# 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, 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

***** To whom correspondence should be addressed. 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.

**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 RNA-sequencing and proteomics analysis on cytokine-producing fixed cells revealed that IL-$2^+$ T cells produce helper cytokines, and that IFN-$\gamma^+$ T cells produce cytotoxic molecules. IFN-$\gamma^+$ 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 selecting for $CD29^+$ T cells could boost the anti-tumoral activity of T cell therapeutics.

## 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, 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 killing capacity of 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, while only a small proportion 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 $CD8^+$ 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 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. Lastly, it remains to be determined if the differential cytokine production is imprinted during *in vitro* T cell culture, a feature that is critical when one generates T cell products for therapeutic purposes.

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 us to efficiently isolate RNA and protein from fixed T cells post intracellular cytokine staining. With this top-down approach we performed an unbiased RNA-seq and Mass spectrometry analysis 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. Using these two markers allowed us to differentiate T cells with a distinct cytokine production profile and cytotoxic potential already prior to T cell stimulation. Notably, the cytotoxic phenotype of $CD29^+$ T cells was also found in $CD4^+$ T cells. We therefore conclude that one can enrich for cytotoxic $CD8^+$ T cells prior to infusion, which should help boost the anti-tumoral effects of adoptive T cell therapies.

## RESULTS

### Efficient retrieval of RNA and protein from T cells after intracellular cytokine staining.

To determine whether and how the gene expression profile of IFN-γ producing T cells differed from IL-2 producing T 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 retrieval 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 (Supplementary 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). 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 (Supplementary Fig. S1A, B). Importantly, the RNA purity and integrity was compatible with RNA-seq analysis, as revealed by nanodrop analysis and RNAnano Chip assay (Supplementary Fig. S1C).

### 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 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 (Supplementary Table 1). 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).

Principal component analysis (PCA) and Pearson correlation of the differentially expressed genes revealed that the three biological replicates from each cytokine-producing T cell population clustered together (Fig. 1C, D). Interestingly, IFNG cells clustered with DP cells, and IL2 cells clustered with DN cells (Fig. 1C, D). This close relationship of IFNG and DP cells and of IL2 with DN cells was also evident from gene cluster analysis using k-means (k=5), in particular in cluster 2 and 4 that contain 241 and 434 genes, respectively (Fig. 1E, Supplementary Table 1). Yet, cluster 1 with 507 genes was more specific for IL2 T cells, and for IFNG T cells cluster 3 and 5, with 474 and 326 genes, respectively (Fig. 1E, Supplementary Table 1).

We next determined whether IFNG and IL2 producers showed differential expression of gene modulators, like transcription factors and non-coding RNAs. The majority of the 134 differentially expressed transcription factors (TFs) were shared between DN and IL2 producers, and between IFNG and DP producers (Fig. 1F, Supplementary table 2). This included GATA3 and IRF8 for IL2/DN producers, and PRDM1 (Blimp-1), NFATc2 and NFATc3 for IFNG/DP producers (Supplementary Fig. S1D, Supplementary table 2). Conversely, Eomes and RORA showed highest gene expression specifically in IFNG producers, whereas DP producers expressed high levels of RORC (ROR-γt; Supplementary Table 2). A small cluster of TFs was specifically up-regulated in DN cells, which included IKZF2 (Helios; Supplementary Fig. S1D).

We also found 351 differentially expressed long non-coding RNAs (lncRNA, Fig 1G, Supplementary Table 2). Again, their expression profile revealed a close kinship of DN and IL2 producers, and of IFNG and DP producers. For instance, the IFN-γ promoting IFNG-AS1 (NEST)[13] was highly expressed in IFNG and in DP producers (Supplementary Fig. S1E). Conversely, GATA3-AS1 known to re-enforce the Th2 phenotype in $CD4^+$ T cells[14] showed highest gene expression in IL2 producers (Supplementary Fig. S1E).

Another important class of regulators that define RNA splicing, stability and translational control are RNA-binding proteins (RBP)[15]. Also the 180 differentially expressed RBPs showed similarities in IFNG/DP producers and in IL2/DN producers (Fig. 1H, Supplementary Table 2). Combined, these data indicate that IFNG and IL2 producers show a differential gene expression profile, which correlates with distinct regulatory gene expression.

**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 Mass spectrometry (MS). Also for MS analysis, the formaldehyde cross-linking needed to be reversed. We used the Filter Aided Sample Preparation (FASP) protocol[16], which is compatible with proteomic sample preparation from formalin fixed material because it contains the required 95°C heating step[17,18].

We identified in total 5995 proteins from the 3 biological replicates. Removal of possible contaminants and filtering for proteins present in each biological replicate of at least one of the four populations resulted in a total of 3833 quantified proteins (Fig 2A, red line), of which each sample contained over 3500 identified proteins (Fig. 2A, Supplementary Table 1). PCA and Pearson correlation showed clustering of IFNG cells with DP cells, and of IL2 cells with DN cells, albeit less stringently than with RNA-seq analysis (Fig. 2B, C). This clustering also became apparent when the 81 differentially expressed proteins were clustered with k-means in a heat map (9 clusters; Fig. 2D). Of note, 42 (52%) of the differentially expressed proteins overlapped with the differential gene expression profile, including IL-2 and IFN-γ (Fig. 2E). In conclusion, IL2 and IFNG producers display a differential gene and protein expression signature.

**IFN-γ producing T cells display a cytotoxic profile**

We next interrogated whether IFNG and IL2 producers differed also in the expression of other secreted effector molecules. To identify genes that encode secreted proteins, we selected differentially expressed genes from Fig. 1 that showed expression levels >100 RPM, and that were annotated as “secreted” in the

human protein atlas[19]. After removing nuclear components, as well as mitochondrial and ribosomal proteins (see methods), we found 67 differentially expressed genes encoding secreted proteins (Fig. 3A). For IL2 producers, this included the prototypic Th2 helper effector molecules IL-4, IL-3, and the leukemia inhibitory factor (LIF)[20,21], as well as the Th2-inducing cytokine IL-11 and prostaglandin-endoperoxide synthase 2 (PTGS2, or COX-2)[22,23] (Fig. 3A). Of these, IL-2 and PTGS2 were also found differentially expressed in the MS analysis (Fig. 3B).

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). Further, the chemokines CCL5 (RANTES), CCL3 (MIP1α), CCL4 (MIP1β), and the alternative splicing variant of CCL4, 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 (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 activation for 4h with PMA-ionomycin confirmed that IFNG/DP producers also produced 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 of cytokine/chemokine producing cells was higher, but also the production of cytokines/chemokines per cell, as measured by the geometric mean fluorescence intensity of the four different T cell populations (geo-MFI; Supplementary Fig. S2).

We therefore conclude that IFNG and DP $CD8^+$ T cells preferentially express cytotoxic effector molecules.

**CD29 and CD38 enrich for IFN-γ and IL-2 producing $CD8^+$ and $CD4^+$ T cells**

We next set out to determine which protein markers associate with the cytokine production profile. 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 T cells that produced IFN-γ and/or IL-2 (Fig. 4A). RNA-seq analysis identified a total of 72 differentially expressed surface markers (Supplementary Fig. S3A). 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 (Supplementary Fig. S3B). In contrast, CD29 (β1-integrin) that was detected in both MS and RNA-seq analysis (Supplementary Fig. S3B) showed significantly higher cell surface expression in IFNG-DP producers than in IL2-DN producers, as determined by its geo-MFI (Fig. 4B; Supplementary Fig. S3C). IL2 and DN producers, on the other hand, were significantly enriched for CD38 (cyclic ADP ribose hydrolase; Fig. 4B, Supplementary Fig. S3C). Of note, the same expression pattern of CD29 and CD38 was found on $CD4^+$ T cells that were IFNG/DP producers and IL2/DN producers (Supplementary Fig. S3D).

Selecting for the surface expression of CD29 and CD38 T cells substantially enriched for IFNG/DP producers and DN/IL2 producers in activated $CD8^+$ and $CD4^+$

T cells (Fig. 4C, D; Supplementary Fig. S4A, B). Also, the cytokine/chemokine profile of $CD29^+CD8^+$ T cells was similar to IFNG/DP producers. In addition, the production levels of TNF-α, Granzyme A, Granulysin, Perforin, CCL3, CCL4, and CCL5 were significantly higher in $CD29^+CD8^+$ T cells (Supplementary Fig. S4A). Because $CD29^+$ T cells also included DP producers and the $CD38^+$ T cell population contained the IL-2 single producing T cells, the overall IL-2 production between $CD29^+$ and $CD38^+$ $CD8^+$ and $CD4^+$ T cells was similar (Supplementary Fig. S4A-D).

The frequency of $CD29^+$ and $CD38^+$ T cells substantially varied between donors (Supplementary Fig. S4E). Importantly, because CD29 and CD38 expression was already detected prior to T cell stimulation (Supplementary Fig. S4F), these two markers can thus be used to enrich for IFNG-DP producers and DN-IL2 producers.

**$CD29^+$ and $CD38^+$ T cells maintain their phenotype**

We next tested whether the expression of CD29 and CD38, and the respective cytokine production profile was maintained upon culture. We FACS-sorted $CD29^+$ and $CD38^+$ $CD8^+$ T cells on day 6 after activation and cultured them for another 17 days. Not only CD29 and CD38 expression was maintained in both fractions (Fig. 5A, B), but also the cytokine production profile, with $CD29^+$ T cells producing substantially more IFN-γ and TNFα, and $CD38^+$ T cells primarily containing of IL-2 single producers (Fig. 5C, D, Supplementary Fig. S4G). Also $CD4^+$ T cells preserved the expression of CD29 and CD38 and the respective cytokine production profile in long term T cell cultures (Supplementary Fig. S4H). $CD29^+$ T cells also expanded slightly but significantly better than $CD38^+$ T cells, which was again true for $CD8^+$ and for $CD4^+$ T cells (Fig. 5E).

We then examined whether the culture conditions we used influenced the expression pattern of CD29. None of the tested cytokines showed differences in CD29 expression on $CD8^+$ T cells (Supplementary Fig. S4I, J). Combined, these data show that the differences between $CD29^+$ and $CD38^+$ T cells are intrinsic, and maintained over time in cell culture.

**scRNA-seq analysis reveals ex vivo cytotoxic signature of $CD29^+$ $CD8^+$ T cells**

We next investigated whether CD29 expression also correlates with a cytotoxic gene signature of $CD8^+$ T cells directly isolated from blood. Because CD38 expression was very limited on peripheral blood-derived $CD8^+$ T cells (Fig. 6A), we compared $CD29^+$ with $CD29^-$ $CD8^+$ T cells. Also *ex vivo*, $CD29^+$ T cells were effective cytokine producers upon PMA-Ionomycin stimulation (Supplementary Fig. S5A). In concordance with the described expression profile of cytotoxic molecules[27,28], CD29 expression was found throughout T cell differentiation, but the highest percentage of $CD29^+$ T cells was detected in $CD45RA^+CD27^-$ Effector $CD8^+$ T cells ($T_{Eff}$), followed by $CD45RA^-CD27^-$ Effector Memory ($T_{EM}$) and $CD45RA^-CD27^+$ Memory $CD8^+$ T cells ($T_M$; Fig. 6B). Of the $CD45RA^+CD27^+$ naïve T cells ($T_N$), only 6.6±4.2% expressed CD29 (Fig. 6C).

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*[29]. Because of the predominant expression of CD29 in non-naïve $CD8^+$ T cells (Fig. 6C), we focused our analysis on $T_{Eff}$, $T_{EM}$, and $T_M$ cells. Unsupervised clustering followed by differential expression analysis identified T cells with a $T_N$ transcriptome signature[30,31], consisting of high *LEF1*, *CCR7*, *SELL* gene

expression, and low *CCL5*, *GZMB*, *PRF1* gene expression (Supplementary Fig. S5B, C). *CCL5* gene expression highly correlated with non-naïve T cells and was thus used to enrich for those for downstream analysis (Supplementary Fig. S5D-F).

Also in non-naïve $CD8^+$ T cells, *ITGB1* (CD29) gene expression was not uniform (Fig. 6D, left panel). Comparing the gene expression profile of non-naïve $CD8^+$ T cells with high or low *ITGB1* gene expression (Supplementary Fig. S6A) revealed 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[32] (Fig. 6D right panel, Supplementary Fig. S6B). Conversely, in concordance with previously described gene expression profiles[27,33], $ITGB1^{low}$ $CD8^+$ T cells associated with high gene expression levels of *GZMK* (Granzyme K), and the inhibitory receptor *KLRB1* (CD161; Fig. 6D right panel, Supplementary Table 3).

scRNA-seq analysis of $CD8^+$ T cells from lung tissue[29] also revealed a non-uniform expression of the *ITGB1* transcript (Fig. 6E, left panel), and *ITBG1* expression associated with a cytotoxic gene expression profile (Fig. 6E right panel; Supplementary Fig. S6C, D). 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[34], 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. 6F, G). Thus, *ITGB1* gene expression may differentially distinguish human T cell populations that correlate in blood and in tissues with a cytotoxic gene expression profile.

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

To determine if the difference in behavior of CD29$^+$ and CD38$^+$ T cells is also observed with a physiological stimulus, we transduced T cells with the codon-optimized MART1-TCR that recognizes the HLA-A*0201 restricted 26-35 epitope of MART1[35]. FACS-sorted CD29$^+$ and CD38$^+$ MART1 TCR-engineered CD8$^+$ 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-)[36,37]. MART1 TCR-engineered CD29$^+$ T cells produced substantially higher levels of IFN-$\gamma$ and TNF-$\alpha$, and Granulysin and CD107$\alpha$ in response to MART1 expressing tumor cells than MART1 TCR-engineered CD38$^+$ T cells (Fig 7A-C). Again, the superior production of IFN-$\gamma$ and TNF-$\alpha$ was conserved in CD29-expressing CD4$^+$ T cells (Supplementary Fig. S7A). In line with our previous results (Fig. 4B), CD29$^+$ T cells were the major source of IL-2$^+$IFN-$\gamma^+$ double producing cells in both CD8$^+$ and CD4$^+$ MART1 TCR-engineered T cells (Fig. 7B). Accordingly, MART1 TCR-engineered CD38$^+$ T cells were enriched for IL-2 single producers (Fig. 7A, B, Supplementary Fig. S7A).

We then determined whether MART1 TCR-engineered CD29$^+$ T cells were also more potent in killing tumor cells than CD38$^+$ T cells. To this end, CFSE-labelled MART1+ and MART1- tumor cells were co-cultured for 20h with FACS-sorted MART1 TCR-engineered CD29$^+$ or CD38$^+$ or CD8$^+$ T cells alone. We found that MART1+ tumor cells were effectively killed by MART1 TCR-engineered T cells (Fig. 7D). Notably, in line with their cytotoxic expression profile (Fig. 6), MART1 TCR-engineered CD29$^+$ T cells were superior in killing MART1$^+$ tumor cells when compared to CD38$^+$ T cells (Fig. 7D, E). In line with previous reports[38], also we found

low, but significant, target cell killing by MART1 TCR-engineered $CD4^+$ T cells (Supplementary Fig. S7B). This was again primarily mediated by $CD29^+$ T cells (Supplementary Fig. S7B). In conclusion, CD29 expression identifies the T cell populations with the highest cytotoxic potential in response to tumor cells.

## DISCUSSION

Here, we present RNA-seq and Mass spectrometry analyses from formaldehyde-fixed, permeabilized primary 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.

With this method, 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. The subsequent identification of CD29 and CD38 allowed for the distinction of IFNG-DP producers from DN-IL2 producers, respectively, even prior to T cell stimulation. The division between highly cytotoxic $CD29^+$ T cells and $CD38^+$ T cells was observed in both $CD8^+$ and $CD4^+$ T cells. Notably, the cytotoxic core signature of $CD29^+$ T cells was not limited to *in vitro* T cell cultures. scRNA-seq analysis on blood, lung- and liver-derived $CD29^+CD8^+$ T cells also showed high gene expression of cytotoxic molecules. Comparable gene expression patterns were also described for CMV-specific cytotoxic $CD4^+$ T cells[39].

CD29 expression is primarily found in non-naive T cells[40]. We here reveal that CD29 in fact enriches for cytotoxic T cells, a feature that is maintained upon prolonged *in vitro* cultures. Importantly, selecting for $CD29^+$ T cells significantly improves the killing capacity of TCR-engineered T cell products at least *in vitro*. A cytotoxic gene

expression of T cells infiltrating tumors is associated with better patient survival[41,42]. Furthermore, previous studies showed that the cytokine production profile of CAR T cells was highly diverse[2]. In line with our findings that CD29 expression can be highly variable between donors, isolating $CD29^+$ T cells may thus help improve the anti-tumoral potential of genetically engineered T cell products.

Selecting for $CD29^+$ cytotoxic T cells to generate T cell products may also come with another 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[43]. In fact, IL2 and DN populations produce PTGS2 (COX-2) that was recently shown to suppress anti-tumoral immunity[44,45]. 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[42,43]. Whereas β1 (CD29) and β2 integrins appear not to be required for T cells migration into at least some tumor types[46], the expression of metalloproteinases is critical for this process[47]. 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 can be used to define T cell populations. This allowed us to uncover the differential make-up of $CD8^+$ T cell responses that is conserved

and sustained. Our findings may also help improve T cell products for therapeutic purposes in the future.

## 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.

### *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-TCR-B 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), 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[48], 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[49] on the human genome hg38-release 92 from ENSEMBL[50]. STAR was used to count reads per genes and DESeq2[51] 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[50,52] to discriminate between different types of coding and non-coding genes.

**Mass spectrometry data analysis**

Sample preparation was performed as previously described[53]. 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 $1.5 \times 10^4$. 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 transformed in log2 scale. Data were imported in R and processed with the DEP tool[54], 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.[29] and Zheng et al[34]. 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[55,56]. First, the batch-effect from the different patients was corrected using ComBaT in SCTK[57]. Cells were filtered for >1000 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 Supplementary 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 (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 clusters***

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 [58–60], and filtered from the DEG.

***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[35]. 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 [36], 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/).

## 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 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**) PCA analysis, (**D**) sample-wise Pearson correlation coefficient of differentially expressed genes (n=2349). (**E**) Heatmap of differentially expressed protein-coding genes (n=1984), numbers indicate k-means clusters (k=5). (**F**) Heatmap of differentially expressed transcription factors (TF) with >100 reads per million (rpm; n=134), (**G**) non-coding genes (n=351), and (**H**) RNA-binding proteins (RBP) with >10rpm (n=180). (F,G,H) 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**) PCA analysis, and (**C**) sample-wise Pearson correlation coefficient of (n=81) differentially expressed proteins. Color-coding of cell populations as in (A). (**D**) Heatmap of differentially expressed proteins, numbers indicate k-means clusters (k=9). (**E**) 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 a new donor cohort (n=4) after PMA-ionomycin activation for 4h in the presence of monensin. 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**) Geometric-mean fluorescence intensity (geoMFI) of 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 on CD29 and CD38 expression (representative of 17 donors). (**D**) CD29 and CD38 expression as in panel (C) for $CD4^+$ T cells (representative of 7 donors). In (B), repeated measures ANOVA with Tukey post-test was performed (*p<0.05; **p<0.01; *** p<0.001; **** p<0.0001).

**Figure 5: CD29 and CD38 populations are stable during in vitro cultures**

(**A-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. Representative CD29 and CD38 expression (**A**) and compiled data of 4 donors (**B**) at day 23 post sort (paired t-test). (**C**) Representative dot plot, and (**D**) compiled data of IFN-γ, IL-2 and TNF-α production by $CD29^+$ and $CD38^+$ FACS-sorted $CD8^+$ T cells that were reactivated for 4h with PMA-Ionomycin at day 23 post sort. (n=4, paired t-test). (**E**) Cells were enumerated at indicated time points post FACS-sort of $CD29^+$ or $CD38^+$ $CD8^+$ (top panel) and $CD4^+$ (bottom panel) T cells (n=4, ratio paired t-test). (n.s.: not significant; *p<0.05; **p<0.01).

**Figure 6: Single-cell analysis reveals cytotoxic signature of $CD29^+CD8^+$ T cells ex vivo**

(**A**) Representative expression profile of CD29 and CD38 of peripheral blood-derived $CD8^+$ T cells *ex vivo*. (**B**) Representative CD29 expression in $CD8^+$ T cell subsets in blood defined as: Naïve ($CD27^+CD45RA^+$), Memory ($CD27^+CD45RA^-$), Effector memory ($CD27^-CD45RA^-$) and Effector ($CD27^-CD45RA^+$). (**C**) Quantification of (B) (n=3). (**D**) 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[29], for details see Supplementary Fig. S7A-D). Left panel: tSNE projection of ITGB1 (CD29) expression in $CD8^+$ T cells excluding naïve T cells (n=1023). Right panel: volcano plot of differentially expressed genes in $ITGB1^{high}$ and $ITGB1^{low}$ $CD8^+$ T cells. ITGB1 gene expression cut-offs are depicted in Supplementary Fig. S7B. **(E**) Similar analysis as in (**D**) for Lung-derived $CD8^+$ T cells (data retrieved from[29]). ITGB1 gene expression cut-offs are depicted in Supplementary Fig. S7D). (**F**) Venn diagram depicting the core signature of $CD29^+CD8^+$ T cells from Blood and Lung (the top 50 up-regulated genes in each tissue). (**G**) Venn diagram of the top 50 up-regulated genes in $CD29^+CD8^+$ T cells from Lung and Liver (retrieved from[29,34]). Differential expression with absolute(LFC)>0.5 and p-adjusted<0.05. (*p<0.05; **p<0.01).

**Figure 7: $CD29^+$MART1-specific T cells are superior in killing tumor cells than $CD38^+$ T cells**

(**A**) $CD29^+$ 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. (**B**) Production of IFN-γ, IL-2, IFN-$\gamma^+$ of the IL-$2^+$ T cells, and of TNF-α. Granulysin levels were determined from non-activated T cells (n=4, paired t-test). (**C**) CD107α staining in response to MART1+ tumor cells (n=6, paired t-test). (**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 (effector to target) ratio. Representative measurement of dead tumor cells as determined by Near IR live-dead marker. (**E)**

Top panel: tumor killing at different effector : target ratios. Bottom panel: tumor killing of $CD29^+$ and $CD38^+$ MART1-TCR expressing T cells compared to total MART1-TCR expressing T cells (n=6, repeated measurement ANOVA with Tukey post hoc test. n.s.: not significant, *$p<0.05$; ** $p<0.01$).

## DATA AVAILABILITY

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

## 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.

## 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. 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.).

## COMPETING INTEREST

Authors declare no conflict of interest.

## REFERENCES

1. Donia, M. *et al.* PD-1+ polyfunctional T cells dominate the periphery after tumor-infiltrating lymphocyte therapy for cancer. *Clin. Cancer Res.* **23,** 5779–5788 (2017).
2. Xue, Q. *et al.* 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. Ma, C. *et al.* Multifunctional T-cell analyses to study response and progression in adoptive cell transfer immunotherapy. *Cancer Discov.* **3,** 418–429 (2013).
4. Han, Q. *et al.* Polyfunctional responses by human T cells result from sequential release of cytokines. *Proc. Natl. Acad. Sci.* **109,** 1607–1612 (2012).
5. Nicolet, B. P., Guislain, A. & Wolkers, M. C. Combined Single-Cell Measurement of Cytokine mRNA and Protein Identifies T Cells with Persistent Effector Function. *J. Immunol.* **198,** 962–970 (2017).
6. Huang, S. Non-genetic heterogeneity of cells in development: more than just noise. *Development* **136,** 3853–3862 (2009).
7. Frentsch, M. *et al.* CD40L expression permits CD8+ T cells to execute immunologic helper functions. *Blood* **122,** 405–412 (2013).
8. Hamann, D. *et al.* Phenotypic and Functional Separation of Memory and Effector Human CD8+ T Cells. *J. Exp. Med.* **186,** 1407–1418 (1997).
9. Wong, K. L., Lew, F. C., MacAry, P. A. & Kemeny, D. M. CD40L-expressing CD8 T cells prime CD8α+ DC for IL-12p70 production. *Eur. J. Immunol.* **38,** 2251–2262 (2008).

10. Hernandez, M. G. H., Shen, L. & Rock, K. L. 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. Okoye, I. S. *et al.* 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. Nilsson, H., Krawczyk, K. M. & Johansson., M. E. High salt buffer improves integrity of RNA after fluorescence-activated cell sorting of intracellular labeled cells. *J. Biotechnol.* **192,** 62–65 (2014).

13. Gomez, J. A. *et al.* The NeST Long ncRNA Controls Microbial Susceptibility and Epigenetic Activation of the Interferon-γ Locus. *Cell* **152,** 743–754 (2013).

14. Gibbons, H. R. *et al.* Divergent lncRNA GATA3-AS1 Regulates GATA3 Transcription in T-Helper 2 Cells. *Front. Immunol.* **9,** 1–9 (2018).

15. Newman, R., McHugh, J. & Turner, M. RNA binding proteins as regulators of immune cell biology. *Clin. Exp. Immunol.* **183,** 37–49 (2016).

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

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

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

19. Uhlén, M. *et al.* Tissue-based map of the human proteome. *Science (80-. ).* **347,** (2015).

20. Mosmann, T. R., Cherwinski, H., Bond, M. W., Giedlin, M. A. & Coffman, R. L. 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. Piccinni, M. P. *et al.* 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. Iñiguez, M. A., Punzón, C. & Fresno, M. Induction of cyclooxygenase-2 on activated T lymphocytes: regulation of T cell activation by cyclooxygenase-2 inhibitors. *J. Immunol.* **163,** 111–9 (1999).
23. Curti, A. *et al.* Interleukin-11 induces Th2 polarization of human CD4+T cells. *Blood* **97,** 2758–2763 (2001).
24. Cocchi, F. *et al.* Identification the Major HIV-Suppressive Factors Produced by. *Science (80-. ).* **270,** 1811–1815 (1995).
25. Auerbach, D. J. *et al.* 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. Narni-Mancinelli, E. *et al.* Memory CD8 + T cells mediate antibacterial immunity via CCL3 activation of TNF/ROI + phagocytes. *J. Exp. Med.* **204,** 2075–2087 (2007).
27. Bengsch, B. *et al.* 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. Böttcher, J. P. *et al.* Functional classification of memory CD8 + T cells by CX 3 CR1 expression. *Nat. Commun.* **6,** (2015).
29. Guo, X. *et al.* Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat. Med.* **24,** 978–985 (2018).
30. Willinger, T. *et al.* 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).
31. Yang, C. *et al.* Transcriptome Signatures Reveal Rapid Induction of Immune-Responsive Genes in Human Memory CD8+ T Cells. *Sci. Rep.* **6,** 1–8 (2016).
32. Vieira Braga, F. A. *et al.* Blimp-1 homolog Hobit identifies effector-type lymphocytes in humans. *Eur. J. Immunol.* **45,** 2945–2958 (2015).

33. Kragten, N. A. M. *et al.* Blimp-1 induces and Hobit maintains the cytotoxic mediator granzyme B in CD8 T cells. *Eur. J. Immunol.* **48,** 1644–1662 (2018).

34. Zheng, C. *et al.* Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell* **169,** 1342–1356.e16 (2017).

35. Gomez-Eerland, R. *et al.* Manufacture of Gene-Modified Human T-Cells with a Memory Stem/Central Memory Phenotype. *Hum. Gene Ther. Methods* **25,** 277–287 (2014).

36. Marincola, F. M. *et al.* Locus-Specific Analysis of Human Leukocyte Antigen Class I Expression in Melanoma Cell Lines. *J. Immunother.* **16,** 13–23 (1994).

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

38. Chhabra, A. *et al.* CD4+CD25- T cells transduced to express MHC class I-restricted epitope-specific TCR synthesize Th1 cytokines and exhibit MHC class I-restricted cytolytic effector function in a human melanoma model. *J. Immunol.* **181,** 1063–70 (2008).

39. Oja, A. E. *et al.* The transcription factor hobit identifies human cytotoxic CD4+T cells. *Front. Immunol.* **8,** 1–11 (2017).

40. Sohen, S. *et al.* The functional heterogeneity of CD8+ cells defined by anti-CD45RA (2H4) and anti-CD29 (4B4) antibodies. *Cell. Immunol.* **128,** 314–28 (1990).

41. Becht, E. *et al.* Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* **17,** 218 (2016).

42. Mlecnik, B. *et al.* Integrative Analyses of Colorectal Cancer Show Immunoscore Is a Stronger Predictor of Patient Survival Than Microsatellite Instability. *Immunity* **44,** 698–711 (2016).

43. Chen, D. S. & Mellman, I. Oncology meets immunology: The cancer-immunity cycle. *Immunity* **39,** 1–10 (2013).

44. Zelenay, S. *et al.* Cyclooxygenase-Dependent Tumor Growth through Evasion of Immunity. *Cell* **162,** 1257–1270 (2015).

45. Böttcher, J. P. *et al.* NK Cells Stimulate Recruitment of cDC1 into the Tumor Microenvironment Promoting Cancer Immune Control. *Cell* **172,** 1022–1028.e14 (2018).

46. Salmon, H. *et al.* 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).

47. Ahrends, T. *et al.* 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).

48. Ye, J. *et al.* Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* **13,** 134 (2012).

49. Dobin, A. & Gingeras, T. R. Mapping RNA-seq Reads with STAR. in *Current Protocols in Bioinformatics* **51,** 11.14.1-11.14.19 (John Wiley & Sons, Inc., 2015).

50. Zerbino, D. R. *et al.* Ensembl 2018. *Nucleic Acids Res.* **46,** D754–D761 (2018).

51. 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).

52. Kinsella, R. J. *et al.* Ensembl BioMarts: A hub for data retrieval across taxonomic space. *Database* **2011,** 1–9 (2011).

53. van Aalderen, M. C. *et al.* 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).

54. Zhang, X. *et al.* Proteome-wide identification of ubiquitin interactions using UbIA-MS. *Nat. Protoc.* **13,** 530–550 (2018).

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

56. Jenkins, D. F. *et al.* Interactive single cell RNA-Seq analysis with the Single Cell Toolkit (SCTK). *bioRxiv* 329755 (2018). doi:10.1101/329755

57. Li, C., Rabinovic, A., Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in

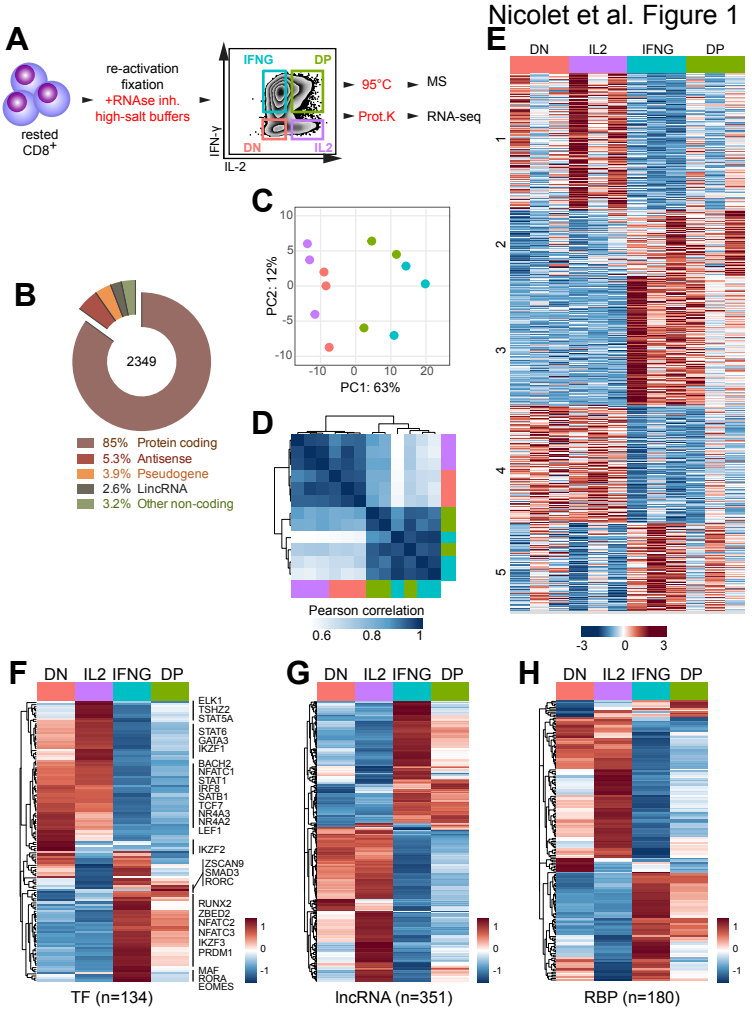
microarray expression data using empirical Bayes methods. *Biostatistics* **8,** 118–127 (2007).

58. Castello, A. *et al.* Comprehensive Identification of RNA-Binding Domains in Human Cells. *Mol. Cell* **63,** 696–710 (2016).

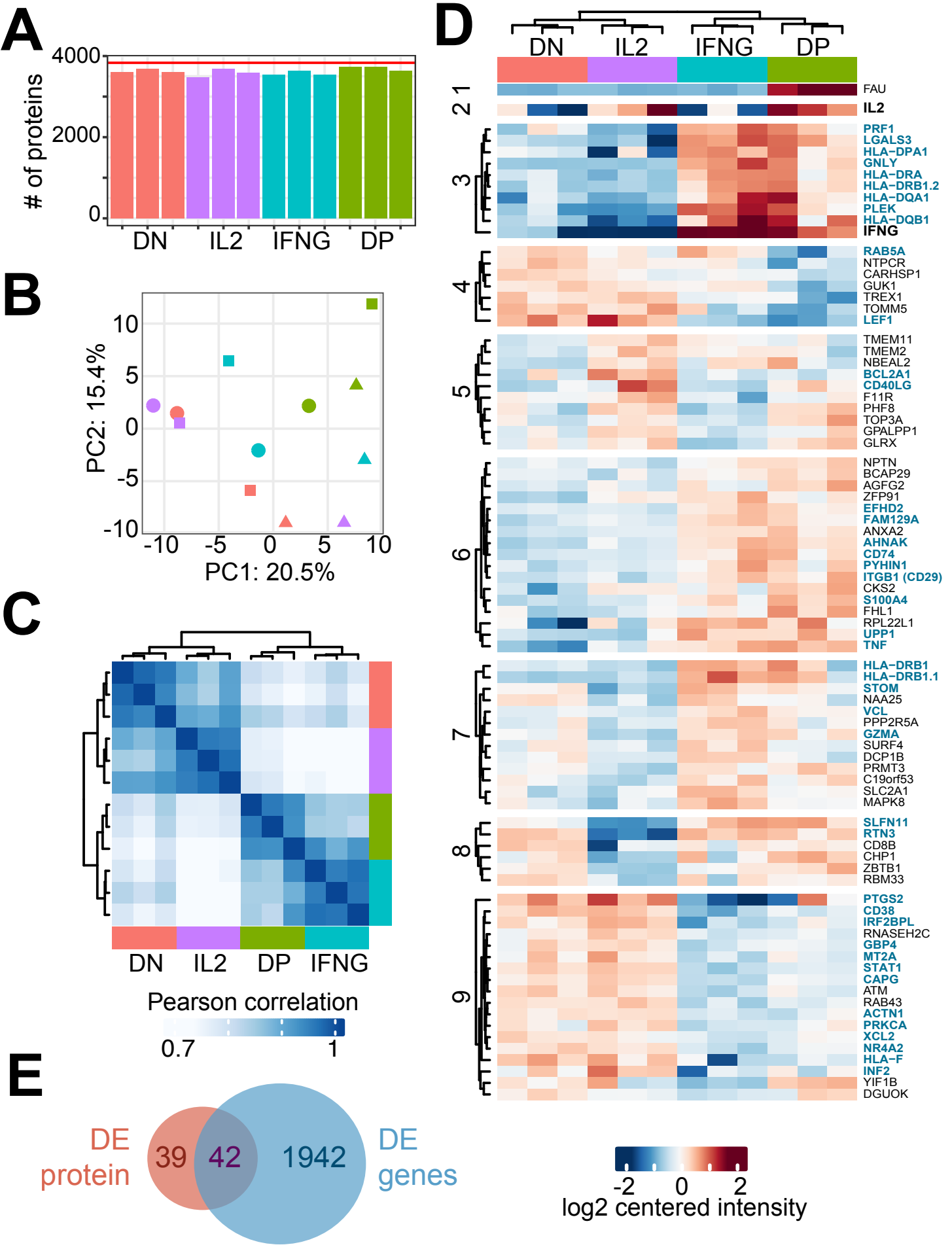
59. Perez-Perri, J. I. *et al.* Discovery of RNA-binding proteins and characterization of their dynamic responses by enhanced RNA interactome capture. *Nat. Commun.* **9,** 4408 (2018).

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

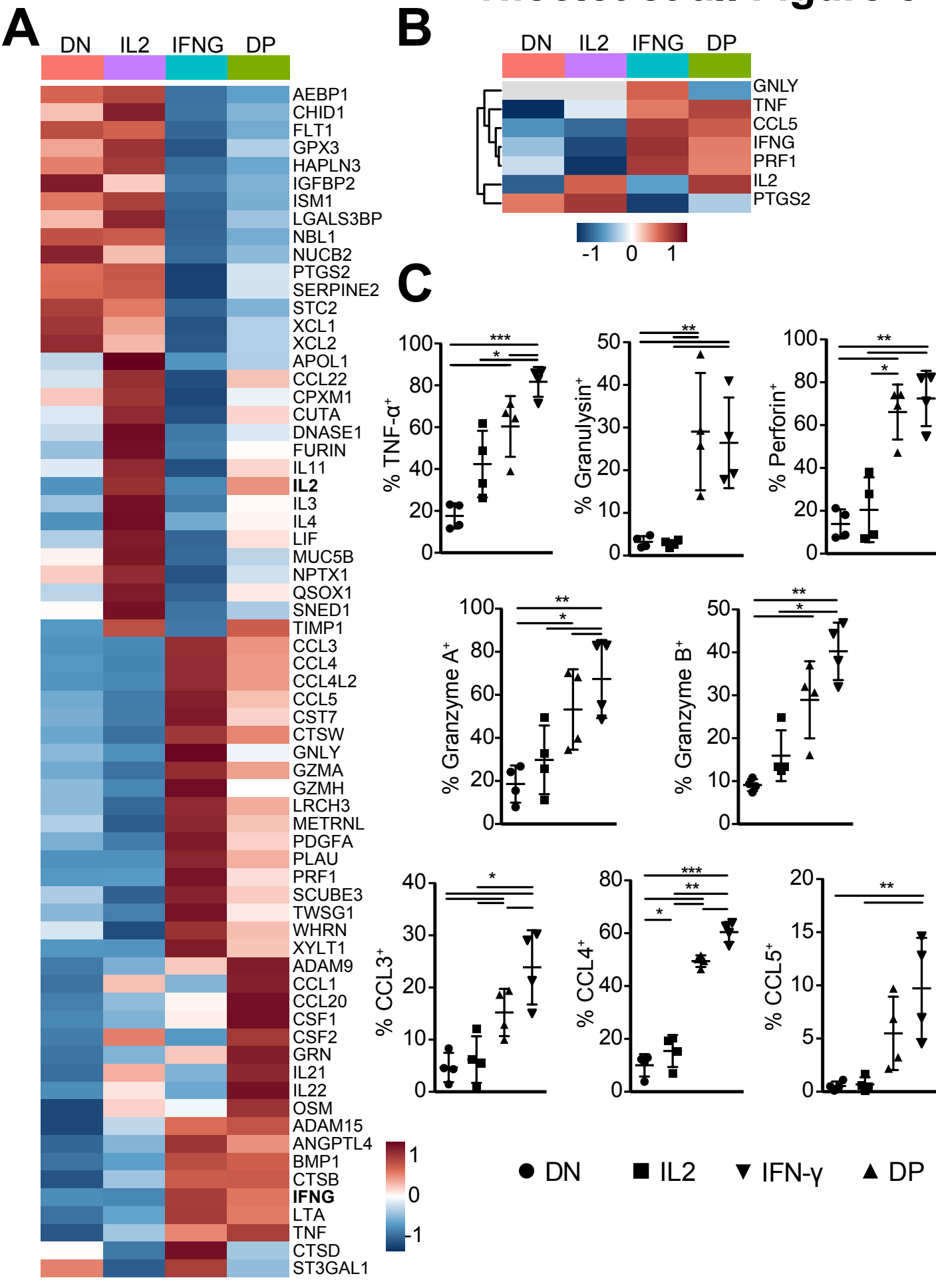
Nicolet et al. Figure 1



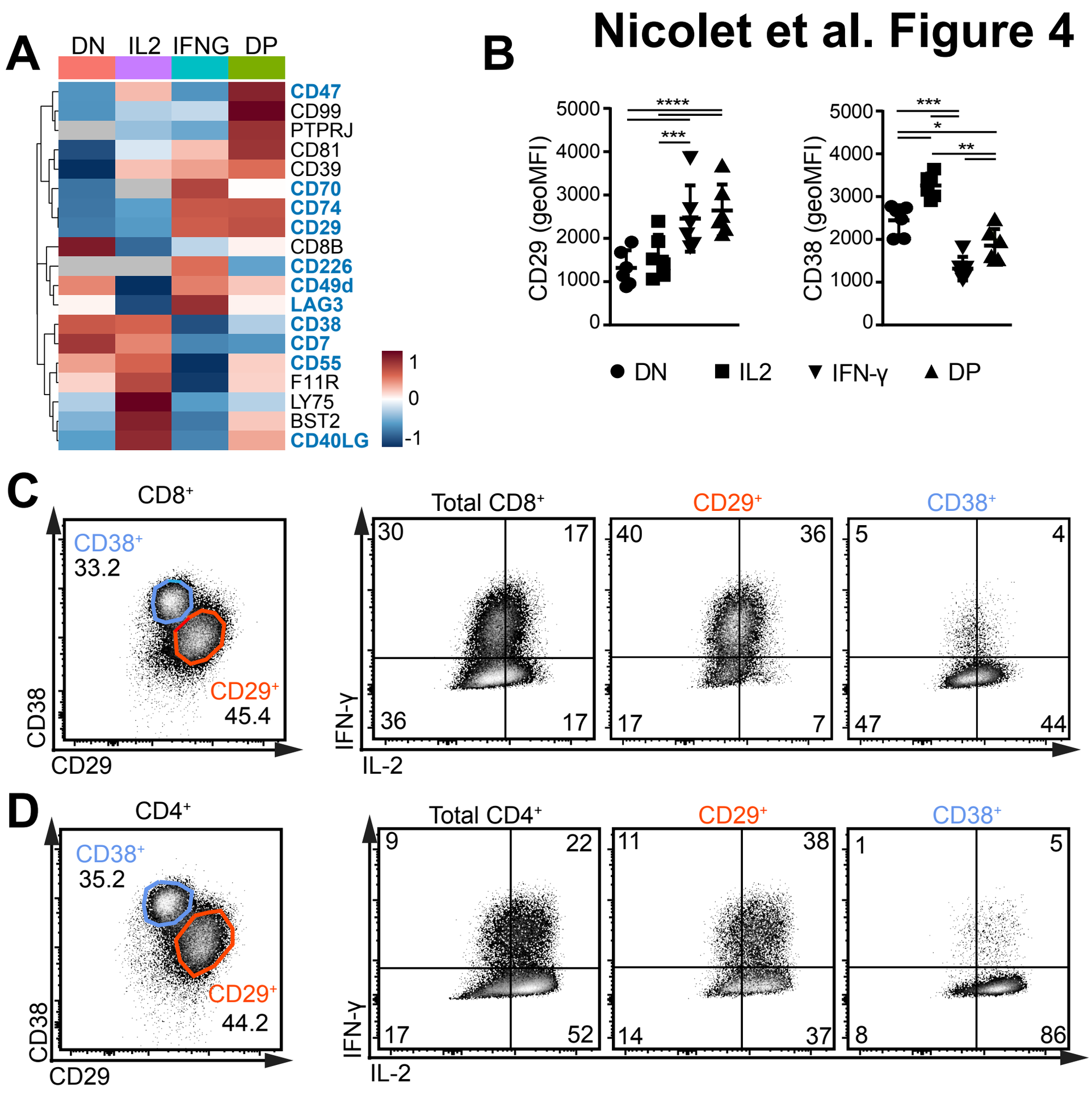
# Nicolet et al. Figure 2



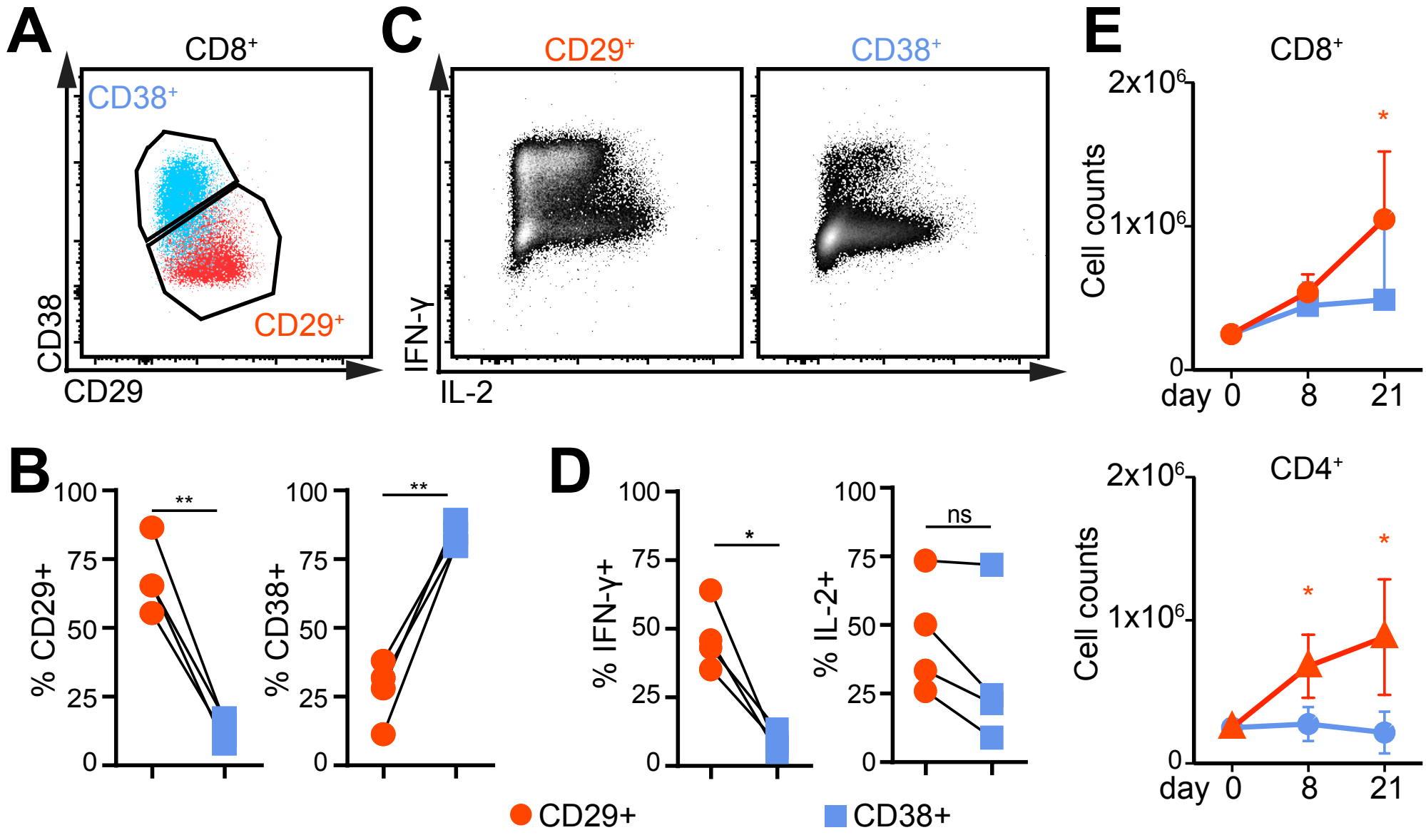
# Nicolet et al. Figure 3



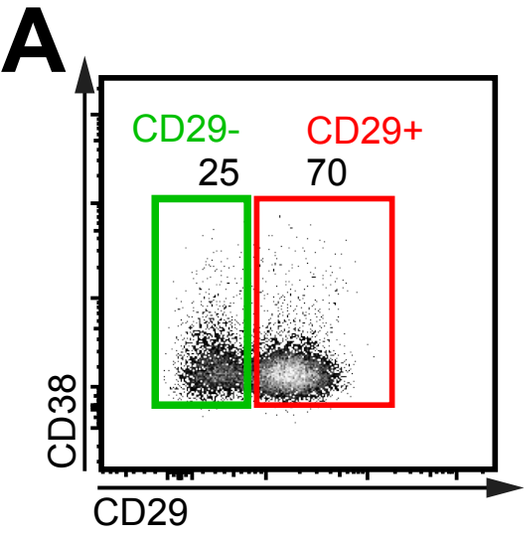
# Nicolet et al. Figure 4

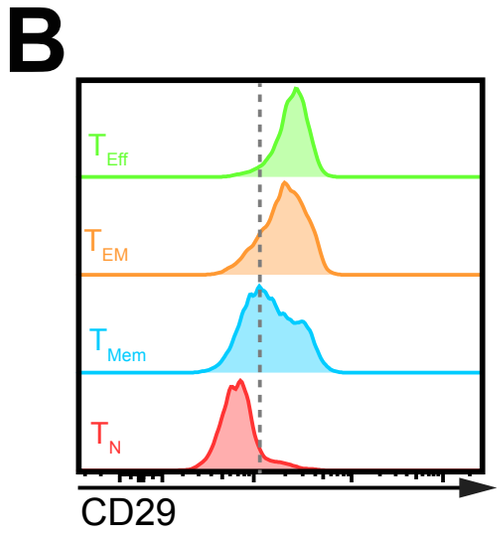


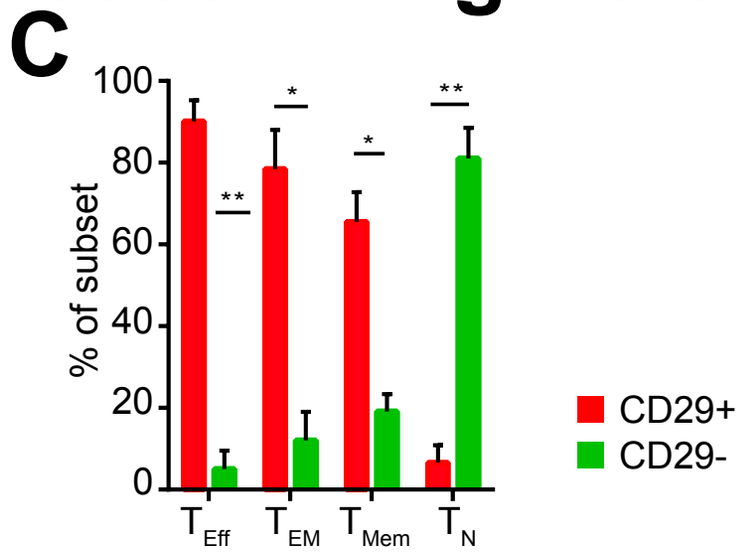
# Nicolet et al. Figure 5

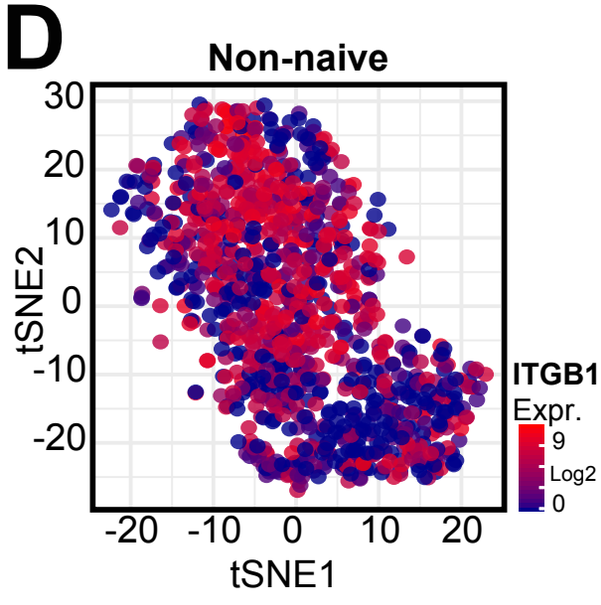


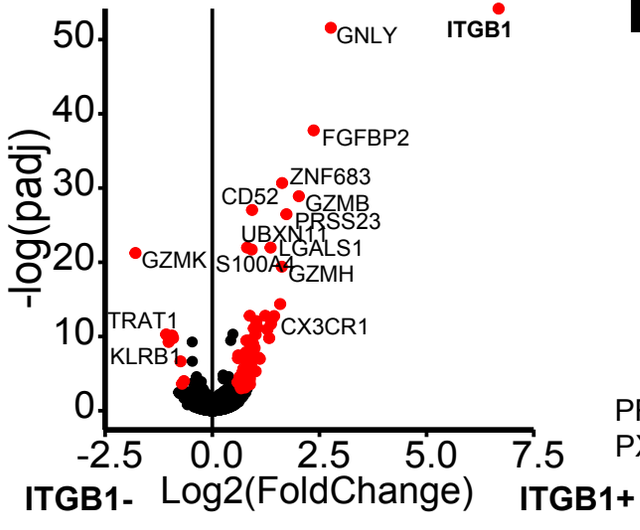
# Nicolet et al. Figure 6

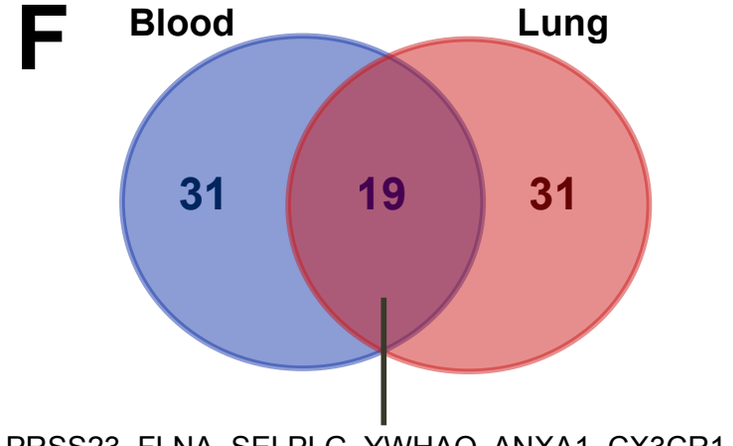


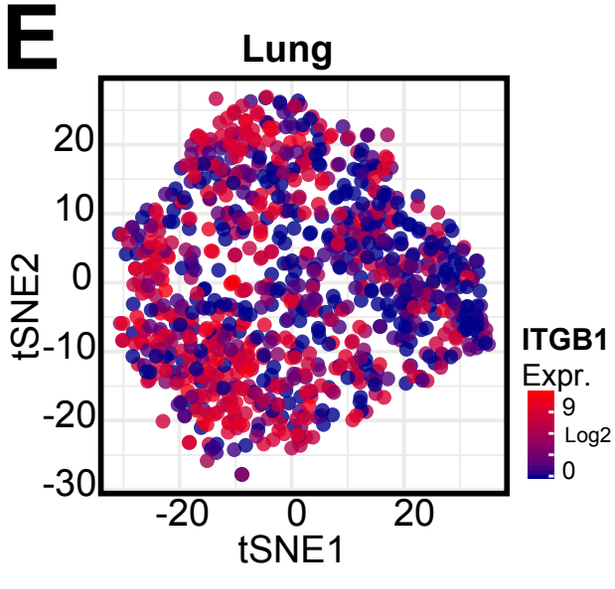



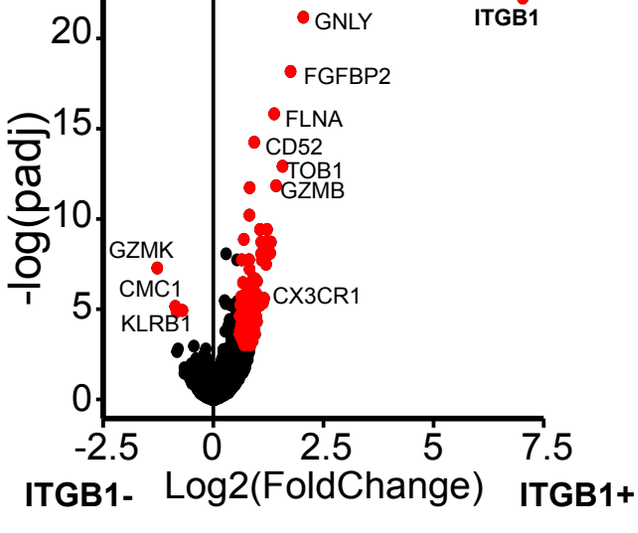



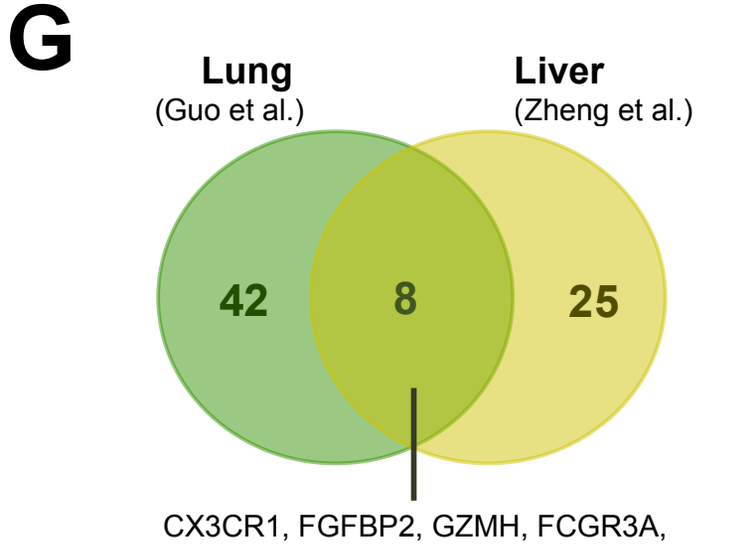



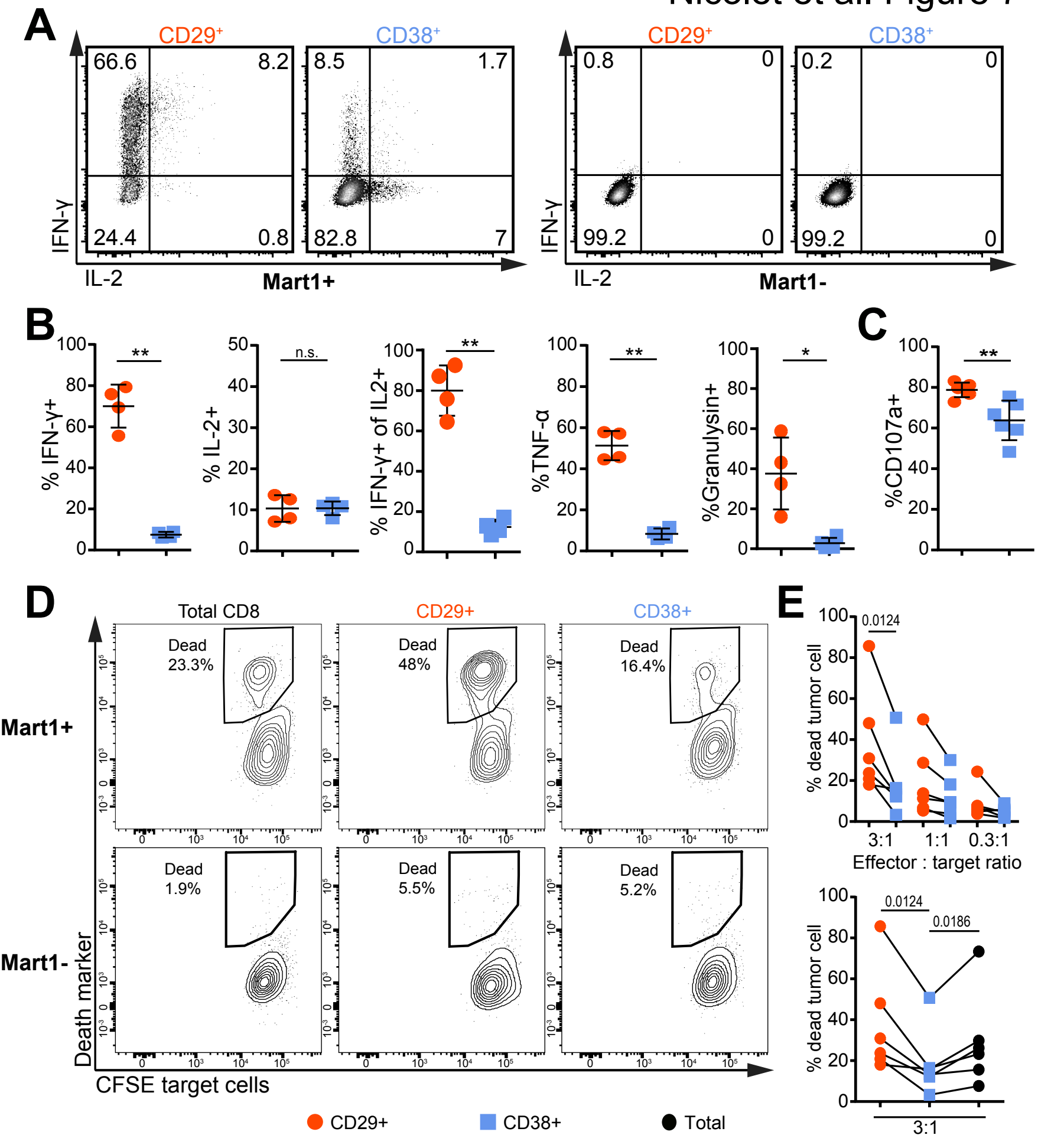




PRSS23, FLNA, SELPLG, YWHAQ, ANXA1, CX3CR1, PXN, FGFBP2, GZMH, TOB1, CD52, LITAF, FCGR3A, **ITGB1**, LGALS1, HOPX, GZMB, GNLY, PATL2




CX3CR1, FGFBP2, GZMH, FCGR3A, **ITGB1**, LGALS1, GZMB, GNLY

Nicolet et al. Figure 7