection, resulting in a reliable isolation of different populations without the need of genetic modifications (*11*). We anticipate that one can readily adapt this rapid and cost-effective method to select cells with other features, such as differential chemokine or transcription factor expression.

Here, we uncovered that the differential cytokine production profile of human $CD8^+$ T cells reflects stable T cell-intrinsic properties. Whereas IL2 producers display a gene expression profile resembling helper T cells, IFNG producers express a cytotoxic core signature. With CD29 and CD38, IFNG-DP producers can be identified and separated from DN-IL2 producers even prior to T cell stimulation, a feature that was maintained upon prolonged *in vitro* cultures. The division between highly cytotoxic $CD29^+$ T cells and $CD38^+$ T cells was not only observed in $CD8^+$ T cells but was also true for $CD4^+$ T cells.

Importantly, scRNA-sequencing analysis on blood and lung derived $CD8^+$ T cells revealed that the cytotoxic core signature of $CD29^+$ T cells is not limited to *in vitro* T cell cultures but is also found *ex vivo*. It has been previously reported that a cytotoxic gene expression of tumor-infiltrating T cells is associated with better patient survival (*45*, *46*). In line with that finding, TCGA analysis in cutaneous melanoma patients also indicates that the CD29 core gene signature is a good prognostic marker for long term survival. In fact, it further stratifies long-term survival in patients with high CD8 infiltration, which is already a good predictor.

We found that CD29 expression marks non-naive T cells. CX3CR1^{+} T cells are also included in this population. However, because CX3CR1 is primarily expressed in effector T cells in humans that are known to poorly expand in vitro (*28*), we consider CD29 superior to identify cytotoxic T cells with a good proliferative capacity. Selecting for CD29^{+} T cells resulted in the generation of TCR engineered T cell products with improved killing of target cells at least *in vitro*.

Between blood donors, we found a great variation in the percentage of CD29^{+} T cells. It is conceivable that this donor-individual difference in CD29 expression stems from a differential pathogenic and microbial exposure. Yet, CMV and EBV infections efficiently generated CD29^{+} IFN-γ-producing T cells, indicating that a variety of pathogens may yield cytotoxic CD29^{+} T cells. Of note, the clear correlation with CD29 expression and cytokine production was not detected in mice housed under specific pathogen free (SPF) conditions (unpublished observation). This may again point to the possibility that pathogen exposure contributes to the generation of non-naïve cytotoxic CD29^{+} T cells.

Previous studies showed that the cytokine production profile of CAR T cells was highly diverse (*2*). As culturing T cells with different cytokines did not alter the relative distribution of CD29^{+} and CD29^{-} T cells, we propose to select for CD29 expression for the generation of highly toxic T cell products, in particular when donors have a low percentage of CD29^{+} T cells. Selecting for CD29^{+} cytotoxic T cells to generate T cell products may also come with a positive side effect. IL2 producers displayed a Th2 gene expression profile, which includes IL-3, IL-4, IL-11, LIF and PTGS2 (COX-2). A Th2 cytokine profile may impede anti-tumoral responses (*47*). In fact, IL2

and DN populations produce PTGS2 (COX-2) that was recently shown to suppress anti-tumoral immunity (*48*, *49*). It is therefore conceivable that removing IL2/DN producers from T cell products may already support the anti-tumoral responses of $CD29^+$ T cells.

Another factor that impedes effective anti-tumoral T cell responses is the failure to migrate into the tumor tissue (*46*, *47*). Whereas β1 (CD29) and β2 integrins appear not to be required for T cells migration into at least some tumor types (*50*), the expression of metalloproteinases is critical for this process (*51*). Intriguingly, we found that IFNG/DP producers also express higher transcript levels of the metalloproteinase ADAM15 and ADAM9, which could potentially support their intra-tumoral migratory capacity.

In summary, our data show that RNA-seq and Mass spectrometry analysis after intracellular cytokine staining is a powerful tool to define distinct T cell populations. This allowed us to uncover the differential make-up of $CD8^+$ T cell responses that is conserved and sustained marked by CD29. Not only our findings may help improve T cell products for therapeutic purposes in the future, they provide a robust marker to study cytotoxic T cells.

MATERIAL AND METHODS

T cell activation

Peripheral blood mononuclear cells (PBMCs) from healthy volunteers were isolated by Lymphoprep (density gradient separation; StemCell) and stored in liquid nitrogen until further use. The study was performed according to the Declaration of Helsinki (seventh revision, 2013). Written informed consent was obtained (Sanquin, Amsterdam, The Netherlands).

For RNA-seq and MS analysis, cryopreserved blood from 3 independently obtained pools of 40 blood donors depleted for monocytes was used. $CD8^+$ T cells were enriched to a purity of ±80% with the CD8 MACS positive isolation kit (Miltenyi). Validation essays were performed with cryopreserved blood from individual donors without MACS isolation.

T cells were activated as previously described (*5*). Briefly, non-tissue culture treated 24-well plates (Corning, USA) were pre-coated overnight at 4°C with 4µg/mL rat α-mouse IgG2a (MW1483, Sanquin) in PBS. Plates were washed and coated for >2h with 1µg/mL αCD3 (clone Hit3a, eBioscience) at 37°C. $0.8x10^6$ $CD8^+$ enriched PBMCs/well were seeded with 1µg/mL soluble αCD28 (clone CD28.2, eBioscience) in 1mL IMDM (LONZA) supplemented with 8% Fetal Calf Serum, 100U/mL Penicillin, 100µg/mL streptomycin and 2mM L-glutamine. After 48h of incubation at 37°C, 5% CO_2, cells were harvested, washed, and further cultured in standing T25/75 tissue culture flasks (Thermo-Scientific) at a density of $0.8x10^6$/mL, supplemented with 10ng/mL human recombinant IL-15 (Peprotech). Medium was refreshed every 5-7days. Alternatively, cells were cultured without cytokine or, with human recombinant IL-2 (50IU/mL), IL-6 (10ng/mL), IL-9 (10ng/mL), IL-11 (10ng/mL), IL-21 (50ng/mL).

Preparation of activated and fixed T cells for RNA-seq and MS analysis

T cells were activated with 10ng/mL PMA and 1μM Ionomycin (Sigma-Aldrich) for 4h in the presence of 1:1000 monensin (eBioscience). Cells were stained in sterile PBS for 30min at 4°C with anti-CD8α (RPA-T8, BD) and Near-IR live-dead marker (ThermoFisher). Intracellular cytokine staining (ICCS) was performed with CytoFix-CytoPerm kit (BD) according to the manufacturer's protocol. Prior to use, CytoFIX was pre-incubated on ice for 2min with 40IU/mL murine RNAse inhibitors (M0314, NEB). After 15min of incubation at 4°C, cells were washed once with 1xCytoPerm that was pre-incubated with 40IU/mL RNAse inhibitors. Cells were then incubated in 1xCytoPerm (pre-incubated with RNAse inhibitors) for 15 min at 4°C. Cells were then washed and re-suspended in Sorting buffer (2xSSC: 300mM NaCl, 30mM sodium citrate, 80IU/mL RNAse inhibitor). Antibodies directed against IL-2 (MQ1-17H12) and IFN-γ (4S.B3) were incubated overnight in Sorting buffer at 4°C. Cells were washed once and re-suspended in Sorting buffer prior to FACS-sorting.

To evaluate the protein recovery of fixed samples compared to fresh samples, we used $5x10^6$ PBMCs treated with CytoFix-CytoPerm kit (BD) according to the manufacturer's protocol or left in PBS. The samples were washed once with PBS, snap-frozen in liquid nitrogen, and used for MS analysis.

FACS-sorting

The pre-cooled FACS Aria III (BD) sorter was washed once with 70% Ethanol, followed by a 5min wash with Cleaning buffer (20xSSC (3M NaCl, 300mM sodium citrate, 400IU/mL RNAse inhibitors). Sorted cells (excluding cell doublets and dead cells) were collected in 500μL Sorting buffer. Cells were spun for 20min at 4°C at 4000rpm (2820g). The cell pellet was transferred in

1mL Sorting buffer in lo-Bind tubes (Eppendorf) and spun at 10.000g for 5min in table-top centrifuge (Eppendorf). Supernatant was removed, and pellet was frozen at -80°C until further use. Live cells were FACS-sorted according to standard procedures. Antibody staining was performed in PBS + 1% FCS for 30min at 4°C for live-dead marker (Invitrogen), α-CD8 (RPA-T8, SK1), α-CD4 (RPA-T4), α-CD29 (Mar4, BD), α-CD38 (HIT2, Biolegend), and for murine α-TCRβ when selecting for MART1-TCR expressing T cells (H57-597). Cells were washed once with culture medium and sorted on a pre-cooled FACS ARIA III sorter. Sorted cells were collected in culture medium.

Flow-cytometry analysis

Cells were stained in PBS+1%FCS for live-dead marker and antibodies against: CD3 (UCHT1), CD4 (RPA-T4), CD8 (RPA-T8,SK1), CD29 (Mar4), CD38 (HIT2), CD55 (JS11), CD11b (ICRF44), CD27 (CLB-27), CD49a (TS2/7), CD49b (P1E6-C5), CD49d (9F10), CD39 (EbioA1), CD40L (24-31), CD45RA (HI100), CD5 (M1649), CD70 (Ki-24), CD74 (M-B741), CD80 (L307.4), CD81 (JS-81), CD103 (B-ly7), CD122 (Mik-B3), CD137 (4B4-1), CD226 (11-A8), CD360 (17A12), CCR4 (1G1), CX3CR1 (2A9-1) ,GITR (108-17), HLA-DR (LN3), ICOS (C398.4A), IL18R1 (H44), IL7R (HIL-7R-M21), KLRG1 (HP-3D9), lag3 (3DS223H), Ox40 (ACT35), PD-L1 (MIH1), for MART1-TCR selection murine TCR-B (H57-597). The cytokine production profile was determined by ICCS after activation with PMA-Ionomycin for 4h. Cells were prepared with CytoFIX-CytoPerm kit following manufacturer's protocol. Cells were stained with antibodies against IFN-γ (4S.B3), IL-2 (MQ1-17H12), TNF-α (MAb11), Granulysin (DH2), Granzyme A (CB9), Granzyme B (GB11), CCL3 (CR3M), CCL4 (FL34Z3L), CCL5 (21445),

Perforin (dG49). Cells were acquired on LSR II, Fortessa or Symphony (all BD) using FACS Diva v8.0.1 (BD). Data were analysed with FlowJo VX (TreeStar).

Sample preparation for RNA-sequencing

For validation assays, cell pellets were defrosted on ice, and resuspended in proteinase K digestion buffer (20 mM Tris-HCl (pH 8.0), 1 mM CaCl2, 0.5% sodium dodecyl sulfate) and 75 µg molecular grade proteinase K (Life-Technologies). After 1h incubation at 55°C, Trizol-chloroform extraction was performed with 500ul according to the manufacturer's protocol (Invitrogen).

For RNA-seq analysis, RNA was isolated with the RNAeasy-FFPE kit (Quiagen) according to the manufacturer's protocol, omitting the steps directed at deparaffinization. RNA was resuspended in RNAse-free water. RNA concentration was measured with nanodrop (ThermoFisher), and RNA integrity was determined with the RNA 6000 Nano assay on the Bioanalyser 2100 (Agilent) Sequencing was performed with stranded "FFPE-library preparation" (GenomeScan, Leiden, the Netherlands), including Ribosomal RNA depletion with NEB Next rRNA depletion kit (NEB# E6310). Sequencing was performed on Illumina HiSeq 4000, with an average depth of 21 million paired-end 150nt (PE150) reads.

Quantitative PCR analysis

For validation experiments, cDNA preparation was performed with random hexamers using Super Script III reverse transcription (Invitrogen) according to the manufacturer's protocol. RT-qPCR Primers were designed with the NCBI primer blast tool (*52*), and manufactured by Sigma. qPCR Primer pairs were used when standard curves showed an r^2>0.98. RT-qPCR was performed with Power SYBR-green (ThermoFisher) with the standard protocol (T_m = 60°C for 1 min) on a 7500

Real-time qPCR system (Applied Biosystems). RT-qPCR data were analysed with 7500 Software v2.3 (Applied Biosystems). Ct difference was calculated by subtracting the Ct values from fixed or fixed-reversed to the Ct values from unfixed sample and plotted using GraphPad PRISM. Primers: *β-Actin* (Fwd: AGAGCTACGAGCTGCCTGAC, Rev.: AGCACTGTGTTGGCGTACAG) *RPS18*: (Fwd: CAGAAGTGACGCAGCCCTCTA, Rev.: AGACAACAAGCTCCGTGAAGA), *GAPDH* (Fwd: GAGACTCGTGCAATGGAGATTCT, Rev.: ACCCTGTTGCTGTAGCCA). *IL2* and *IFNG* primers were previously described (*5*).

RNA-seq data analysis

Reads quality was inspected using fastQC version 0.11.7 (S. Andrews, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Reads were aligned with STAR version 2.5.0a (*53*) on the human genome hg38-release 92 from ENSEMBL (*54*). STAR was used to count reads per genes (option --quantMode GeneCounts) and DESeq2 (*55*) using Wald test with p-adjusted <0.01 and log2 fold change (LFC) >0.5 to isolate differentially expressed genes. Further annotation was obtained from Ensembl BioMart (*54*, *56*) to discriminate between different types of coding and non-coding genes. Reads per million mapped reads (RPM) were calculated by dividing the read counts per gene by the total amount of mapped reads.

Mass spectrometry analysis

Sample preparation was performed as previously described (*57*). Tryptic peptides were separated by nanoscale C18 reverse phase chromatography coupled on line to an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific) via a nanoelectrospray ion source (Nanospray Flex Ion Source, Thermo Scientific). The MS^2 ion count target was set to 1.5×10^4. Only precursors with

charge state 2–7 were sampled for MS^2. The instrument was run in top speed mode with 3s cycles. All data were acquired with Xcalibur software.

The RAW mass spectrometry files were processed with the MaxQuant computational platform, 1.5.2.8. Proteins and peptides were identified using the Andromeda search engine by querying the human Uniprot database (downloaded February 2015). Standard settings with the additional options match between runs, Label Free Quantification (LFQ), and razor + unique peptides for quantification were selected. The generated 'proteingroups.txt' table was filtered for potential contaminants, reverse hits and 'only identified by site' using Perseus 1.5.1.6. The LFQ values were normalized using a log2 transformation. Data were imported in R and processed with the DEP package (*58*), and filtered for proteins that were found in all biological replicates of at least one population. Filtered data were normalized and imputed using random draws from a Gaussian distribution centered around a minimal value (q=0.01). Differential protein expression was determined with DEP (which uses Limma) with p-adj. <0.05 and log2 fold change >0.5.

Single cell RNA-seq analysis:

Single cell RNA-seq datasets were reanalysed from Guo et al. (*33*) and Zheng et al (*38*). Count matrix was filtered for "PTC" for peripheral blood $CD8^+$ T cells, "NTC" for normal lung or liver $CD8^+$ T cells. Single cell toolkit (SCTK) or ASAP (for clustering of Naïve-like cells) was used scRNA-seq data analysis (*59*, *60*). First, the batch-effect from the different patients was corrected using ComBaT in SCTK (*61*). Cells were filtered for >1000 expressed genes per cell, keeping the 50% most expressed genes after zero removal. We then log-transformed and performed differential expression (DE) analysis (absolute Log fold change >=0.5 and p-adjusted <0.05). CD29 grouping was determined based on the "double peak" expression distribution of *ITGB1* (CD29), "PTC" from

blood $CD8^+$ (cut-off: 10reads/cell), "NTC" from Lung or Liver $CD8^+$ cells (cut-off: 32 reads/cell (lung) and 10 reads/cell (liver)) (see Fig S7). Unsupervised clustering was performed in ASAP, using k-means (set to k=4) on tSNE. For validation, naïve-like cells were compared to the remaining "non-naïve-like" cells with ASAP (using Limma, LFC>1 and p-adjusted < 0.05). All DE genes were used for volcano-plots, and Venn diagrams show the top50 most upregulated DE genes.

Filtering of differentially expressed gene for functional annotation

Genes encoding secreted proteins were extracted from https://www.proteinatlas.org/humanproteome/tissue/secretome (downloaded Jan. 2018), and differentially expressed genes (DEG), or proteins with an absolute(LFC)>1.5, were identified. Of note, we manually curated the list for secreted genes for obvious mis-annotations (histones, membrane, TF, RNA-binding, CD molecules, protein without protein evidence, collagen, nuclear, nucleus, ribosomal, mitochondrial). Similarly, genes and proteins encoding CD molecules were extracted with the corresponding gene and protein names (www.uniprot.org/docs/cdlis; downloaded July 2017), and DEG, or proteins with an absolute(LFC)>1.5, were identified.

Genes encoding for Transcription Factor were obtained from https://www.proteinatlas.org/humanproteome/proteinclasses (downloaded Jan. 2019), and filtered from the DEG. LncRNA were filtered from DEG, excluding protein-coding genes and TCR/IG pseudogenes. Genes of experimentally validated RNA-binding proteins were obtained from (*62–64*), and filtered from the DEG.

Survival analysis

Oncolink (*65*) was used to assess the effect of *CD8B* expression on different types of cancer. A Cox regression analysis for CD8B was perfomed and the –log(FDR) was used to identify responsive cancer types (FDR<0.01).

Datasets for SKCM (skin cutaneous melanoma) RNA-sequencing and patient survival were obtained from TCGA (www.cancer.gov/tcga) using UCSC Xena (xenabrowser.net/datapages/; BioRxiv: doi: 10.1101/326470v5). Gene signature of CD29+ CD8+ T cells (Figure 5G) was used to isolate the gene expression from the RNA-sequencing SKMC TCGA dataset (in TPM+1). TPM+1 counts were transformed into a gene-wise Z-score. The average of all Z-score was calculated per patient. Patients were then stratified in two groups (high or low CD29 signature expression) using the median of the Z-score average (as described in (*33*)). Patients were also stratified for $CD8^+$ T cell infiltrates using estimates obtained by CIBERSORT (*41*) on SKCM TPM+1 counts. Survminer R package was used to prepare and visualize the patient survival with Kaplan-Meier plot. Log-rank test was used to determine statistical differences in survival between group.

Generation of MART1-TCR expressing T cells

PBMCs from individual donors were activated for 48h with αCD3/αCD28 as described above. Cells were harvested and retrovirally transduced with the MART1-TCR, as previously described (*42*). Briefly, non-tissue cultured treated 24 well plates were coated with 50µg/mL Retronectin (Takara) overnight and washed once with 1mL/well PBS. 300µL/well viral supernatant was added to the plate, and plates were centrifuged for 1h at 4°C at 4000 rpm (2820g). $0.5x10^6$ T cells were added per well, spun for 10min at 1000rpm, and incubated overnight at 37°C. The following day,

cells were harvested and cultured in T25 flasks at a concentration of $0.8x10^6$ cells/mL for 6-8 days in presence of 10ng/mL rhIL-15.

Functional assays with MART1-TCR expressing T cells

MART1-TCR transduced $CD4^+$ and $CD8^+$ T cells were FACS-sorted based on TCR expression, and rested over-night in medium at 37°C. Cytokine production was determined, as described above, by ICCS after 6h of co-culture with HLA-$A2^+$ $MART1^{hi}$ Mel 526 (MART1+), or HLA-$A2^-$ $MART1^{lo}$ Mel 888 (MART1-) tumor cell lines (*43*), in a 3 to 1 effector to target (E:T) ratio. For CD107α staining, co-cultured cells were supplemented with anti-CD107a, Brefeldin A and Monensin, in the culture medium for 6h. Cells were subsequently measured by flow cytometry. For killing assays, total $CD4^+$ and $CD8^+$ MART1 TCR-engineered T cells were FACS-sorted, or sorted based on their CD29 and CD38 expression profile ($CD29^+$ and $CD38^+$). Tumor cells were labelled with 1.5μM CFSE for 10 min at 37°C in FCS-free medium, and washed 3 times with warm culture medium. $15x10^3$ tumor cells were co-cultured with MART1-TCR^+ FACS-sorted T cells for 20h, in a 3:1, 1:1 and 0.3:1 E:T ratio. Dead cells were quantified by flow-cytometry with Near IR live-dead marker on $CFSE^+$ tumor cells.

Statistical analysis and Data visualisation

Data generated with flow cytometry were analysed with paired t-test, repeated measurement ANOVA using GraphPad PRISM version 7. Differences were considered significant with a p-value < 0.05.

Plots were generated with ggplot2 version 3.0, DEP version 1.2.0, and with GraphPad. Heatmaps were generated with pHeatmap version 1.0.10 and DEP. Venn diagrams were generated with (http://bioinformatics.psb.ugent.be/webtools/Venn/).

SUPPLEMENTARY MATERIALS

Fig. S1: RNA recovery after fixation of T cells.

Fig. S2: Differential cytokine production and surface molecule expression profile of IL2 and IFNG producers

Fig. S3: Characterization of $CD29^+$ and $CD38^+$ T cells

Fig. S4: Ex vivo analysis of $CD29^+CD8^+$ T cells.

Fig. S5: CX3CR1 marks a subset of $CD29^+CD8^+$ T cells

Fig. S6: $CD29^+$MART1-specific T cells are superior cytokine producers and show higher killing capacity than $CD38^+$ T cells

Table S1: Differentially expressed genes and proteins between DN, IL2, IFNG, and DP populations.

Table S2: List of differentially expressed genes and proteins annotated as: secreted proteins, transcription factors and RNA-binding proteins.

Table S3: Differentially expressed genes between ITGB1high and low from scRNA-seq data.

REFERENCES

1. M. Donia, J. W. Kjeldsen, R. Andersen, M. C. W. Westergaard, V. Bianchi, M. Legut, M. Attaf, B. Szomolay, S. Ott, G. Dolton, R. Lyngaa, S. R. Hadrup, A. K. Sewell, I. M. Svane, PD-1+ polyfunctional T cells dominate the periphery after tumor-infiltrating lymphocyte therapy for cancer. *Clin. Cancer Res.* **23**, 5779–5788 (2017).
2. Q. Xue, E. Bettini, P. Paczkowski, C. Ng, A. Kaiser, T. McConnell, O. Kodrasi, M. F. Quigley, J. Heath, R. Fan, S. Mackay, M. E. Dudley, S. H. Kassim, J. Zhou, Single-cell multiplexed cytokine profiling of CD19 CAR-T cells reveals a diverse landscape of polyfunctional antigen-specific response. *J. Immunother. Cancer*. **5**, 1–16 (2017).
3. C. Ma, A. F. Cheung, T. Chodon, R. C. Koya, Z. Wu, C. Ng, E. Avramis, A. J. Cochran, O. N. Witte, D. Baltimore, B. Chmielowski, J. S. Economou, B. Comin-Anduix, A. Ribas, J. R. Heath, Multifunctional T-cell analyses to study response and progression in adoptive cell transfer immunotherapy. *Cancer Discov.* **3**, 418–429 (2013).
4. Q. Han, N. Bagheri, E. M. Bradshaw, D. A. Hafler, D. A. Lauffenburger, J. C. Love, Polyfunctional responses by human T cells result from sequential release of cytokines. *Proc. Natl. Acad. Sci.* **109**, 1607–1612 (2012).
5. B. P. Nicolet, A. Guislain, M. C. Wolkers, Combined Single-Cell Measurement of Cytokine mRNA and Protein Identifies T Cells with Persistent Effector Function. *J. Immunol.* **198**, 962–970 (2017).
6. S. Huang, Non-genetic heterogeneity of cells in development: more than just noise. *Development*. **136**, 3853–3862 (2009).
7. M. Frentsch, R. Stark, N. Matzmohr, S. Meier, S. Durlanik, A. R. Schulz, U. Stervbo, K. Jürchott, F. Gebhardt, G. Heine, M. A. Reuter, M. R. Betts, D. Busch, A. Thiel, CD40L expression permits CD8+ T cells to execute immunologic helper functions. *Blood*. **122**, 405–412 (2013).
8. D. Hamann, P. A. Baars, M. H. G. Rep, B. Hooibrink, S. R. Kerkhof-Garde, M. R. Klein, R. A. W. v. Lier, Phenotypic and Functional Separation of Memory and Effector Human CD8+ T Cells. *J.*

Exp. Med. **186**, 1407–1418 (1997).

9. K. L. Wong, F. C. Lew, P. A. MacAry, D. M. Kemeny, CD40L-expressing CD8 T cells prime CD8α+ DC for IL-12p70 production. *Eur. J. Immunol.* **38**, 2251–2262 (2008).

10. M. G. H. Hernandez, L. Shen, K. L. Rock, CD40-CD40 Ligand Interaction between Dendritic Cells and CD8+ T Cells Is Needed to Stimulate Maximal T Cell Responses in the Absence of CD4+ T Cell Help. *J. Immunol.* **178**, 2844–2852 (2007).

11. I. S. Okoye, S. Czieso, E. Ktistaki, K. Roderick, S. M. Coomes, V. S. Pelly, Y. Kannan, J. Perez-Lloret, J. L. Zhao, D. Baltimore, J. Langhorne, M. S. Wilson, Transcriptomics identified a critical role for Th2 cell-intrinsic miR-155 in mediating allergy and antihelminth immunity. *Proc. Natl. Acad. Sci.* **111**, E3081–E3090 (2014).

12. H. Nilsson, K. M. Krawczyk, M. E. Johansson., High salt buffer improves integrity of RNA after fluorescence-activated cell sorting of intracellular labeled cells. *J. Biotechnol.* **192**, 62–65 (2014).

13. J. R. Wiśniewski, A. Zougman, N. Nagaraj, M. Mann, Universal sample preparation method for proteome analysis. *Nat. Methods*. **6**, 359–362 (2009).

14. S. R. Shi, C. Liu, B. M. Balgley, C. Lee, C. R. Taylor, Protein extraction from formalin-fixed, paraffin-embedded tissue sections: Quality evaluation by mass spectrometry. *J. Histochem. Cytochem.* **54**, 739–743 (2006).

15. S. Magdeldin, T. Yamamoto, Toward deciphering proteomes of formalin-fixed paraffin-embedded (FFPE) tissues. *Proteomics*. **12**, 1045–1058 (2012).

16. J. A. Gomez, O. L. Wapinski, Y. W. Yang, J.-F. Bureau, S. Gopinath, D. M. Monack, H. Y. Chang, M. Brahic, K. Kirkegaard, The NeST Long ncRNA Controls Microbial Susceptibility and Epigenetic Activation of the Interferon-γ Locus. *Cell*. **152**, 743–754 (2013).

17. H. R. Gibbons, G. Shaginurova, L. C. Kim, N. Chapman, C. F. Spurlock, T. M. Aune, Divergent lncRNA GATA3-AS1 Regulates GATA3 Transcription in T-Helper 2 Cells. *Front. Immunol.* **9**, 1–9 (2018).

18. R. Newman, J. McHugh, M. Turner, RNA binding proteins as regulators of immune cell biology.

Clin. Exp. Immunol. **183**, 37–49 (2016).

19. M. Uhlén, L. Fagerberg, B. M. Hallström, C. Lindskog, P. Oksvold, A. Mardinoglu, Å. Sivertsson, C. Kampf, E. Sjöstedt, A. Asplund, I. M. Olsson, K. Edlund, E. Lundberg, S. Navani, C. A. K. Szigyarto, J. Odeberg, D. Djureinovic, J. O. Takanen, S. Hober, T. Alm, P. H. Edqvist, H. Berling, H. Tegel, J. Mulder, J. Rockberg, P. Nilsson, J. M. Schwenk, M. Hamsten, K. Von Feilitzen, M. Forsberg, L. Persson, F. Johansson, M. Zwahlen, G. Von Heijne, J. Nielsen, F. Pontén, Tissue-based map of the human proteome. *Science (80-.).* **347** (2015), doi:10.4324/9781315766188.

20. T. R. Mosmann, H. Cherwinski, M. W. Bond, M. A. Giedlin, R. L. Coffman, Two types of murine helper T cell clone. I. Definition according to profiles of lymphokine activities and secreted proteins. *J. Immunol.* **136**, 2348–57 (1986).

21. M. P. Piccinni, L. Beloni, C. Livi, E. Maggi, G. Scarselli, S. Romagnani, Defective production of both leukemia inhibitory factor and type 2 T- helper cytokines by decidual T cells in unexplained recurrent abortions. *Nat. Med.* **4**, 1020–1024 (1998).

22. M. A. Iñiguez, C. Punzón, M. Fresno, Induction of cyclooxygenase-2 on activated T lymphocytes: regulation of T cell activation by cyclooxygenase-2 inhibitors. *J. Immunol.* **163**, 111–9 (1999).

23. A. Curti, M. Ratta, S. Corinti, G. Girolomoni, F. Ricci, P. Tazzari, M. Siena, A. Grande, M. Fogli, S. Tura, R. M. Lemoli, Interleukin-11 induces Th2 polarization of human CD4+T cells. *Blood*. **97**, 2758–2763 (2001).

24. F. Cocchi, A. L. Devico, A. Garzino-demo, S. K. Arya, R. C. Gallo, P. Lussot, Identification the Major HIV-Suppressive Factors Produced by. *Science (80-.).* **270**, 1811–1815 (1995).

25. D. J. Auerbach, Y. Lin, H. Miao, R. Cimbro, M. J. DiFiore, M. E. Gianolini, L. Furci, P. Biswas, A. S. Fauci, P. Lusso, Identification of the platelet-derived chemokine CXCL4/PF-4 as a broad-spectrum HIV-1 inhibitor. *Proc. Natl. Acad. Sci.* **109**, 9569–9574 (2012).

26. E. Narni-Mancinelli, L. Campisi, D. Bassand, J. Cazareth, P. Gounon, N. Glaichenhaus, G. Lauvau, Memory CD8 + T cells mediate antibacterial immunity via CCL3 activation of TNF/ROI + phagocytes. *J. Exp. Med.* **204**, 2075–2087 (2007).

27. B. Bengsch, T. Ohtani, R. S. Herati, N. Bovenschen, K. M. Chang, E. J. Wherry, Deep immune profiling by mass cytometry links human T and NK cell differentiation and cytotoxic molecule expression patterns. *J. Immunol. Methods*. **453**, 3–10 (2018).

28. J. P. Böttcher, M. Beyer, F. Meissner, Z. Abdullah, J. Sander, B. Höchst, S. Eickhoff, J. C. Rieckmann, C. Russo, T. Bauer, T. Flecken, D. Giesen, D. Engel, S. Jung, D. H. Busch, U. Protzer, R. Thimme, M. Mann, C. Kurts, J. L. Schultze, W. Kastenmüller, P. A. Knolle, Functional classification of memory CD8 + T cells by CX 3 CR1 expression. *Nat. Commun.* **6** (2015), doi:10.1038/ncomms9306.

29. S. Sohen, D. M. Rothstein, T. Tallman, D. Gaudette, S. F. Schlossman, C. Morimoto, The functional heterogeneity of CD8+ cells defined by anti-CD45RA (2H4) and anti-CD29 (4B4) antibodies. *Cell. Immunol.* **128**, 314–28 (1990).

30. V. Pulko, J. S. Davies, C. Martinez, M. C. Lanteri, M. P. Busch, M. S. Diamond, K. Knox, E. C. Bush, P. A. Sims, S. Sinari, D. Billheimer, E. K. Haddad, K. O. Murray, A. M. Wertheimer, J. Nikolich-Žugich, Human memory T cells with a naive phenotype accumulate with aging and respond to persistent viruses. *Nat. Immunol.* **17**, 966–975 (2016).

31. V. Appay, P. R. Dunbar, M. Callan, P. Klenerman, G. M. A. Gillespie, L. Papagno, G. S. Ogg, A. King, F. Lechner, C. A. Spina, S. Little, D. V. Havlir, D. D. Richman, N. Gruener, G. Pape, A. Waters, P. Easterbrook, M. Salio, V. Cerundolo, A. J. McMichael, S. L. Rowland-Jones, Memory CD8+ T cells vary in differentiation phenotype in different persistent virus infections. *Nat. Med.* **8**, 379–385 (2002).

32. K. M. L. Hertoghs, P. D. Moerland, A. van Stijn, E. B. M. Remmerswaal, S. L. Yong, P. J. E. J. van de Berg, S. M. van Ham, F. Baas, I. J. M. ten Berge, R. A. W. van Lier, Molecular profiling of cytomegalovirus-induced human CD8+ T cell differentiation. *J. Clin. Invest.* **120**, 4077–4090 (2010).

33. X. Guo, Y. Zhang, L. Zheng, C. Zheng, J. Song, Q. Zhang, B. Kang, Z. Liu, L. Jin, R. Xing, R. Gao, L. Zhang, M. Dong, X. Hu, X. Ren, D. Kirchhoff, H. G. Roider, T. Yan, Z. Zhang, Global

characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat. Med.* **24**, 978–985 (2018).

34. T. Willinger, T. Freeman, M. Herbert, H. Hasegawa, A. J. McMichael, M. F. C. Callan, Human Naive CD8 T Cells Down-Regulate Expression of the WNT Pathway Transcription Factors Lymphoid Enhancer Binding Factor 1 and Transcription Factor 7 (T Cell Factor-1) following Antigen Encounter In Vitro and In Vivo. *J. Immunol.* **176**, 1439–1446 (2006).

35. C. Yang, A. Khanniche, J. R. Dispirito, P. Ji, S. Wang, Y. Wang, H. Shen, Transcriptome Signatures Reveal Rapid Induction of Immune-Responsive Genes in Human Memory CD8+ T Cells. *Sci. Rep.* **6**, 1–8 (2016).

36. F. A. Vieira Braga, K. M. L. Hertoghs, N. A. M. Kragten, G. M. Doody, N. A. Barnes, E. B. M. Remmerswaal, C. C. Hsiao, P. D. Moerland, D. Wouters, I. A. M. Derks, A. van Stijn, M. Demkes, J. Hamann, E. Eldering, M. A. Nolte, R. M. Tooze, I. J. M. ten Berge, K. P. J. M. van Gisbergen, R. A. W. van Lier, Blimp-1 homolog Hobit identifies effector-type lymphocytes in humans. *Eur. J. Immunol.* **45**, 2945–2958 (2015).

37. N. A. M. Kragten, F. M. Behr, F. A. Vieira Braga, E. B. M. Remmerswaal, T. H. Wesselink, A. E. Oja, P. Hombrink, A. Kallies, R. A. W. van Lier, R. Stark, K. P. J. M. van Gisbergen, Blimp-1 induces and Hobit maintains the cytotoxic mediator granzyme B in CD8 T cells. *Eur. J. Immunol.* **48**, 1644–1662 (2018).

38. C. Zheng, L. Zheng, J. K. Yoo, H. Guo, Y. Zhang, X. Guo, B. Kang, R. Hu, J. Y. Huang, Q. Zhang, Z. Liu, M. Dong, X. Hu, W. Ouyang, J. Peng, Z. Zhang, Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell*. **169**, 1342–1356.e16 (2017).

39. M. Nishimura, H. Umehara, T. Nakayama, O. Yoneda, K. Hieshima, M. Kakizaki, N. Dohmae, O. Yoshie, T. Imai, Dual Functions of Fractalkine/CX3C Ligand 1 in Trafficking of Perforin+/Granzyme B+ Cytotoxic Effector Lymphocytes That Are Defined by CX3CR1 Expression. *J. Immunol.* **168**, 6173–6180 (2002).

40. Q. Sun, C. Zhou, R. Ma, Q. Guo, H. Huang, J. Hao, H. Liu, R. Shi, B. Liu, Prognostic value of

increased integrin-beta 1 expression in solid cancers: a meta-analysis. *Onco. Targets. Ther.* **Volume 11**, 1787–1799 (2018).

41. A. M. Newman, C. L. Liu, M. R. Green, A. J. Gentles, W. Feng, Y. Xu, C. D. Hoang, M. Diehn, A. A. Alizadeh, Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods*. **12**, 453–457 (2015).

42. R. Gomez-Eerland, B. Nuijen, B. Heemskerk, N. van Rooij, J. H. van den Berg, J. H. Beijnen, W. Uckert, P. Kvistborg, T. N. Schumacher, J. B. A. G. Haanen, A. Jorritsma, Manufacture of Gene-Modified Human T-Cells with a Memory Stem/Central Memory Phenotype. *Hum. Gene Ther. Methods*. **25**, 277–287 (2014).

43. F. M. Marincola, P. Shamamian, T. B. Simonis, A. Abati, J. Hackett, T. O'Dea, P. Fetsch, J. Yannelli, N. P. Restifo, J. J. Mule, S. A. Rosenberg, Locus-Specific Analysis of Human Leukocyte Antigen Class I Expression in Melanoma Cell Lines. *J. Immunother.* **16**, 13–23 (1994).

44. S. L. Topalian, D. Solomon, S. A. Rosenberg, Tumor-specific cytolysis by lymphocytes infiltrating human melanomas. *J. Immunol.* **142**, 3714–25 (1989).

45. E. Becht, N. A. Giraldo, L. Lacroix, B. Buttard, N. Elarouci, F. Petitprez, J. Selves, P. Laurent-Puig, C. Sautès-Fridman, W. H. Fridman, A. de Reyniès, Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* **17**, 218 (2016).

46. B. Mlecnik, G. Bindea, H. K. Angell, P. Maby, M. Angelova, D. Tougeron, S. E. Church, L. Lafontaine, M. Fischer, T. Fredriksen, M. Sasso, A. M. Bilocq, A. Kirilovsky, A. C. Obenauf, M. Hamieh, A. Berger, P. Bruneval, J. J. Tuech, J. C. Sabourin, F. Le Pessot, J. Mauillon, A. Rafii, P. Laurent-Puig, M. R. Speicher, Z. Trajanoski, P. Michel, R. Sesboüe, T. Frebourg, F. Pagès, V. Valge-Archer, J. B. Latouche, J. Galon, Integrative Analyses of Colorectal Cancer Show Immunoscore Is a Stronger Predictor of Patient Survival Than Microsatellite Instability. *Immunity*. **44**, 698–711 (2016).

47. D. S. Chen, I. Mellman, Oncology meets immunology: The cancer-immunity cycle. *Immunity*. **39**,

1–10 (2013).

48. S. Zelenay, A. G. Van Der Veen, J. P. Böttcher, K. J. Snelgrove, N. Rogers, S. E. Acton, P. Chakravarty, M. R. Girotti, R. Marais, S. A. Quezada, E. Sahai, C. Reis E Sousa, Cyclooxygenase-Dependent Tumor Growth through Evasion of Immunity. *Cell*. **162**, 1257–1270 (2015).

49. J. P. Böttcher, E. Bonavita, P. Chakravarty, H. Blees, M. Cabeza-Cabrerizo, S. Sammicheli, N. C. Rogers, E. Sahai, S. Zelenay, C. Reis e Sousa, NK Cells Stimulate Recruitment of cDC1 into the Tumor Microenvironment Promoting Cancer Immune Control. *Cell*. **172**, 1022–1028.e14 (2018).

50. H. Salmon, K. Franciszkiewicz, D. Damotte, M. C. Dieu-Nosjean, P. Validire, A. Trautmann, F. Mami-Chouaib, E. Donnadieu, Matrix architecture defines the preferential localization and migration of T cells into the stroma of human lung tumors. *J. Clin. Invest.* **122**, 899–910 (2012).

51. T. Ahrends, A. Spanjaard, B. Pilzecker, N. Bąbała, A. Bovens, Y. Xiao, H. Jacobs, J. Borst, CD4+T Cell Help Confers a Cytotoxic T Cell Effector Program Including Coinhibitory Receptor Downregulation and Increased Tissue Invasiveness. *Immunity*. **47**, 848–861.e5 (2017).

52. J. Ye, G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen, T. L. Madden, Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*. **13**, 134 (2012).

53. A. Dobin, T. R. Gingeras, in *Current Protocols in Bioinformatics* (John Wiley & Sons, Inc., Hoboken, NJ, USA, 2015; https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4631051/), vol. 51, p. 11.14.1-11.14.19.

54. D. R. Zerbino, P. Achuthan, W. Akanni, M. R. Amode, D. Barrell, J. Bhai, K. Billis, C. Cummins, A. Gall, C. G. Girón, L. Gil, L. Gordon, L. Haggerty, E. Haskell, T. Hourlier, O. G. Izuogu, S. H. Janacek, T. Juettemann, J. K. To, M. R. Laird, I. Lavidas, Z. Liu, J. E. Loveland, T. Maurel, W. McLaren, B. Moore, J. Mudge, D. N. Murphy, V. Newman, M. Nuhn, D. Ogeh, C. K. Ong, A. Parker, M. Patricio, H. S. Riat, H. Schuilenburg, D. Sheppard, H. Sparrow, K. Taylor, A. Thormann, A. Vullo, B. Walts, A. Zadissa, A. Frankish, S. E. Hunt, M. Kostadima, N. Langridge, F. J. Martin, M. Muffato, E. Perry, M. Ruffier, D. M. Staines, S. J. Trevanion, B. L. Aken, F.

Cunningham, A. Yates, P. Flicek, Ensembl 2018. *Nucleic Acids Res.* **46**, D754–D761 (2018).

55. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

56. R. J. Kinsella, A. Kähäri, S. Haider, J. Zamora, G. Proctor, G. Spudich, J. Almeida-King, D. Staines, P. Derwent, A. Kerhornou, P. Kersey, P. Flicek, Ensembl BioMarts: A hub for data retrieval across taxonomic space. *Database*. **2011**, 1–9 (2011).

57. M. C. van Aalderen, M. van den Biggelaar, E. B. M. Remmerswaal, F. P. J. van Alphen, A. B. Meijer, I. J. M. ten Berge, R. A. W. van Lier, Label-free Analysis of CD8+T Cell Subset Proteomes Supports a Progressive Differentiation Model of Human-Virus-Specific T Cells. *Cell Rep.* **19**, 1068–1079 (2017).

58. X. Zhang, A. H. Smits, G. B. A. Van Tilburg, H. Ovaa, W. Huber, M. Vermeulen, Proteome-wide identification of ubiquitin interactions using UbIA-MS. *Nat. Protoc.* **13**, 530–550 (2018).

59. V. Gardeux, F. P. A. David, A. Shajkofci, P. C. Schwalie, B. Deplancke, ASAP: A web-based platform for the analysis and interactive visualization of single-cell RNA-seq data. *Bioinformatics*. **33**, 3123–3125 (2017).

60. D. F. Jenkins, T. Faits, E. Briars, S. C. Pro, S. Cunningham, J. D. Campbell, M. Yajima, W. E. Johnson, Interactive single cell RNA-Seq analysis with the Single Cell Toolkit (SCTK). *bioRxiv*, 329755 (2018).

61. C. Li, A. Rabinovic, W. E. Johnson, C. Li, A. Rabinovic, Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. **8**, 118–127 (2007).

62. A. Castello, B. Fischer, C. K. Frese, R. Horos, A.-M. Alleaume, S. Foehr, T. Curk, J. Krijgsveld, M. W. Hentze, Comprehensive Identification of RNA-Binding Domains in Human Cells. *Mol. Cell*. **63**, 696–710 (2016).

63. J. I. Perez-Perri, B. Rogell, T. Schwarzl, F. Stein, Y. Zhou, M. Rettel, A. Brosig, M. W. Hentze, Discovery of RNA-binding proteins and characterization of their dynamic responses by enhanced RNA interactome capture. *Nat. Commun.* **9**, 4408 (2018).

64. S. Gerstberger, M. Hafner, T. Tuschl, A census of human RNA-binding proteins. *Nat. Rev. Genet.* **15**, 829–845 (2014).

65. J. Anaya, OncoLnc: linking TCGA survival data to mRNAs, miRNAs, and lncRNAs. *PeerJ Comput. Sci.* **2**, e67 (2016).

ACKNOWLEDGMENTS

We would like to thank the blood donors for donation and E.Mul, M.Hoogenboezem and S.Tol for FACS sorting. We are grateful to B. Popovic and D. Amsen for critical reading of this manuscript.

FUNDING

This research was supported by the Landsteiner Foundation of Blood Transfusion Research, by the Dutch Science Foundation, and by the Dutch Cancer Society (LSBR-Fellowship 1373, VIDI grant 917.14.214, and KWF grant 10132 to M.C.W.).

AUTHOR CONTRIBUTIONS

B.P.N. and M.C.W. designed experiments; B.P.N., A.G. F.P.J.v.A. and M.v.d.B. performed experiments and analysed data, R.G.-E. and T.N.M.S. provided the MART1 TCR system. M.C.W. directed the study; B.P.N. and M.C.W wrote the manuscript.

COMPETING INTEREST

Authors declare no conflict of interest.

DATA AVAILABILITY

All sequencing data are deposited on NCBI GEO under the accession no: GSE125497, and all MS data were deposited on PRIDE PXD012874.

FIGURE LEGENDS

Figure 1: RNA-seq analysis on CD8$^+$ T cells after intracellular cytokine staining. (**A**) Scheme to isolate RNA and protein from cytokine-producing, fixed T cells. Primary human CD8$^+$ T cells were activated for 48h with αCD3/αCD28, and rested for 4 days in the presence of 10ng/mL rhIL-15. T cells were reactivated for 4h with 10ng/mL PMA and 1μM Ionomycin. Intracellular cytokine staining was performed in RNA-protecting buffers. Cells were FACS-sorted under RNA-protecting conditions. Total RNA and protein was recovered for RNA-seq and Mass Spectrometry analysis (fig. 2) by reversing formaldehyde cross-linking. *IL2*: IL-2 single producers; *IFNG*: IFN-γ single producers; *DP*: double positive: IL-2 and IFN-γ co-producers; *DN*: double negative for IL-2 and IFN-γ production. (**B-H**) RNA-seq analysis of CD8$^+$ T cells with a differential IL2 and/or IFN-γ production profile. (**B**) Gene biotypes (n=2349) that are differentially expressed between all 4 T cell populations. (**C**) Sample-wise Pearson correlation coefficient of differentially expressed genes (n=2349). (**D**) Heatmap of differentially expressed protein-coding genes (n=1984), numbers indicate k-means clusters (k=5). (**E**) Heatmap of differentially expressed transcription factors (TF) with >100 reads per million (rpm; n=134), (**F**) long non-coding RNAs (lncRNA) (n=351), and (**G**) RNA-binding proteins (RBP) with >10rpm (n=180). (E-G) Expression levels of biological replicates were averaged (n=3 per population).

Figure 2: Mass-spectrometry analysis on cytokine producing, fixed CD8$^+$ T cells. (**A**) Number of identified proteins by Mass-spectrometry analysis of FACS-sorted IL2, IFNG, DN and DP producers per sample (bars) and of all samples combined after filtering (n=3833, red line). (**B**) Sample-wise Pearson correlation coefficient of (n=81) differentially expressed proteins. Color-coding of cell populations as in (A). (**C**) Heatmap of differentially expressed proteins, numbers

indicate k-means clusters (k=9). (**D**) Venn diagram of differentially expressed proteins and genes (fig.1). Proteins that show differential expression in mRNA and protein are indicated in (D) in blue.

Figure 3: IL2 and IFNG producers express unique sets of secreted proteins

(**A-B**) Heatmap of differentially expressed genes (A, n=67) and proteins (B, n=7) that are annotated as secreted proteins (see mat/methods section). Expression levels of biological replicates were averaged (n=3 per population). (**C**) The production of secreted effector molecules identified in (A, B) was measured by intracellular cytokine staining in 4 new individual donors. Cells were activated with αCD3/αCD28 and rested as in fig. 1, and reactivated with PMA-ionomycin for 4h in the presence of monensin to measure cytokine production. The percentage of indicated effector protein-producing cells in the IL2, IFNG, DP, and DN population was determined by flow cytometry. Repeated measurement ANOVA with Tukey post-test was performed. (*$p<0.05$, ** $p<0.01$, *** $p<0.001$).

Figure 4: CD29 and CD38 identify IFNG and IL2 producers

(**A**) Heat map of differentially expressed CD molecules identified by MS analysis with LFC>1.5 (grey: not detected). Expression levels of biological replicates were averaged (n=3 per population). (**B**) CD29 (left) and CD38 (right) expression of $CD8^+$ DN, IL2, IFNG and DP producers, as determined by flow cytometry after 4h reactivation with PMA-Ionomycin (n=6 donors). (**C**) Left panel: representative CD29 and CD38 expression of $CD8^+$ T cells re-stimulated for 4h with PMA-Ionomycin. Right panel: IL-2 and IFN-γ production of total $CD8^+$ T cell population, and of T cells gated based on CD29 and CD38 expression (representative of 17 donors). (**D-E**) Rested $CD8^+$ T

cells were FACS-sorted based on CD29 and CD38 expression and re-activated for 2 days with αCD3/αCD28, harvested and rested for another 21 days with rhIL-15. (**D**) Representative CD29 and CD38 expression at day 23 post sort. (**E**) Representative dot plot of IFN- and IL-2 production by $CD29^+$ and $CD38^+$ FACS-sorted $CD8^+$ T cells that were reactivated for 4h with PMA-Ionomycin at day 23 post sort (representative of 4 donors).

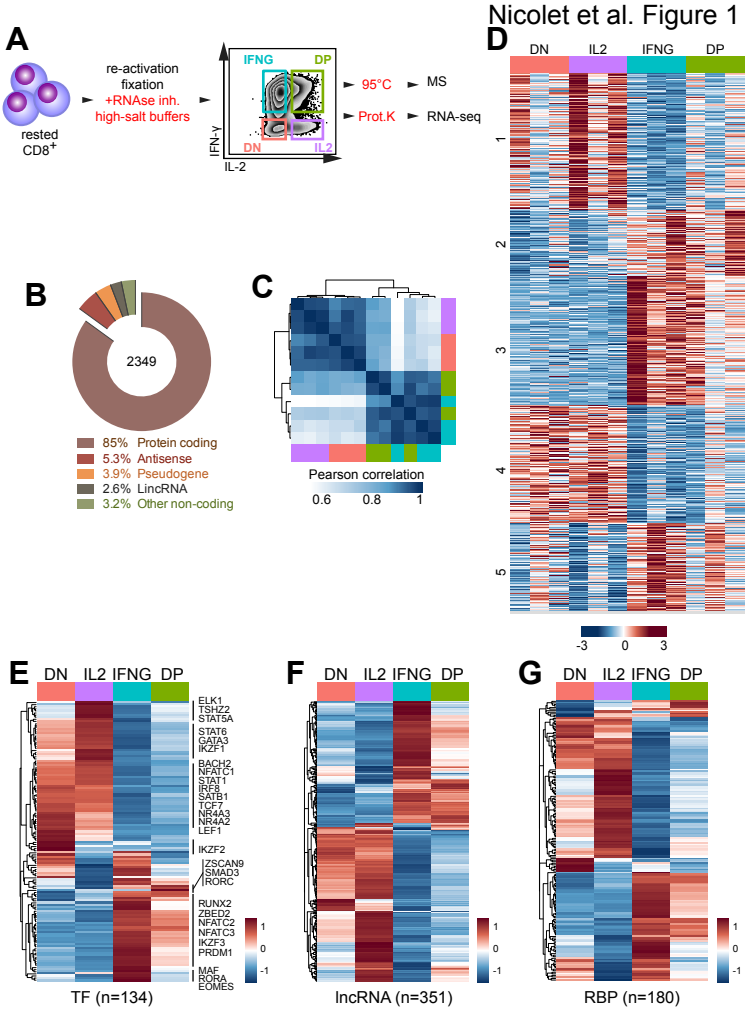
Figure 5: CD29 marks non-naïve $CD8^+$ T cells ex vivo with a cytotoxic signature (**A**) Representative expression profile of CD29 and CD38 of peripheral blood-derived $CD8^+$ T cells *ex vivo*. (**B**) Percentage of CD29+ $CD8^+$ T cell per subset, in blood, defined as: Naïve ($CD27^+CD45RA^+$), Memory ($CD27^+CD45RA^-$), Effector memory ($CD27^-CD45RA^-$) and Effector ($CD27^-CD45RA^+$, n=8 donors). (**C**) Representative dot plot of CD45RA and CD27 expression in CD8+ T cells activated with PMA-Ionomycin for 4h (left panel). T_N were gated and IFN-γ production was plotted against CD29 expression (right panel, representative of 5 donors). (**D**) *Ex vivo* CD49d expression of CD29+ (black) and CD29- (light grey) T_N CD8+ T cells from blood (representative of 4 donors). (**E-F**) Peripheral blood-derived $CD8^+$ T cells were activated for 6h with MHC-I restricted peptide pools for CMV (**E**) or EBV (**F**). (E-F) representative dot plots of IFN-γ and CD29 expression (left panels) and quantification (right panels, n=5). (**G**) Single-cell gene expression analysis of peripheral blood-derived $CD8^+$ T cells excluding naïve T cells (data retrieved from Single-cell RNA-seq data from(*33*), for details see fig. S5B-F). Left panel: tSNE projection of ITGB1 (CD29) expression in $CD8^+$ T cells excluding naïve T cells (n=1023, log2(TPM)). Right panel: volcano plot of differentially expressed genes in $ITGB1^{high}$ and $ITGB1^{low}$ $CD8^+$ T cells. ITGB1 gene expression cut-off was determined as depicted in fig. S4G. (**H**) Venn diagram of the core signature of $CD29^+CD8^+$ T cells from Blood and Lung (the top 50

up-regulated genes in each tissue). Differential expression with absolute(LFC)>0.5 and p-adjusted<0.05. (*p<0.05; **p<0.01).

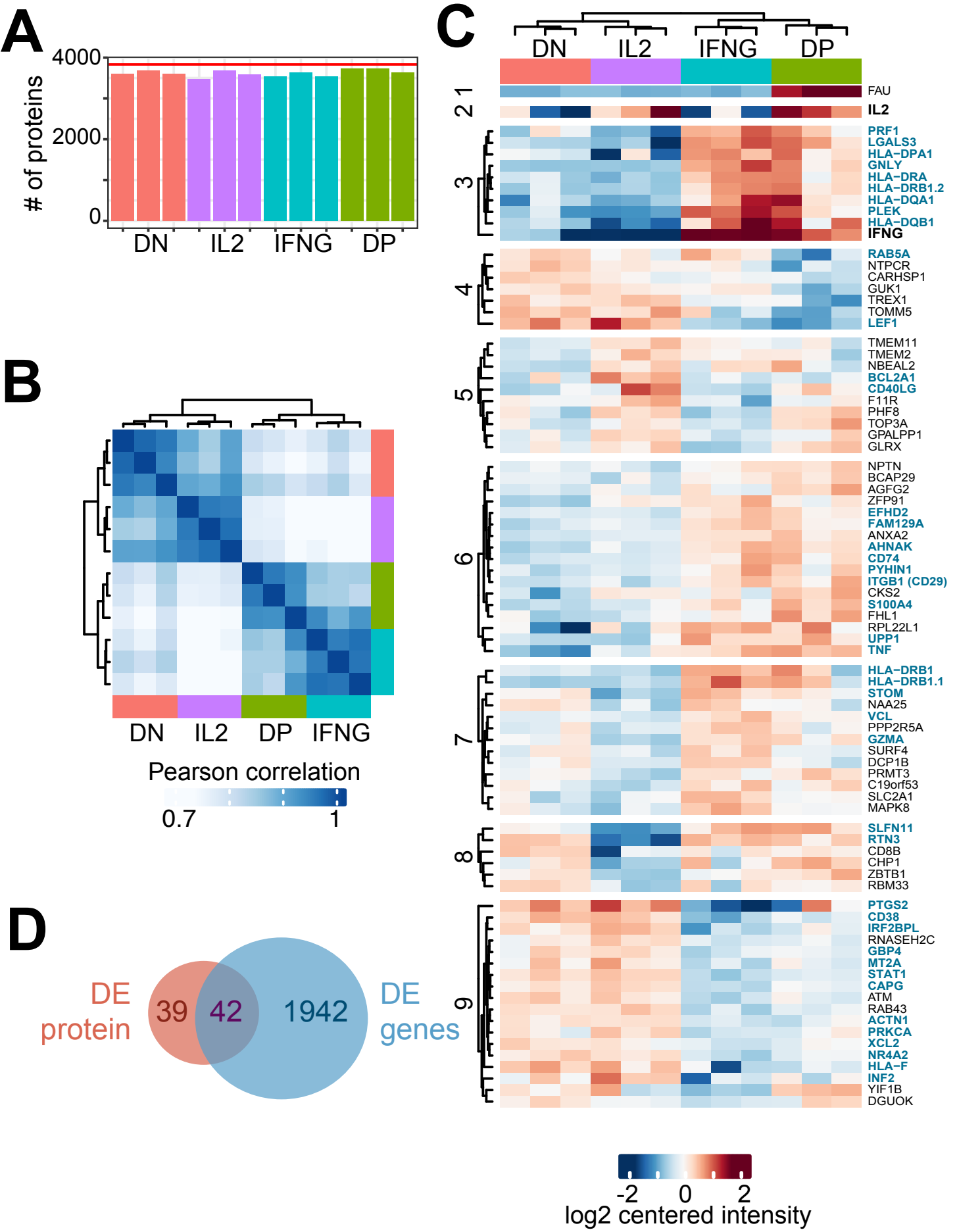
Figure 6: CD29 gene signature is associated with better survival, and CD29$^+$T cells are superior in killing tumor cells

(**A-B**) Kaplan-Meier plot showing the overall survival of total skin cutaneous melanoma (SKCM) patients (**A**) or patients with high CD8$^+$ T cell infiltrate estimates (**B**). Patients were stratified for a high or low CD29 gene signature as defined in Figure 5H (See method section). The lower panels depict the number of patient at risk for each time points. The indicated p-value was calculated using a log-rank test. (**C**) CD29$^+$ or CD38$^+$ MART1 TCR-engineered FACS-sorted CD8$^+$ T cells were cultured for 6h with MART1+ or MART1- tumor cells at an effector : target (E:T) ratio of 3:1. Representative dot plot of IFN-γ and IL-2 production as measured by intracellular cytokine staining. (**D**) Total, or CD29$^+$ and CD38$^+$ FACS-sorted MART1-TCR expressing CD8$^+$ T cells were co-cultured with CFSE-labeled MART1+ (upper panel) or MART1- (lower panel) tumor cells for 20h at a 3:1 E:T ratio. Representative measurements of dead tumor cells as determined by Near IR live-dead marker. (**E)** Tumor cell killing at 3:1 E:T ratio (n=6, repeated measurement ANOVA with Tukey post-test was performed). (**F**) Correlation of tumor killing of total CD8$^+$ MART1-TCR expressing T cells with the percentage of CD29$^+$CD8$^+$ T cells present in the T cell product (n=34, linear regression).

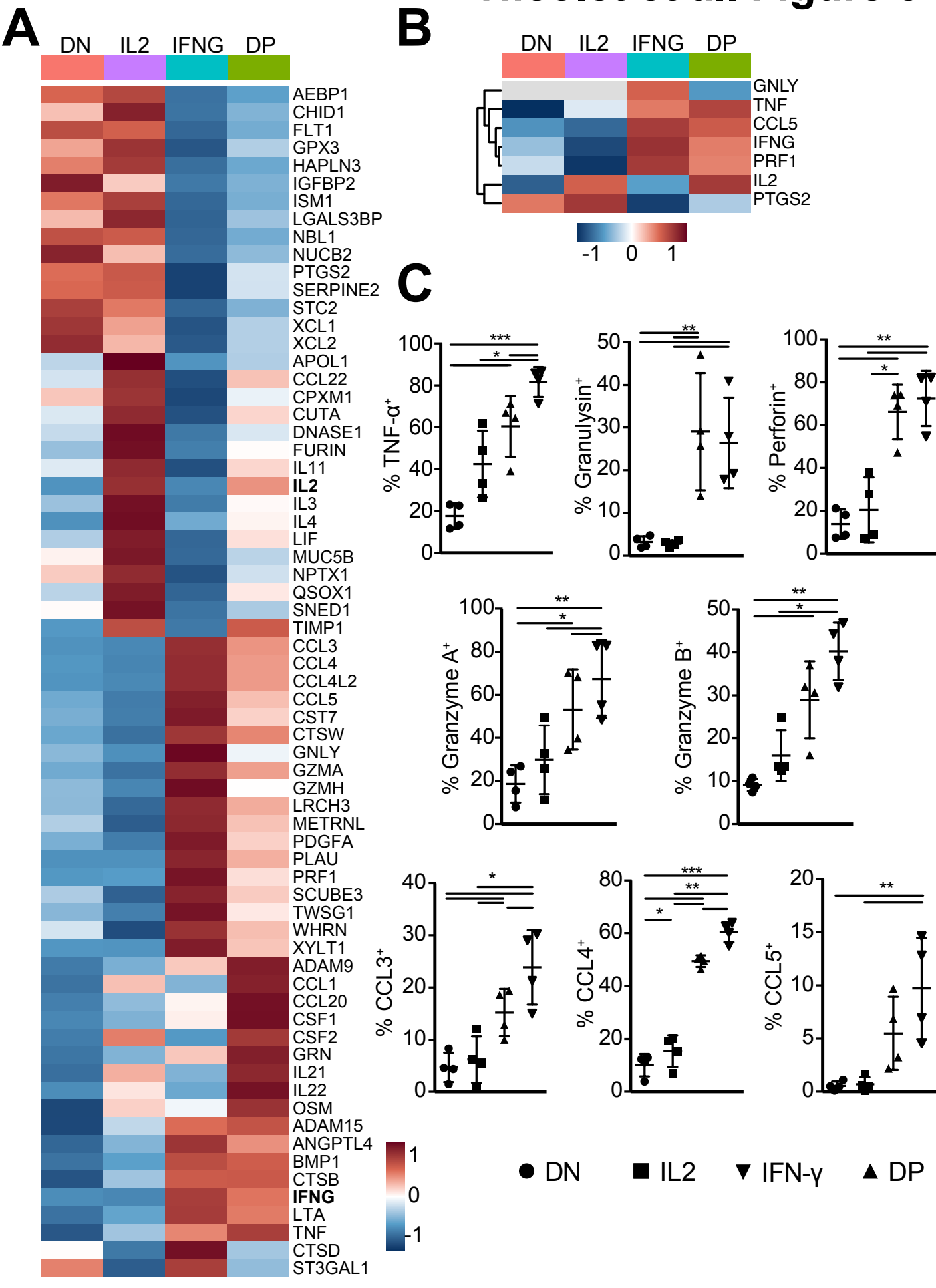
Nicolet et al. Figure 1



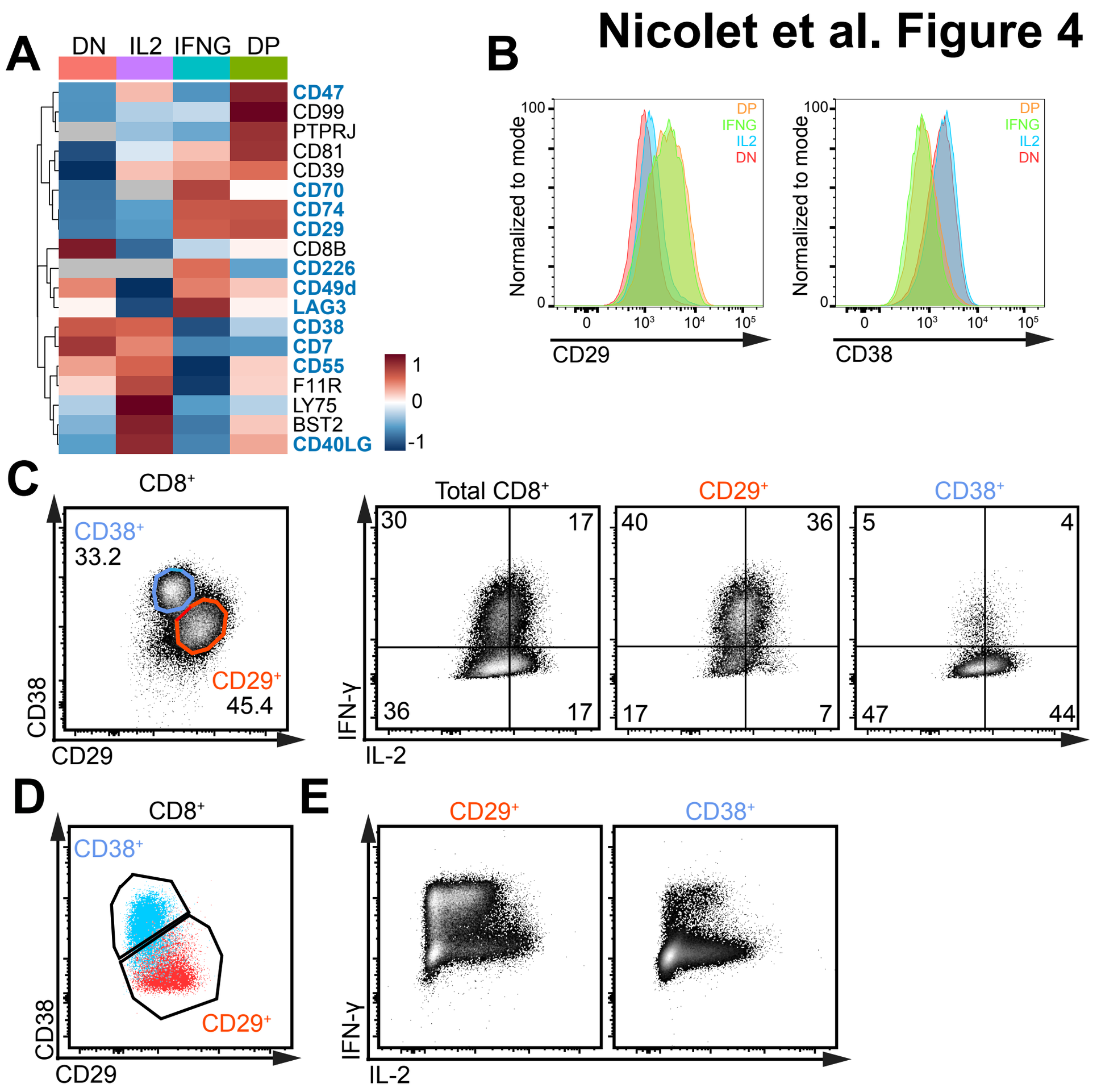
Nicolet et al. Figure 2



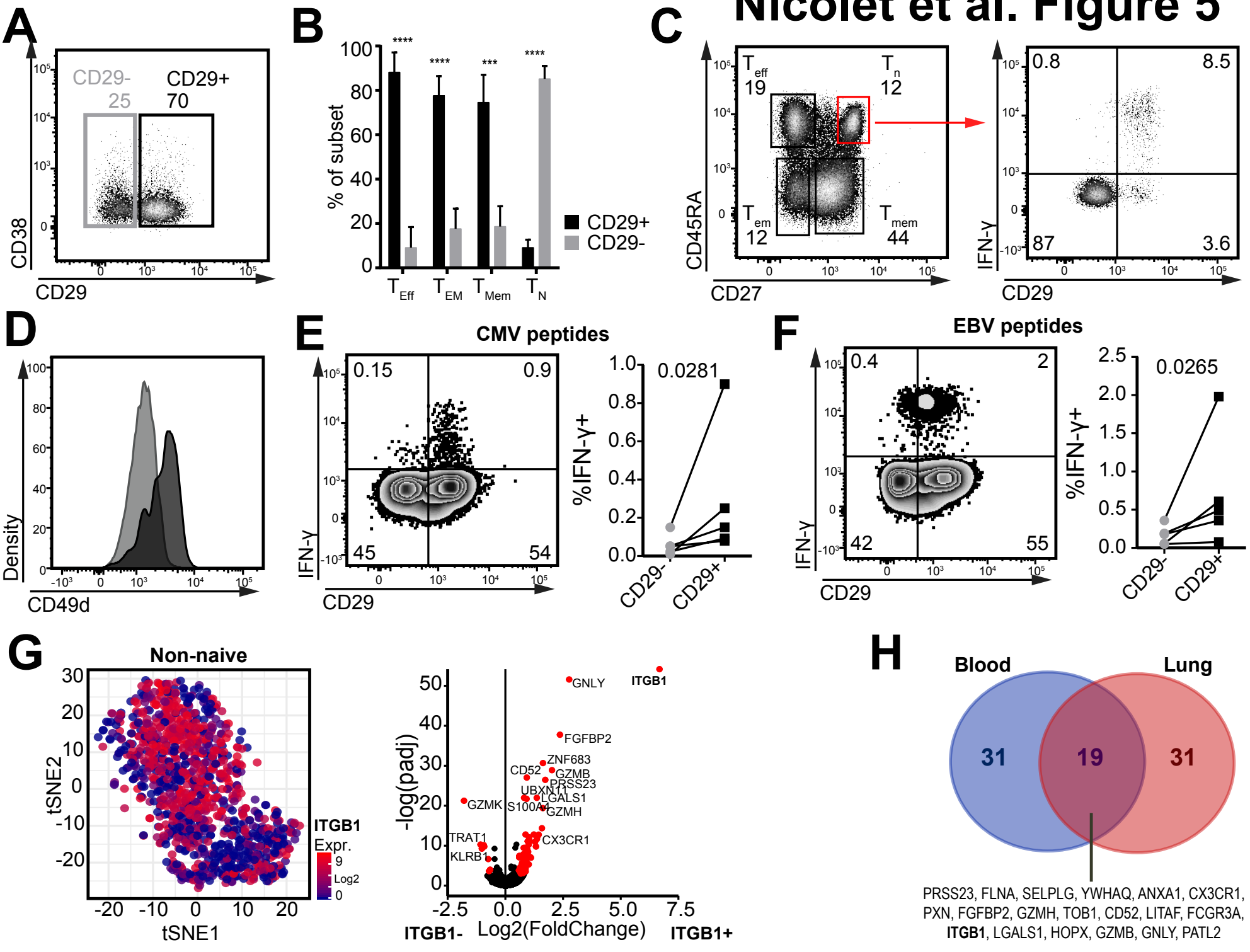
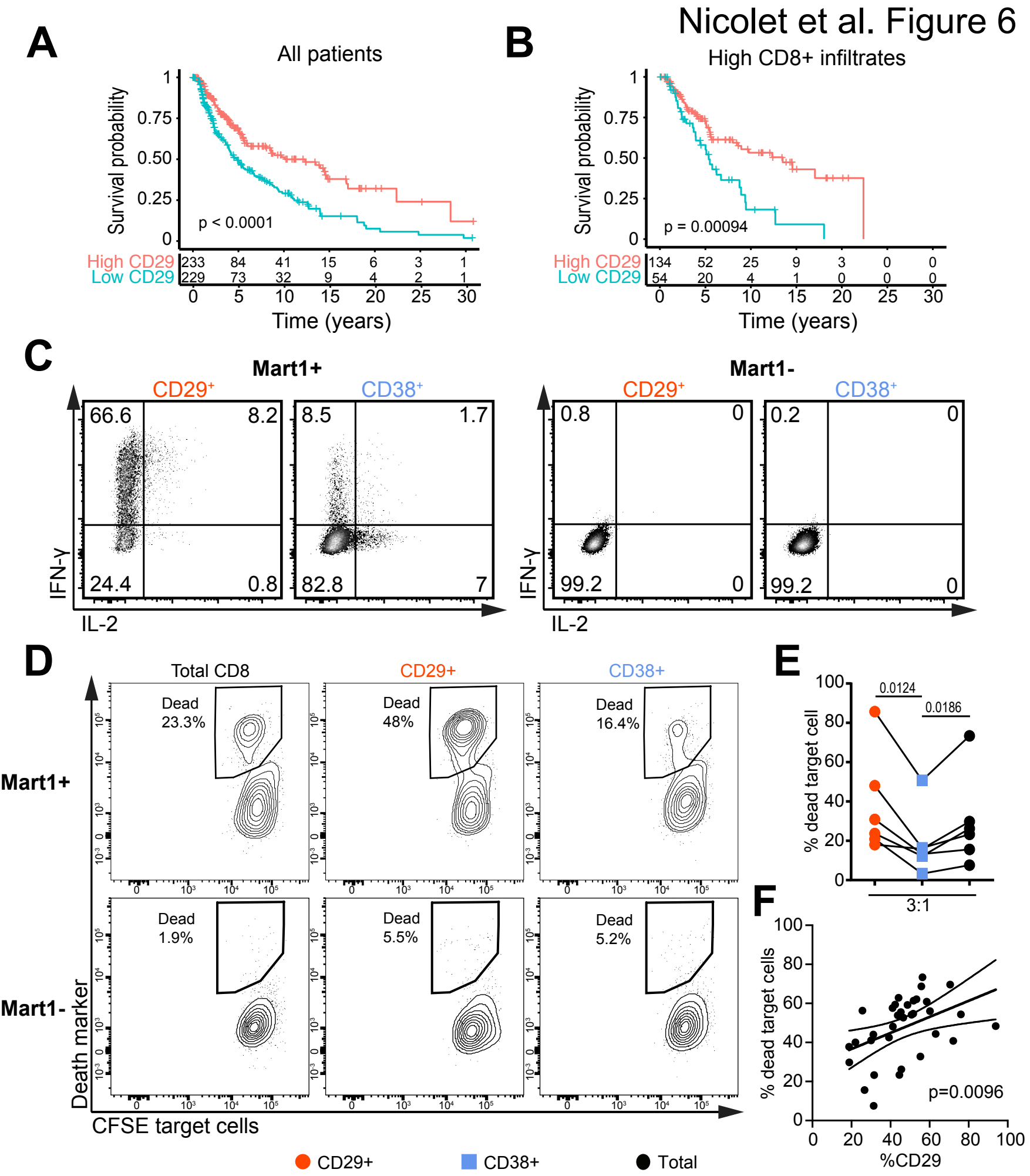
Nicolet et al. Figure 3



Nicolet et al. Figure 4



Nicolet et al. Figure 5



Nicolet et al. Figure 6