Supplemental Methods

I. Statistical Methodology, Part A

Statistical analyses of non-microarray data:

For two-group comparisons with continuous valuables, the 2-tailed Student's *t* test with unequal variance, either paired or impaired depending upon the relationship of the samples in the groups compared, was used to generate *P* values.

For comparisons between groups with non-continuous variables, the Pearson's chi-squared (χ2) test was used for *P* value generation.

For a Pearson correlation between two groups, correlation coefficients (r) and sample sizes (*n*) were used to calculate the *t* values using the formula $t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$ before being converted to *P* values applying 2-tailed Student's *t* distribution function.

For qRT-PCR experimental data, *P* values were calculated using the ΔCt values. $P < 0.05$ is considered as a significant difference.

II. Statistical Methodology, Part B

Abbreviations:
Fc = fold change; $|Fc|$ = absolute fold change

A. Discovery set samples: $CD4^+$ T cells were isolated from the tumors (TIL), lymph nodes (LN) and peripheral blood (PB) of 10 breast cancer patients and PB from 4 healthy donors (clinico-pathological characteristics detailed in Supplemental Table 1B).

Clustering analyses

The dendrogram of unsupervised hierarchical clustering analysis was generated in R with pvclust (Suzuki & Shimodaira, 2006) using the top 5% (*n* = 2,734) most variable probe sets across all samples. Correlation distance and average linkage were used for the clustering shown in Figure 1; other methods gave similar results (data not shown). Robustness of each branch separation in the dendrogram was estimated by bootstrap analysis.

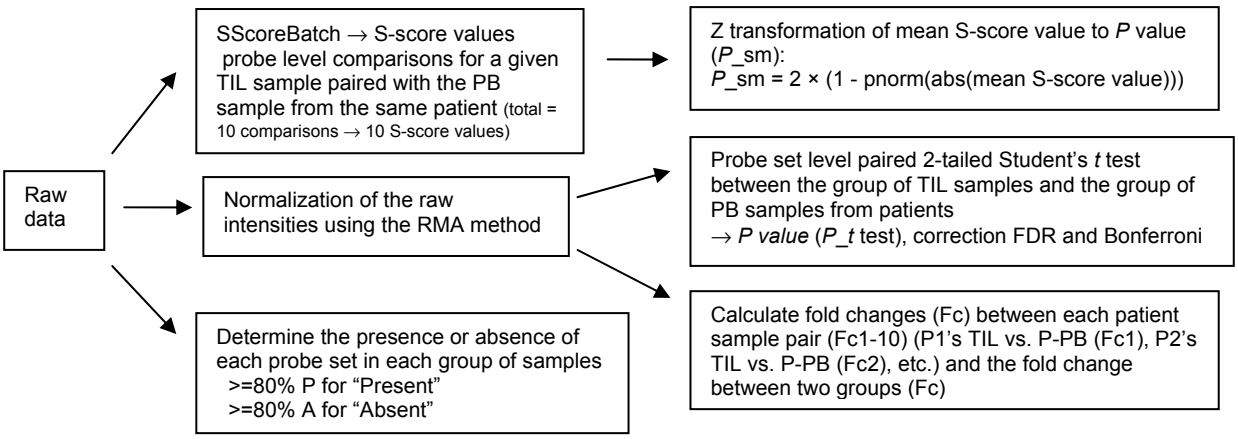
Gene selection

Analysis of the discovery set microarray data was a derivative of methodology used in a previous study of patients with hypereosinophilic syndrome (Ravoet *et al*., 2009). The initial criteria from our previous study were deliberately set to be very stringent because of the small patient number (*n* = 3) in that study. Because the number of patients in this study's discovery set was higher (*n* = 10), we slightly loosened these criteria. Therefore, we combined *t* test *P* values together with those obtained using the S-score algorithm and other criteria (detailed in Figure A below) to generate a list of "significant" genes for a given comparison (i.e. TIL vs. P-PB, TIL vs. LN, LN vs. P-PB, P-PB vs. D-PB, ER^- vs. ER^+ TIL or Ext vs. Min TIL in Supplemental Table 2).

In (Ravoet *et al*., 2009) we found that small sample sizes of microarray data required different analytical approaches than those traditionally used for large datasets and used uncorrected S-score *P* values with additional stringent filters. We found this was better than using FDR or Bonferroni corrected *t* test *P* values to select consistent gene changes, including those with low significance but potential biological importance. The S-score algorithm (Zhang *et al*., 2002) permits direct testing of the hypotheses using probe level data and increases the accuracy of differentially expressed gene identification based on a straightforward error model, offering higher statistical power for small sample sizes (Kennedy *et al*., 2006). This method showed excellent sensitivity and specificity in detecting low-level gene expression changes with the rank ordering of S-score values more accurately reflecting known fold change values compared to other algorithms (RMA, dChip and MAS5). The traditionally used *t* test (using probe set level normalized data) is an accurate tool for large sample size data analyses but accuracy decreases with sample size (Subramaniam and Hsiao, 2012). Thus, for small sample size microarray data the *t* test is a low-powered statistical test and inference could be based on an error model.

In this study, group comparisons were made for data from 10 patients (e.g. TIL vs. P-PB) as well as for data from ±5 patients (e.g. Ext vs. Min TIL). For these comparisons, we needed to employ the same method and have this method be equivalent for each. The S-score algorithm only permits comparisons between two samples and generates an S-score value for each probe set and its associated *P* value [*P* value = 2 * (1 - pnorm(abs(S-score value)))]. To address this problem, we converted the mean S-score value to generate a combined S-score *P* value for each two-group comparison (Figure A). This combination *P* value method can be more stringent for large sample size datasets than for small ones (Whitlock, 2005), which is the opposite of the *t* test. On the other hand, limitations using S-score exist such that a large observed S-score could indicate a defective chip (or other unexplained factors) rather than a biologically significant change. Other statistical tests using normalized probe set level expression data, such as the *t* test, could potentially help to reduce this type of error risk. For our small size dataset, we chose to combine two different statistical methods (S-score and *t* test), each considering a different level of data (probe- and probe set-level, respectively) and select genes determined to be significant by both methods, independent of their relative degree of significance (to include low-significance gene changes with potential biological importance). Additional filtering was applied to further eliminate inconsistent gene changes of low significance (Figure A).

Figure A: Statistical methodology used for discovery set samples
(example shown for the TIL vs. P-PB comparison)



S-score algorithm is used to generate one S-score value per probe set for each patient sample pair at the probe level, the mean S-score values for all 10 patients were transformed to *P* value (*P*_sm), 2,955 probe sets have *P*_sm inferior to 0.05 (299 = |Fc| <2 and 0 = |Fc| <1.5).
Student's *t* test (2-tailed with unequal variance) is performed in parallel using RMA normalized probe set level Log2 intensity values between the two groups of samples (i.e. TIL and P-PB) from all 10 patients. This generates a second *P* value (*P_t* test) per probe set, with 20,808 probe sets having a *P_t* test inferior to 0.05 where 14,581 remain after the FDR correction (10,507 = |Fc|<2, and 5761= |Fc|<1.5) and 283 after the Bonferroni correction (54 = |Fc|<2, 9 = |Fc|<1.5).

1st: We selected six primary lists of probe sets based on different criteria:
1. List with both *P*_sm and *P_t* test inferior to 0.05 and |Fc| superior to 2;
2. List with both *P*_sm and *P_t* test inferior to 0.05 and all of the |Fc1-10| (patients 1-10) superior to 1.2;
3. List with *P_t* test_Bonferroni inferior to 0.05;
4. List with *P_t* test_FDR inferior to 0.002 and |Fc| superior to 2;
5. List with *P_t* test_FDR inferior to 0.05, |Fc| superior to 2 and all of the |Fc1-10| (patients 1-10) superior to 1.2;
6. List with *P_t* test_FDR inferior to 0.05 and all of the |Fc1-10| (patients 1-10) superior to 1.4.

2nd: In all of these lists, we eliminated probe sets with either a Mean Log2 Intensity value inferior to 3.5 or detected as "Absent" in the up-regulated group in all patients. For probe sets detected as "Absent" in >5 patients of the up-regulated group, we only selected those with both types of *P* values inferior to 0.01.

Thus, for the TIL vs. P-PB we generated a list of 3,412 probe sets with 2,632 from the first list, 108 from the second list, 63 from the third list, 376 from the fourth list, 190 from the fifth list and 43 from the sixth list.

Briefly, as shown in Figure A, the raw data (CEL files) were analyzed with the *SScoreBatch* function of the *SScore* package in R (a language and environment for statistical computing and graphics, available from http://www.r-project.org/) version 2.3.0 (Kennedy *et al*., 2006; Zhang *et al*., 2002). S-score values for each two-sample comparison (either paired or impaired depending upon their relationship in the groups compared) were generated then the mean S-score values of all possible two-chip comparisons were used to calculate the combined S-score *P* values (*P*_sm) [*P*_sm = 2 * (1 - pnorm(abs(mean S-score value)))]. RMA-normalized Log2(intensity) values were used for the calculation of *t* test *P* values and the fold change values. For gene selection, we considered several lists of probe sets with different levels of significance based on the S-score *P* values, the *t* test *P* values and the fold change values. The detection values ("present", "absent" or "marginal") were further considered to eliminate genes detected "absent" in the group with increased expression levels. The selected genes are ranked based on the S-score *P* values. Additional genes (indicated with a "+") were selected on the basis of all criteria except the S-score *P* values.

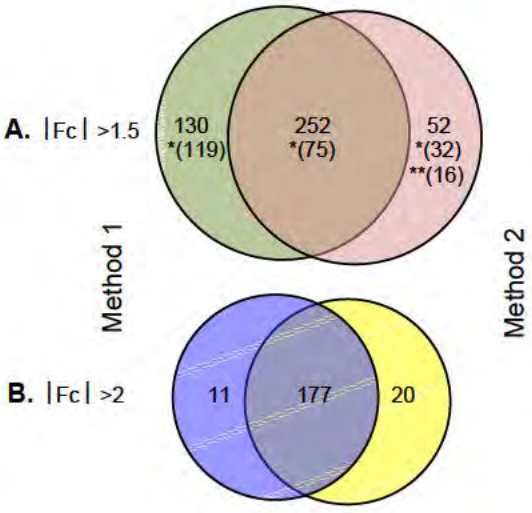
We recognize that this approach is less stringent than traditionally used methodology. Our goal for these analyses was to obtain an initial picture of the data and then use other

experimental approaches to subsequently confirm some of the important changes detected. Therefore, we considered a larger list of genes, including some consistent changes of low-significance, in our small sample size datasets (rather than a highly restricted smaller list that removed some consistent low-significant gene changes of real biological interest). It is interesting to note that even this less stringent approach overlooked some true positive gene expression changes such as *IFNG* and *CXCL9* (and *PDCD1* [PD-1], a gene poorly detected by microarrays) in the Ext vs. Min TIL comparison. Due to their known biological importance, several of these genes, including those mentioned above, were assessed by sensitive qRT-PCR and found to be significantly increased in extensively infiltrated tumors.

B. Discovery set, part 1: non-stimulated (NS) and stimulated (S) memory ($CD45RO^+$) $CD4^+$ T cells from a healthy donor blood were treated with SN from 4 tumors (two with minimal, one extensive and one borderline extensive lymphoid infiltrate). NS total $CD4^+$ T cells from another donor blood were also treated with 3 SN (diluted on the basis of total RNA obtained) giving low level changes (due to the SN dilution) that were consistent with the above-mentioned data.

In (Ravoet *et al.*, 2009), a critical filter (containing several sub-filters) was applied to the three patient groups, based on the well-accepted notion in biology that an expression value superior to 3-times the background standard deviation can be considered to be differentially expressed. This filter permitted selection of consistent gene changes by eliminating all inconsistent ones (stringent but particularly useful for small sample size datasets containing only triplicates or duplicates). We tested the power of this filter on our previously published data and found that it alone (method 2 in Figure B below), without *P* value consideration, permitted the selection of remarkably similar (83% similarity) gene changes to the relatively "complicated" approach (method 1) used in Ravoet *et al*. This was particularly true for genes with an absolute fold change > 2 (Figure B).

Figure B: Comparison of the method used in Ravoet, *et al.* (Method 1) with the filter alone (Method 2)



E.g. comparison P1-yr+6[a] vs P1-yr0[b]
[one patient; triplicate samples]

A.
The number of deregulated probe sets selected using Method 1 (only those with |Fc| >1.5 are included in the comparison between the two methods) or Method 2 → 83% of probe sets selected with Method 2 are also in the list obtained with Method 1.
* Number of probe sets with |Fc| inferior to 2.
** Number of probe sets with S-score *P* values superior to 0.05.

B.
If we consider only the probe sets with |Fc| superior to 2 then the convergence of the two methods is excellent.

[a]P1-yr+6 = P1 had progressed to T lymphoma
[b]P1-yr0 = P1 with chronic premalignant disease

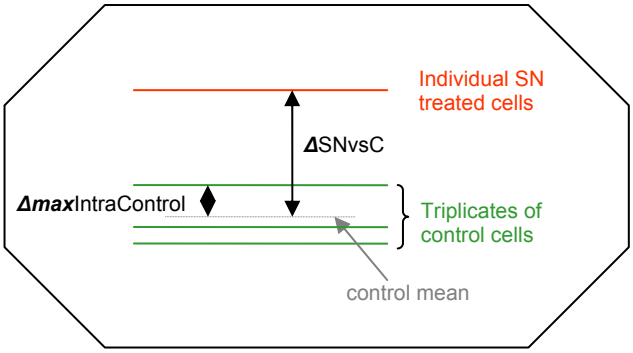
For this dataset (tumor SN treatment of normal $CD4^+$ T cells), untreated control cells (either non-stimulated [NS] or stimulated [S]) were in triplicate, and each sample treated with a different tumor SN (NS+SN or S+SN) was compared to the three control samples using the same concept shown in Method 2 (Figure B, above), which is detailed in Figure C (below). We elected to separately compare each tumor SN to the three control samples because there were potentially differences in individual SN's of important biological significance (e.g. SN019, from an extensively infiltrated tumor, was less immunosuppressive than SN's from minimally infiltrated tumors).

Figure C: Statistical methodology used for each tumor SN treated sample

Filters were applied using fold change values obtained by the RMA normalization method:

a. for |Fc| between 1.5 and 2, ***Δ***SN vs. C 5-fold of ***Δmax***IntraControl. Probe sets detected "Absent" or "Marginal" in the up-regulated sample(s) were eliminated.

b. for |Fc| superior to 2, ***Δ***SNvsC 3-fold of ***Δmax***IntraControl. Probe sets detected "Absent" or one incidence of "Marginal" and |Fc| inferior to 3 were eliminated.



Using these filters, we generated a relatively accurate list of deregulated genes for each tumor SN-treated sample (the high level of similarity between our microarray data and qRT-PCR results for a list of randomly selected genes confirmed this accuracy [Supplemental Table 5G, Tumor SN Expt#2]). Donor cells treated with SN from 2 extensive and 2 minimally infiltrated tumors were considered separately; genes that were commonly changed by both SN's (minimal or extensive) were selected for further analysis. For NS+SN treatment, we performed preliminary experiments where total $CD4^+$ T cells were treated with 3 diluted SN. Genes from this experiment with consistent changes in at least two SN's were included in our subsequent analyses despite their lower overall fold change levels due to SN dilution. Our rationale was because we found that in SN-treated memory ($CD45RO^+$) $CD4^+$ T cells some gene changes were not detectable by microarrays but were using qRT-PCR (e.g. *GNLY*; Supplemental Table 5, Tumor SN Expt#1). For the comparison between untreated S and NS cells, the same filters were applied to individual S samples compared to the triplicate NS samples and the gene changes commonly detected in all three S samples were considered to be significant.

C. <u>Discovery set, part 2</u>: $CD4^+$ TIL (shown to be >95% $CD45RO^+$) from two tumors (TIL062 and TIL064) and $CD4^+CD45RO^+$ T cells from a healthy donor blood (control) were split, with half of the cells extracted immediately (time 0) and the remaining half extracted

after a 24h rest ex vivo prior to gene expression analysis using microarrays (Clinico-pathological characteristics detailed in Supplemental Table 1C).

First, each 24h rested TIL was compared with its corresponding time 0h fresh TIL. Genes with an absolute fold change >2 in these TIL comparisons (2 tumors) were initially selected with subsequent elimination of any detected as "Absent" or "Marginal" in the upregulated group. Genes that changed in only one TIL were considered only if the absolute fold change was >5. Because the two tumors used in this experiment were differentially infiltrated with lymphocytes (TIL064=extensive and TIL062=minimal), they were analyzed separately. This fold change limit was stringently applied to any gene changes that were specific for only one tumor but less stringent for changes common to both tumors.

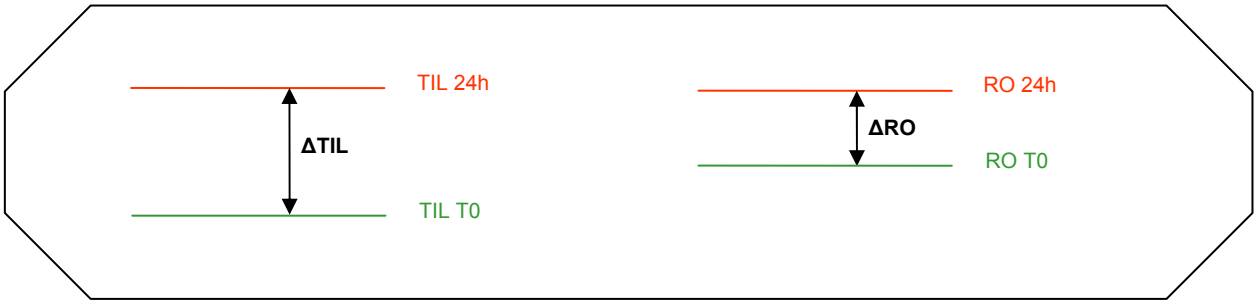
Second, we also compared gene changes in the rested TIL with rested memory $CD4^+$ $CD45RO^+$ T cells from a healthy donor (RO in Table A) and changes detected in the rested TIL but <u>not</u> the rested RO cells were selected. Genes were selected if they had an absolute fold change in the 24h rested RO (compared to RO time 0) of <1.2; some genes with an absolute fold change >1.2 in the rested RO but with high ratio values of ΔTIL vs. ΔRO (variable on the basis of the fold change values of the rested TIL and the rested RO) were also selected (listed in Table A with the definition of ΔTIL and ΔRO shown in Figure D).

These lists were selected using simple methods to quickly identify genes that commonly changed under different circumstances for additional analysis, focusing on genes whose expression was also altered in the TIL vs. P-PB comparison (discovery set).

Table A. Criteria used for data analysis of 24h rested TIL

24h vs time 0									
Common gene changes (up)	TIL062&TIL064	Fc>2		RO	Fc<1.2	or	1.2<Fc<1.5	ΔTIL/ΔRO	>5
		Fc>5					Fc>1.5	ΔTIL/ΔRO	>3
Common gene changes (down)		Fc<-2			Fc>-1.2	or	-1.5<Fc<1.2	ΔTIL/ΔRO	<-5
		Fc<-5					Fc<-1.5	ΔTIL/ΔRO	<-3
TIL Min (up)	TIL062	Fc>5	and	RO	Fc<1.2	or	1.2<Fc<1.5	ΔTIL/ΔRO	>5
		Fc>10					Fc>1.5	ΔTIL/ΔRO	>5
		Fc>5					Fc>1.5	ΔTIL/ΔRO	>10
TIL Min (down)		Fc<-5			Fc>-1.2	or	-1.5<Fc<1.2	ΔTIL/ΔRO	<-5
		Fc<-10					Fc<-1.5	ΔTIL/ΔRO	<-5
		Fc<-5					Fc<-1.5	ΔTIL/ΔRO	<-10
TIL Ext (up)	TIL064	Fc>5		RO	Fc<1.2	or	1.2<Fc<1.5	ΔTIL/ΔRO	>5
		Fc>10					Fc>1.5	ΔTIL/ΔRO	>5
		Fc>5					Fc>1.5	ΔTIL/ΔRO	>10
TIL Ext (down)		Fc<-5			Fc>-1.2	or	-1.5<Fc<1.2	ΔTIL/ΔRO	<-5
		Fc<-10					Fc<-1.5	ΔTIL/ΔRO	<-5
		Fc<-5					Fc<-1.5	ΔTIL/ΔRO	<-10

Figure D: Definition of ΔTIL and ΔRO



D. Public data (Th subsets): Th1, Th2, Tfh, Tcm and Tem subsets from public dataset #1 (Chtanova *et al.*, 2005); Treg, Th17 enriched population, $CD25^{+}$ and Memory(vs. Naive) subsets from public dataset #3 (Miyara *et al.*, 2009) (characteristic details of each subset in Supplemental Table 3A).

Public dataset #1 (Chtanova *et al.*, 2005) and set #3 (Miyara *et al.*, 2009) were analyzed separately using different criteria and filters that were specifically adapted for the biology and intensity range of each dataset. These methods are empiric because only simplicates or duplicates were available and the cell populations were not equivalent (i.e. the arrays were performed using cell clones or purified cells with different levels of specificity depending on the combination of surface markers employed), making it impractical to apply the same method to each dataset. Based on our selection criteria (the criteria or filter details used in this study are not exclusive and while it is possible to define others using similar concepts and arguments to generate slightly different gene lists, the global view of these expression profiles would remain similar), genes that were relatively specific for a given Th subset (with

some overlap between subsets due to Th plasticity and the relative gene-cell specificity for a given Th subset; e.g. Th1 and Th2) were selected.

Briefly, for public dataset #1 (Chtanova *et al.*, 2005) (this method applies to the Th1, Th2, Tfh, Tcm and Tem subsets), we selected genes with an absolute fold change >2 in both comparisons considered as shown in Table B for the Th1 subset: 1) Th1 subset compared to cord blood naive ($CD45RA^+$) $CD4^+$ T cells ("Th1 vs. naive"; Table B); and 2) Th1 subset compared to the mean intensity of the other Th subsets ("Th1/Ave"; Table B). The Th1 and Th2 cells were polarized clones derived from cord blood naive cells whereas Tfh, Tcm and Tem cells were purified from adult donors (tonsil or blood). Because adult blood naive cell data was absent in this study, we also included cord blood (footnote "e"; Table B) and adult blood (footnote "c"; Table B) naive cells from public dataset #2 (Lee *et al.*, 2004) as controls. This helped to eliminate non- or very low-specific gene changes by only considering genes with an absolute fold change >4 in a given Th (Th1 in this example) subset compared to adult blood naive cells ("Th1/AB4"; Table B). An additional list of gene changes with an absolute fold change >4 in the "Th1 vs. naive" comparison and >2 in the "Th1/AB4" as well as the "Th1/Ave" comparisons was also considered. Because a small number of probe sets were absent from dataset #2 (derived from Affymetrix gene chip U133A plus U133B analysis; footnote "c" and "e"; Table B), adult blood naive cells (footnote "f"; Table B) from public dataset #3 (Miyara *et al.*, 2009) were also incorporated as an additional control. For this small number of probe sets, the ratio value of "Th1/AB4" (as well as for the other Th subsets) is calculated between the Th1 subset samples from dataset #1 and the adult blood naive control sample from dataset #3 (footnote "g"; Table B). Because a direct ratio calculation between samples from different datasets may introduce errors, for these probe sets we only considered genes with an absolute fold change >4 in the "Th1 vs. naive" comparison and >2 in the "Th1/AB4" and "Th1/Ave" comparisons. The calculation details and selection criteria are shown in Table B below.

Table B. Methodology used for Th subsets from public dataset #1

Calculation :										
Probe Set		Th1[a] vs. naive[b]	Th1/AB4[c]	Th1/Ave[d]	Th2[a] vs. naive	Th2/AB4	Th2/Ave	Tfh[a] vs. naive	Tfh/AB4	Tfh/Ave
1007_s_at	with CB4vs.AB4[d]	0.512	0.524	0.881	0.422	0.432	0.699	0.667	0.682	1.228
1552256_a_at	without CB4vs.AB4[e]	1.033	1.012	1.422	0.921	0.902	1.221	0.608	0.595	0.730
Probe Set		Tcm[a] vs. naive	Tcm/AB4	Tcm/Ave	Tem[a] vs. naive	Tem/AB4	Tem/Ave		CB4[e] vs. AB4[c]	cbRA^{+}[b] vs. pb25^{-} RA^{+}[f]
1007_s_at	with CB4vs.AB4[d]	0.633	0.647	1.147	0.605	0.619	1.084		1.022	1.313
1552256_a_at	without CB4vs.AB4[e]	0.669	0.655	0.818	0.709	0.694	0.877			0.979
Selection criteria :										
Example Th1		Th1 vs. naive	Th1/AB4[c]	Th1/Ave[d]	Rank		The selection creteria for Th2, Tfh, Tcm and Tem are similar to Th1.			
with CB4vs.AB4	either	>2 (up) or <0.5 (down)	>4 (up) or <0.25 (down)	>2 (up) or <0.5 (down)	Th1/Ave					
	or	>4 (up) or <0.25 (down)	>2 (up) or <0.5 (down)							
without[g] CB4vs.AB4	only	>4 (up) or <0.25 (down)	>2 (up) or <0.5 (down)	>2 (up) or <0.5 (down)						

[a]Th1 and Th2 are in vitro polarized clones from cord blood naive cells; Tfh, Tcm and Tem are purified cells from tonsil or adult peripheral blood; from public data set #1 from paper PubMed ID: 16339519 ; Affymetrix gene chip U133plus2.0 performed in duplicate

[b]Naive = cbRA^{+} = cord blood CD45RA^{+} naive CD4^{+} T cells from public data set #1 from paper PubMed ID: 16339519 ; Affymetrix gene chip U133plus2.0 in simplicate

[c]AB4 = adult peripheral blood CD45RA^{+}CD62L^{+} naive CD4^{+} T cells from public data set #2 from paper PubMed ID: 15210650 ; Affymetrix gene chip U133A + U133B (a bit smaller number of probe set than U133plus2.0) performed in simplicate

[d]Ave = Mean value of other 4 Th subsets, here for Th1, Ave = (Th2+Tfh+Tcm+Tem)/4

[e]CB4 = cord blood CD45RA^{+}CD62L^{+} naive CD4^{+} T cells from public data set #2 from paper PubMed ID: 15210650 ; Affymetrix gene chip U133A + U133B (a bit smaller number of probe set than U133plus2.0) performed in simplicate

[f]pb25^{-}RA^{+} = adult peripheral blood CD45RA^{+}CD25^{-} naive CD4^{+} T cells from public data set #3 from paper PubMed ID: 19464196 ; Affymetrix gene chip U133plus2.0 performed in simplicate

[g]For probe sets that were absent in Affymetrix gene chip U133A + U133B data (public data set #2), the ratio of Th1/AB4 (idem. for other Th subsets) was calculated relative to adult peripheral blood CD45RA^{+}CD25^{-} naive CD4^{+} T cells (= pb25^{-} RA^{+})[f] from public data set #3

For public dataset #3 (Miyara *et al.*, 2009)(Treg, Th17 enriched population, $CD25^+$ and Memory[vs. Naïve] subsets; purified from adult peripheral blood according to their relative intensity of surface CD45RA and CD25 expression; Table C), different criteria adapted to each subset were used for the selection of subset specific genes. The calculation details and selection criteria are listed in Table C. Specifically we considered the following for the individual subsets:

Treg (=$CD45RA^-CD25^{+++}$ activated Treg population; Table C): Miyara *et al*. identified a new combination of Treg surface markers (CD45RA and CD25) based on the intensity of their expression (CD25 surface expression is proportional to FoxP3 intracellular protein levels in $CD45RA^-$[=$CD45RO^+$] cells; e.g. $CD45RA^-[RO^+]CD25^{+++}$ corresponds to $CD45RA^-[RO^+]FoxP3^{high}$). By considering genes with an absolute fold change (versus naive cells) >2 in the $CD45RA^-CD25^{+++}$ activated Treg subset and eliminating those with similar expression levels in the other $CD25^+$ populations (genes with similar changes in the $CD45RA^-CD25^{++}$ resting Treg were not eliminated because this population is also immunosuppressive; Table C), genes specific for the Treg subset were selected.

Th17 (=$CD45RA^-CD25^{++}$ Th17 enriched population; Table C): Miyara *et al.* also found that the $CD45RA^-CD25^{++}$ cells (corresponding to $CD45RA^-FoxP3^{low}$) were enriched with IL-17 producing Th17 cells (note: IL-17 is also produced at lower levels by other Th subsets including purified Treg)(Supplemental Table 3A; Table C below). We therefore used these data to identify genes associated with a Th17-enriched population, whose expression levels were proportional to the percentage of IL-17 producing cells but not to CD25 (or FoxP3) expression intensity in the purified Th subsets. Briefly, we selected genes with an absolute fold change (versus naive cells) >2 in the $CD45RA^-CD25^{++}$ population and then considered those whose expression levels increased (or decreased for downregulated genes) in parallel with the percentage of IL-17^+ cells in the other populations (detailed in Table C). Because the $CD45RA^+CD25^{++}$ population also contains a very small number of IL-17^+ cells (compared to $CD25^-$ naive cells), genes differently changed in this population (e.g. genes upregulated in Th17 enriched $CD45RA^-CD25^{++}$ but downregulated in this $CD45RA^+CD25^{++}$ population with ratio <0.5; Table C) were eliminated. Due to the lack of a more specific gene expression dataset for the Th17 subset, this analysis provides a valid impression of Th17 gene expression, which is supported by the number of known Th17 specific genes (e.g. *CCL20*, *IL1R1*, *RORC*, etc.; Supplemental Table 3) in the data derived applying this analytical approach. A number of these genes were confirmed by qRT-PCR (Figure 3) or by flow cytometry (Supplemental Figure2).

$CD25^+$ (=four populations containing $CD25^+$ cells; Table C): By selecting genes with an absolute fold change (versus naive cells) >2 in all four $CD25^+$ populations (compared to a $CD25^-$ population), we produced a list of commonly changed genes for $CD25^+$ $CD4^+$ T cells (independent of CD25 expression intensity).

Memory (=three $CD45RA^-$ populations; Table C) vs. Naive: By considering the three $CD45RA^-$ (=$CD45RO^+$) memory populations (compared to two $CD45RA^+$ naive populations) in diminishing importance according to their surface CD25 positivity (the fold change limit set at 2 for the $CD45RA^-CD25^{-/+}$ population [compared to naive cells], 1.5 for the $CD25^{++}$ and 1.5

for the CD25^{+++} CD45RA^{-} populations, respectively) and applying additional criteria (Table C), commonly changed genes in the three memory cell populations were selected.

Table C. Methodology used for Th subsets from public dataset #3

Calculation															
	selected genes with rank				Ratio vs. pb CD4^{+} CD45RA^{+}CD25^{-} cells				Ratio of ratio						
	Th subsets/subpopulations				[1]	[2]	[3]	[4]	Treg		Th17 (enriched)				Memory vs. Naive
Probe set	Treg	Th17	CD25+	Memory vs. Naive	CD45RA^{+} CD25^{++} (FoxP3^{lo}) vs. Naive	CD45RA^{-} CD25^{+++} (FoxP3^{hi}) vs. Naive	CD45RA^{-} CD25$^{-/+}$ (FoxP3^{-}) vs. Naive	CD45RA^{-} CD25^{++} (FoxP3^{lo}) vs. Naive	[2]/[3]	[2]/[4]	[3]/[1]	[4]/[3]	[4]/[1]	[4]/[2]	[3]/[1]
232968_at	up1		up7		2.06	87.99	2.47	29.17	35.5807	3.0161	1.20065	11.79703	14.1641	0.331557	1.2006
224361_s_at		up1			1.04	2.43	1.74	6.70	1.3953	0.3628	1.670386	3.845947	6.424215	2.756451	
226034_at	up8	up115	up1	up192	4.16	215.97	11.28	63.29	19.1413	3.4121	2.711206	5.609761	15.20922	0.29307	2.7112
214470_at				up1	0.83	7.38	30.23	16.35	0.2443	0.4518	36.32376	0.540697	19.64015	2.213511	36.3238
223836_at					0.16	0.16	10.88	0.15	68.1295	0.9338	0.014976	72.95991	1.092625	1.070901	66.7749
1405_i_at	down1				0.35	0.30	19.09	0.72	63.6055	2.3968					54.9054
230983_at		down1			1.37	0.53	0.65	0.42			2.116825	1.541437	3.262952	1.271279	
209840_s_at			down1		0.06	0.02	0.10	0.02	5.1815	1.0376	0.573015	4.993538	2.861372	0.96372	0.5730
231798_at				down1	1.02	0.07	0.08	0.14	1.1236	1.9407	12.34041	0.57897	7.14473	0.515276	12.3404
1553645_at	up20				1.07	2.97	0.28	1.35	10.5045	2.2041					3.7866
	Public data 3 from paper PubMed ID: 19464196; Affymetrix gene chip U133plus2.0 performed in simplicate														
	In color are highlightened the most up/down-regulated gene for each sub-population [Treg, Th17, CD25+, (Memory vs. Naive)] and the concerned calculation values														

Selection criteria :				
Th subset	Considered ratios		Considered ratios of ratio	Rank
Treg	CD45RA^{-} CD25^{+++} (FoxP3^{hi}) vs. Naive [2]: Fc>2 (up) or <0.5 (down)	and	[2]/[3]>2, [2]/[4]>1.5	[2]/[3]
Th17 (enriched)	CD45RA^{-} CD25^{++} (FoxP3^{lo}) vs. Naive [4] CD45RA^{+} CD25^{++} (FoxP3^{lo}) vs. Naive [1] up: Fc[4]>2 and Fc[1]>0.5 down: Fc[4]<0.5 and Fc[1]<2	and	[4]/[1]>2, [4]/[3]>1.5, [3]/[1]>1.5	[4]/[2]
CD25^{+}	[1], [2], [3] and [4]: Fc>2 (up) or <0.5 (down)		no	[4]
Memory vs. Naive	CD45RA^{-} CD25 $^{-/+}$ (FoxP3^{-}) vs. Naive [3] CD45RA^{-} CD25^{++} (FoxP3^{lo}) vs. Naive [4] CD45RA^{-} CD25^{+++} (FoxP3^{hi}) vs. Naive [2] [3]: Fc>2 (up) or <0.5 (down) [4]: Fc>1.5 (up) or < 0.67 (down) [2]: Fc>1 (up) or <1 (down)	and	[3]/[1]>1.5	[3]/[1]
Additional information about the nature of each population	Th subsets/subpopulations		Effect on naive Th cell proliferation	IL-17^{+} cells (%)
	CD4^{+} CD45RA^{+} CD25^{++} (FoxP3^{lo})		low suppressive	3.6
	CD4^{+} CD45RA^{-} CD25^{+++} (FoxP3^{hi})		high suppressive	8.9
	CD4^{+} CD45RA^{-} CD25 $^{-/+}$ (FoxP3^{-})		high enhancive	7.4
	CD4^{+} CD45RA^{-} CD25^{++} (FoxP3^{lo})		low enhancive	16.2

The ratio values from public dataset #2 (Lee *et al.*, 2004) between naive CD4^{+} T cells isolated from adult peripheral blood and those from cord blood were included in the heat maps (Figures 1 and 2) and in Supplemental Table 3 to demonstrate any potential differences.

We decided it was important to re-analyze these public datasets because Chtanova *et al.* did not provide a list of altered genes for the Th subsets in dataset # 1 and Miyara *et al.* provided only minimal analysis of dataset #3. The gene lists derived from our re-analysis are certainly not definitive for Th subsets; however, they do provide considerably extended lists

of relatively specific gene changes reflecting the individual Th subsets. They also include genes commonly changed in several subsets, which mirrors the well-known plasticity of the adaptive immune response. These data were used to make a molecular fingerprint of the individual Th subsets present in the TIL and produce a global image of their presence or absence among the tumor-infiltrating $CD4^+$ T cells. Confirmation experiments using flow cytometry and qRT-PCR were guided by these data, which largely verified the changes in marker gene expression.

Note: Due to the fold change limit which is fixed to 32-fold in the heatmaps (Figures 1 and 2), some differences are not visually distinguishable between the Th subsets that commonly share an increase or decrease in specific gene expression albeit at different levels of expression (these genes are included in Supplemental Table 3).

References

Chtanova, T., Newton, R., Liu, S.M., Weininger, L., Young, T.R., Silva, D.G., Bertoni, F., Rinaldi, A., Chappaz, S., Sallusto, F*., et al.* (2005). Identification of T cell-restricted genes, and signatures for different T cell responses, using a comprehensive collection of microarray datasets. J Immunol *175*, 7837-7847.

Kennedy, R.E., Kerns, R.T., Kong, X., Archer, K.J., and Miles, M.F. (2006). SScore: an R package for detecting differential gene expression without gene expression summaries. Bioinformatics *22*, 1272-1274.

Lee, M.S., Hanspers, K., Barker, C.S., Korn, A.P., and McCune, J.M. (2004). Gene expression profiles during human CD4+ T cell differentiation. International immunology *16*, 1109-1124.

Miyara, M., Yoshioka, Y., Kitoh, A., Shima, T., Wing, K., Niwa, A., Parizot, C., Taflin, C., Heike, T., Valeyre, D*., et al.* (2009). Functional delineation and differentiation dynamics of human CD4+ T cells expressing the FoxP3 transcription factor. Immunity *30*, 899-911.

Ravoet, M., Sibille, C., Gu, C., Libin, M., Haibe-Kains, B., Sotiriou, C., Goldman, M., Roufosse, F., and Willard-Gallo, K. (2009). Molecular profiling of CD3-CD4+ T cells from patients with the lymphocytic variant of hypereosinophilic syndrome reveals targeting of growth control pathways. Blood *114*, 2969-2983.

Subramaniam, S., and Hsiao, G. (2012). Gene-expression measurement: variance-modeling considerations for robust data analysis. Nature immunology *13*, 199-203.

Whitlock, M.C. (2005). Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. J EvolBiol *18*, 1368-1373.

Zhang, L., Wang, L., Ravindranathan, A., and Miles, M.F. (2002). A new algorithm for analysis of oligonucleotide arrays: application to expression profiling in mouse brain regions. J MolBiol *317*, 225-235.

Gu-Trantien, *et al.*
Supplemental
Figures 1-3:

Figure S1: Flow cytometric analysis (FACS) of the cellular subpopulations present in human breast tumors (pages 2-8).

Figure S2: Flow cytometric analysis of surface marker expression on $CD4^+$ TIL from human breast tumors (pages 9-73).

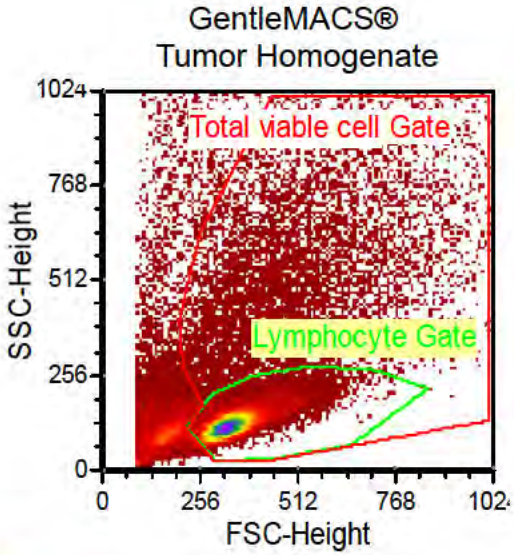
Figure S3: Flow cytometric analysis of cellular subpopulations expressing CXCL13 in human breast tumors (pages 74-77).

Figure S1:

Flow cytometric analysis (FACS) of the cellular subpopulations present in human breast tumors

(Patients from the Confirmation Set; Table S1C)

Figure S1A:
Cell populations in homogenates of fresh breast tumor fragments with different levels of leukocyte infiltration



- Total viable cells
- Lymphocyte gate
- Tumor cells = EpCAM^{hi}CD45^{-} large cells
- Leukocytes = CD45^{+}
- CD10^{hi}CD45^{-} (MSC, other stroma cells)

% = percentage of total viable cells.

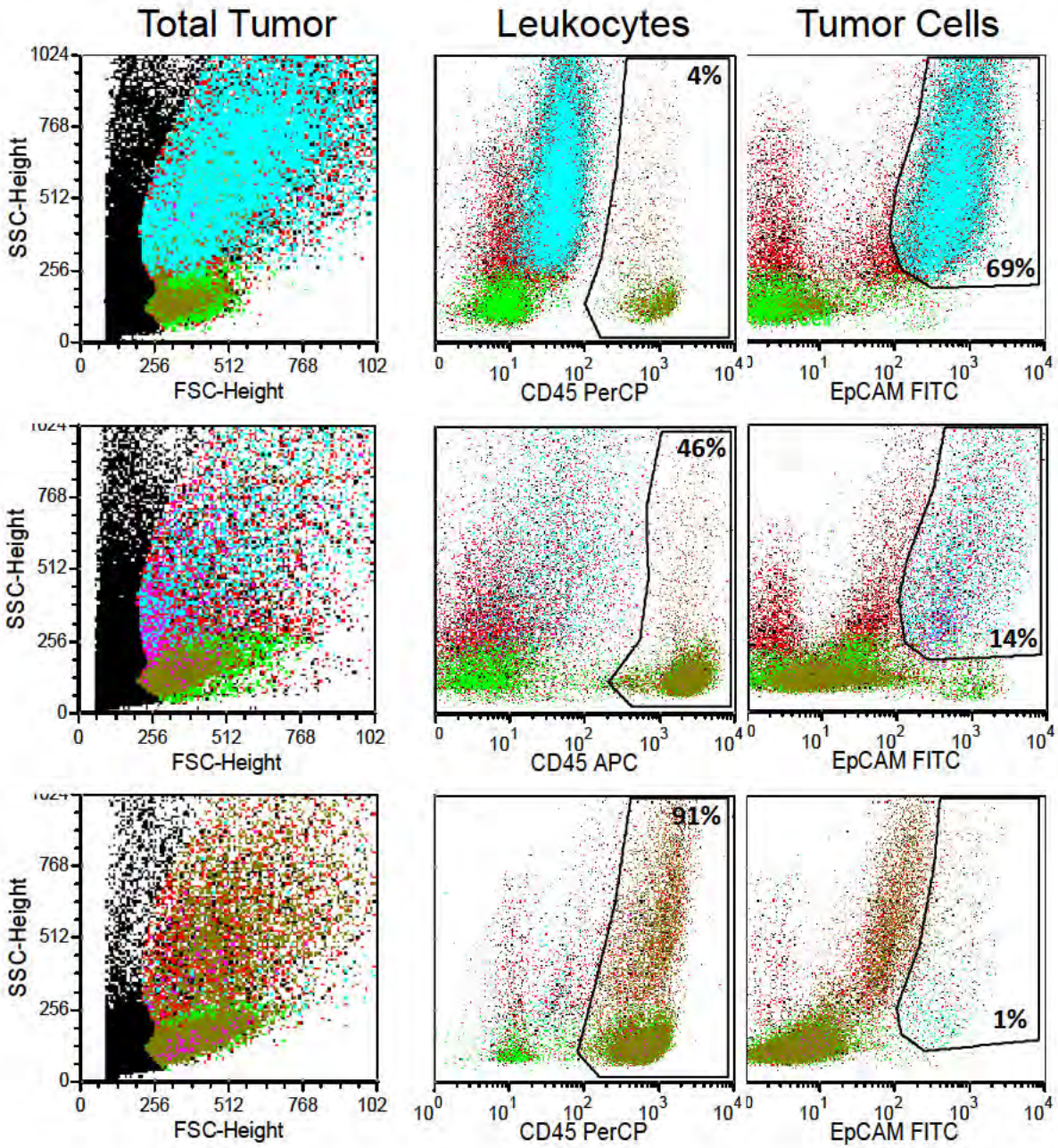
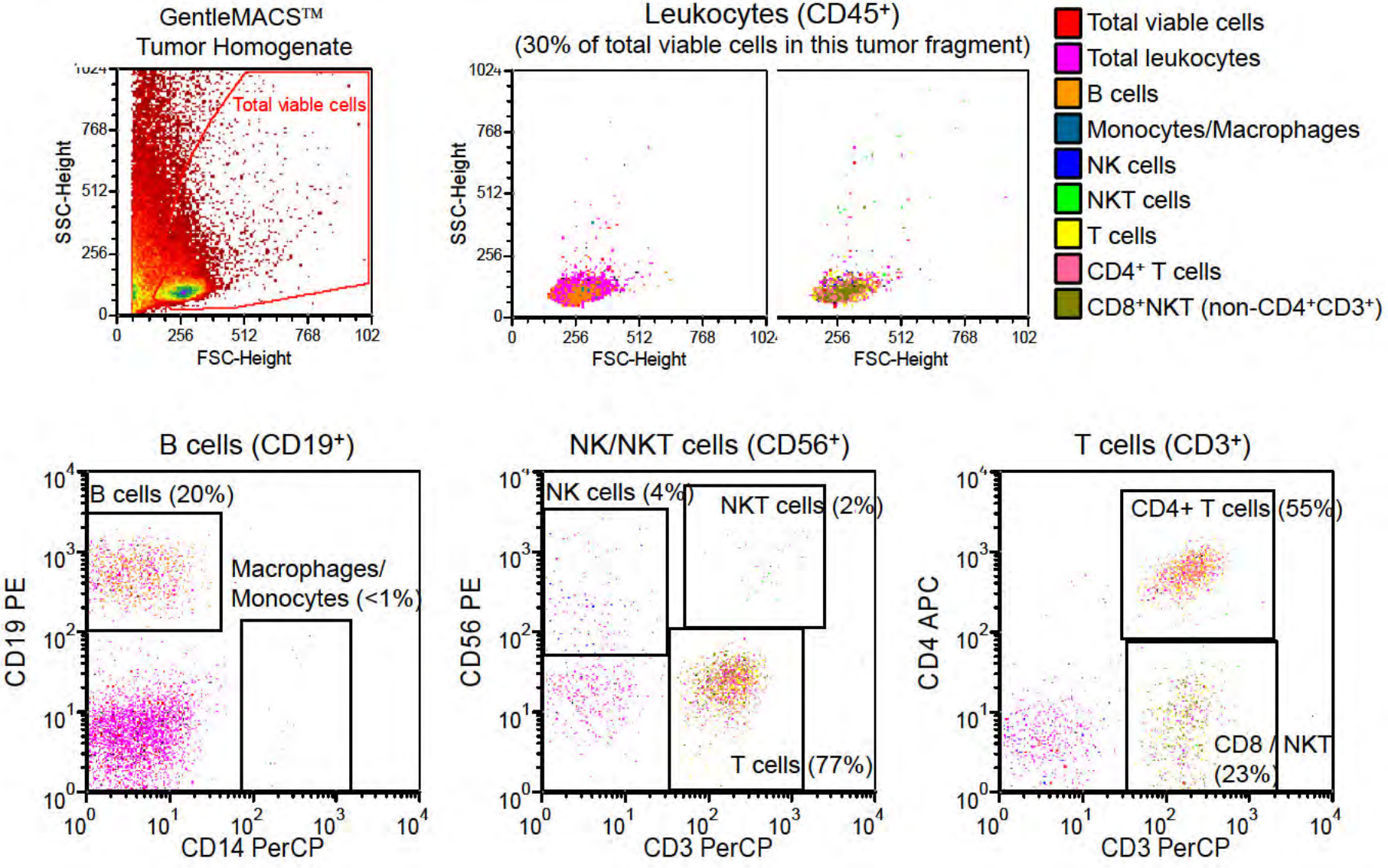
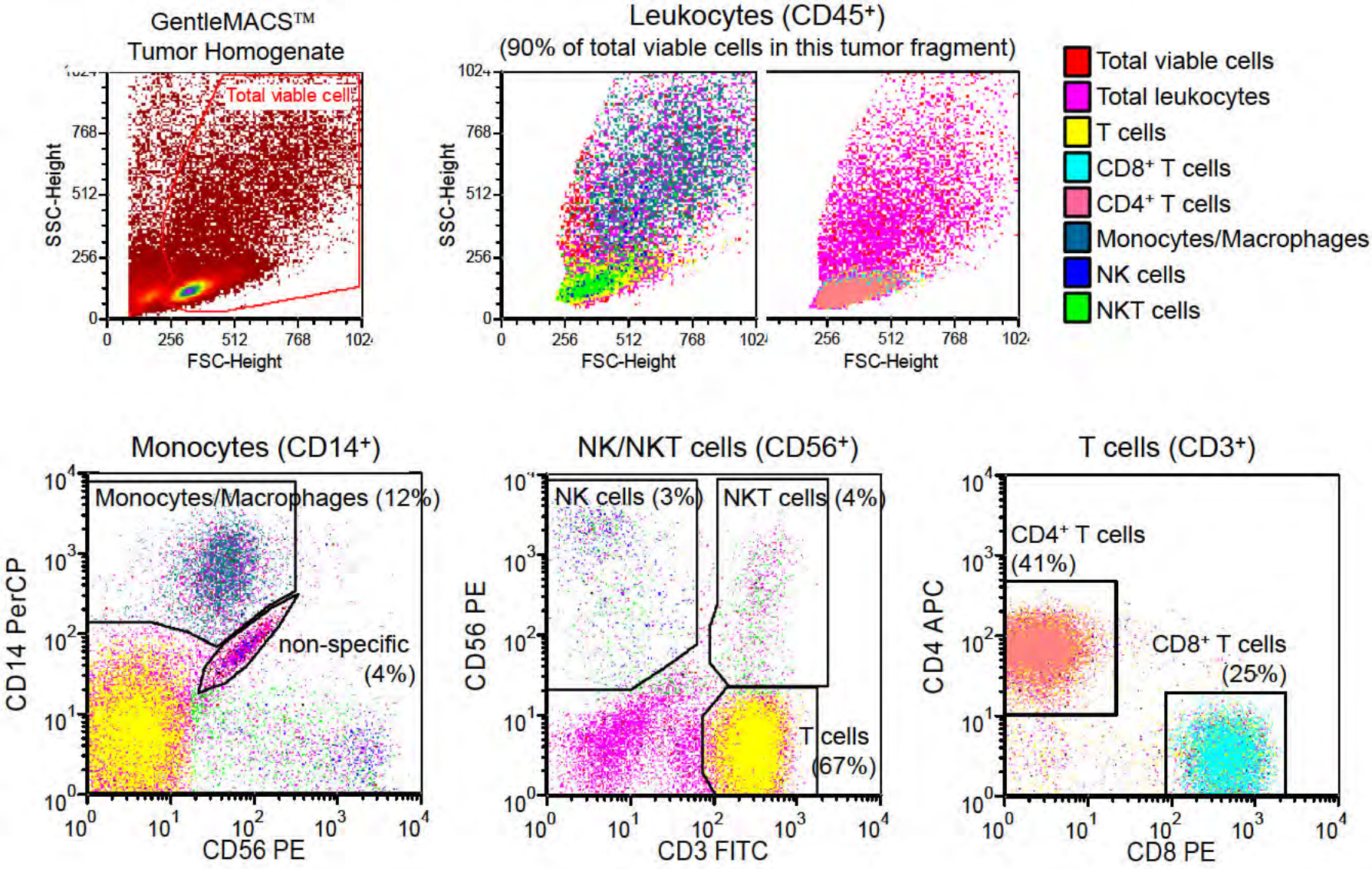


Figure S1B: Leukocyte Subpopulations in Fresh Breast Tumor Homogenates



% = percentage of gated CD45+ cells

Figure S1C: Leukocyte Subpopulations in Fresh Breast Tumor Homogenates



% = percentage of gated CD45+ cells

Figure S1D:
CD4⁺ T cells Infiltrating Breast Tumors

CD4⁺ T cells in tumor tissue

Patient ID[a]	% CD4⁺ in total CD3⁺ [b]	Patient ID[a]	% CD4⁺ in total CD3⁺ [b]
TIL003	31	TIL045	57
TIL004	46	TIL046	76
TIL007	44	TIL048	57
TIL008	43	TIL049	83
TIL011	47	TIL050	65
TIL012	61	TIL053	48
TIL013	55	TIL054	54
TIL014	56	TIL060	50
TIL015	53	TIL065	22
TIL016	50	TIL066	41
TIL019	48	TIL068	41
TIL025	29	TIL069	64
TIL039	47	TIL128	54
TIL040	72	TIL129	62
TIL041	62	TIL131	57
TIL042	53	TIL132	61
TIL043	51	TIL133	72
TIL044	46	Mean = 53% ± 13	

[a]Patient information is detailed in Table S1C

[b]Percentage of CD4⁺ cells within the gated CD3⁺ lymphocyte population assessed by flow cytometry

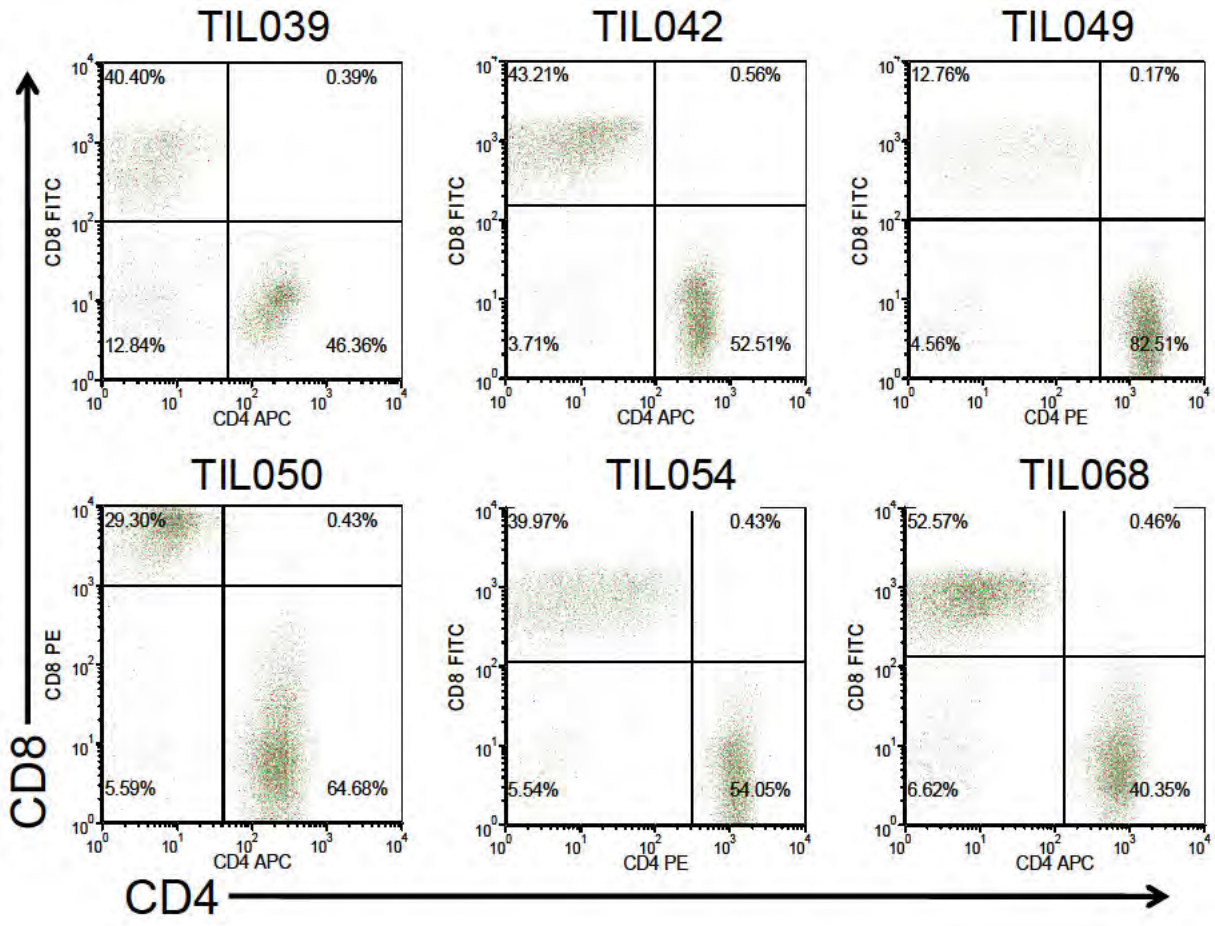


Figure S1E:
CD4⁺ T cells Infiltrating Non-Adjacent Normal Breast Tissue

CD4⁺ T cells in non-adjacent normal tissue

Patient ID[a]	% CD4⁺ in total CD3⁺ [b]	Patient ID[a]	% CD4⁺ in total CD3⁺ [b]
TIL003	54	TIL043	39
TIL004	46	TIL044	10
TIL007	38	TIL045	32
TIL008	35	TIL046	54
TIL011	41	TIL048	50
TIL012	27	TIL050	46
TIL013	57	TIL053	23
TIL014	42	TIL054	21
TIL015	68	TIL060	31
TIL016	43	TIL066	49
TIL019	41	TIL069	50
TIL025	31	TIL128	17
TIL039	35	TIL129	48
TIL040	19	TIL131	32
TIL041	28	TIL133	17
TIL042	46	Mean = 38% ± 13	

[a]Patient information is detailed in Table S1C

[b]Percentage of CD4⁺ cells within the gated CD3⁺ lymphocyte population assessed by flow cytometry

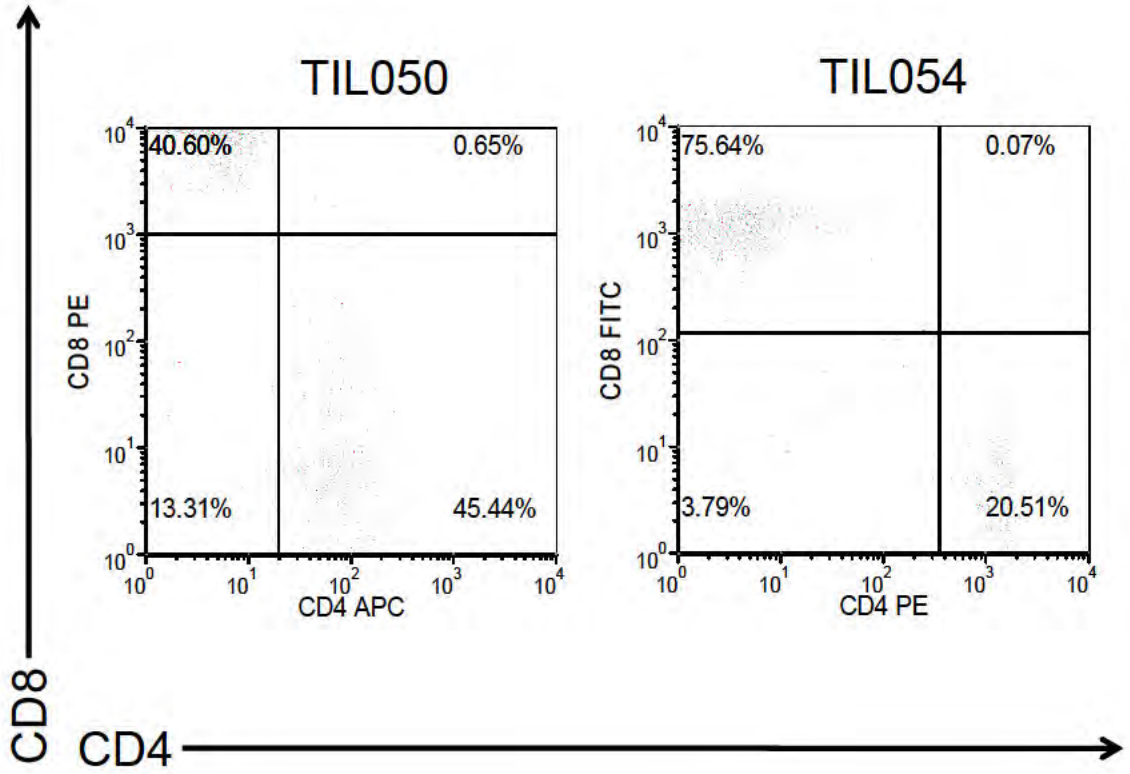


Figure S1F:
Antibodies used in the flow cytometry experiments to identify cellular subpopulations in homogenates of human breast tumor fragments

Antibody	Alternate name	Gene symbol	Fluorochrome	Source	Reference	Page
EpCAM	CD326, TACSTD1	EPCAM	FITC	Myltenyi Biotec	130-080-301	2
CD3		CD3E	PerCP	BD	347344	3
CD4		CD4	PE	BD	345769	4
CD4		CD4	APC	BD Pharmingen	555349	3
CD8		CD8A/CD8B	FITC	Beckman Coulter	A07756	4
CD8		CD8A/CD8B	PE	BD Pharmingen	555367	4
CD10		MME	PE	BD Pharmingen	555375	2
CD14		CD14	PerCP	BD	345786	3
CD19		CD19	PE	BD	345777	3
CD45		PTPRC	PerCP	BD	345809	2
CD45		PTPRC	APC	BD Pharmingen	555485	3
CD56		NCAM1	PE	Beckman Coulter	A07788	3

Figure S2:

Flow cytometric analysis of surface marker expression
on $CD4^+$ tumor infiltrating lymphocytes TIL
from human breast tumors*

(Patients from the Confirmation Set; Table S1C)

*The images shown represent duplicate or triplicate experiments obtaining similar results

Summary of surface expression for conventional and newly defined Th markers commonly altered on the TIL (our flow cytometry/microarray data) and Th subsets (public microarray data)

			Public Data*										Our Data		
Name	Alternate name	Gene Symbol	Th1	Th2	Tfh	Treg	Th17	CD25+	Tem	Tcm	Memory vs Naïve	Name	TIL (FACS)	≠ on TIL	TIL (Array)
CD45RO		PTPRC (isoform)									Up#	CD45RO	Up		
CD6						Down						CD6	Down		-2.18
CD7							Down		Down	Down		CD7	Down	yes	
CD18	LFA-1β	ITGB2	Down	Down	Up					Up		CD18	Up/Down		-2.94
CD25	IL-2Rα	IL2RA	Up	Up		Up		Up				CD25	Up		3.36
CD26		DPP4	Up		Down	Down		Down				CD26	Up/Down	yes	-3.94
CD27			Down	Down					Down			CD27	Down		
CD38			Up									CD38	Up/Down	yes	
CD43		SPN		Up								CD43	Down		
CD54		ICAM1	Up									CD54	Up		3.42
CD55	DAF				Down			Down				CD55	Down		-2.71
CD58			Up			Up		Up			Up	CD58	Up	yes	3.03
CD62L		SELL		Down					Down			CD62L	Down		-2.80
CD69					Up							CD69	Up		
CD71		TFRC	Up	Up		Up						CD71	Up		2.46
CD86			Up									CD86	Up	yes	
CD200	OX2				Up							CD200	Up/Down	yes	11.24
CCR4	CD194			Up#		Up	Up				Up	CCR4	Up		
CCR5	CD195		Up#			Up#			Up		Up	CCR5	Up		3.79
CCR7	CD197				Down	Down			Down			CCR7	Down		
CXCR3	CD183							Up	Up		Up	CXCR3	Up		2.80
CXCR4	CD184		Down	Down								CXCR4	Up		
CXCR5	CD185	BLR1			Up						Up	CXCR5	Up/Down		
BTLA	CD272				Up	Down						BTLA	Up/Down		2.51
CTLA4	CD152		Up			Up		Up				CTLA4	Up		4.87
FAS	CD95					Up		Up			Up	FAS	Up		2.07
GITR	AITR	TNFRSF18	Up			Up#	Up					GITR	Up	yes	
HLA-DR		HLA-DRA/B1-6				Up	Up		Up		Up	HLA-DR	Up		3.77
ICOS	CD278		Up		Up	Up#						ICOS	Up	yes	2.83
IL1R1	CD121α						Up				Up	IL1R1	Up		9.59
IL2Rβ	CD122	IL2RB						Up				IL2Rβ	Up		2.50
IL6ST	CD130			Down	Up						Down	IL6ST	Down		
IL12Rβ2		IL12RB2	Up			Up		Up				IL12Rβ2	Up		5.84
OX40	CD134	TNFRSF4	Up								Up	OX40	Up		
PD1		PDCD1			Up							PD1	Up	yes	
TGFβR2		TGFBR2					Down	Down			Down	TGFβR2	Up		-2.88
TGFβR3		TGFBR3							Up		Up	TGFβR3	Up		

*Public data from Table S3 #Conventional Th markers not included in Table S3 §Variable expression pattern on TIL from different patients

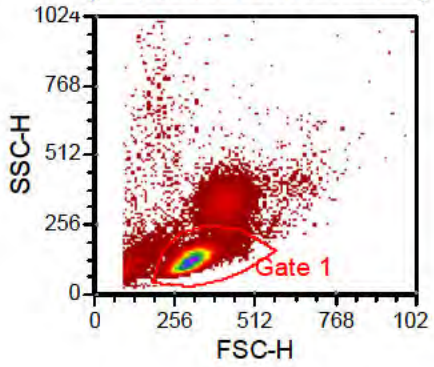
These data confirm that the CD4+ TIL are activated, effector memory cells with all major Th subsets infiltrating the tumors

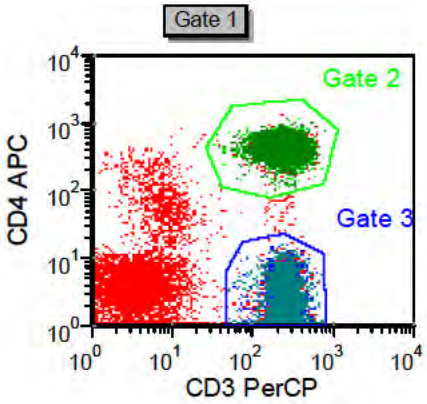
Upregulated
Downregulated
D-PB
TIL

Gu-Trantien, *et al.*

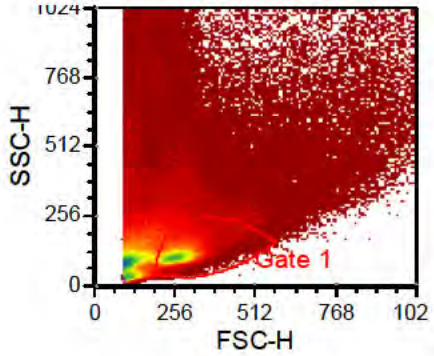
Gating Parameters

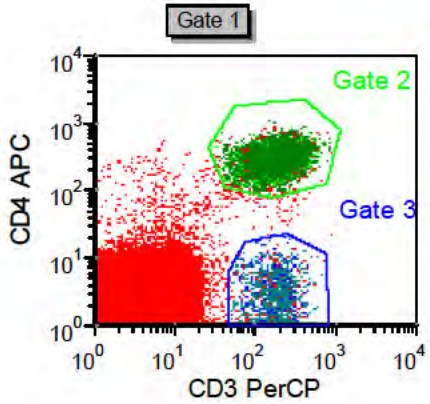
Healthy Donor Blood
(PB; total PBMC shown)





GentleMACS™
Tumor Homogenate

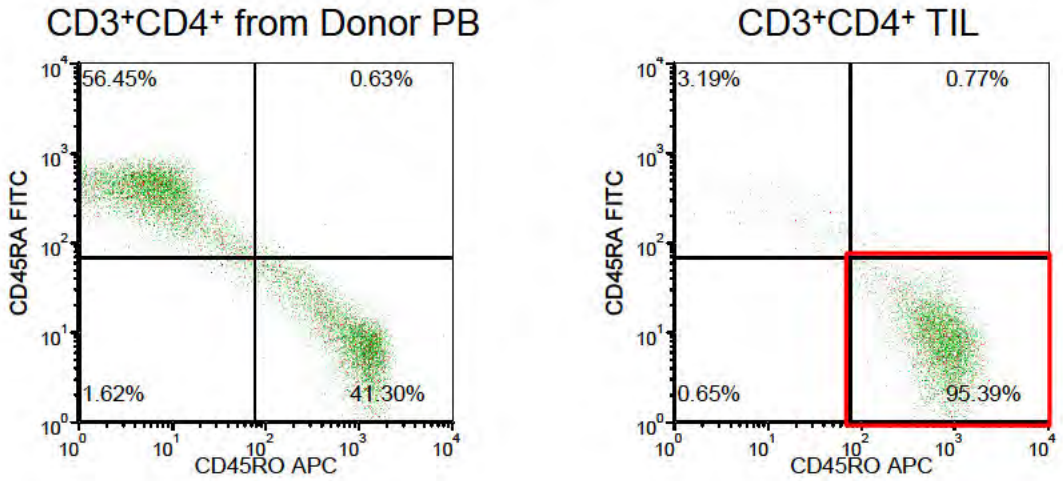




$CD3^+CD4^+$
= Gate 1 plus Gate 2

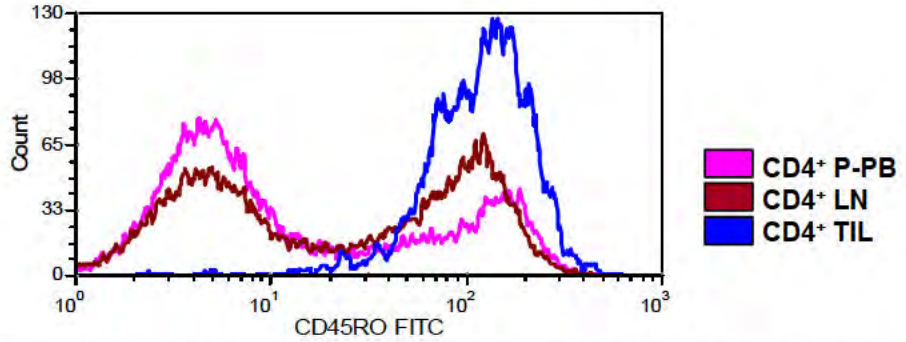
$CD3^+CD4^-$ (majority $CD8^+$)
= Gate 1 plus Gate 3

The number of CD45RA⁺ and CD45RO⁺ CD4⁺ T cells varies in healthy donor PB between 40-60% depending upon their age



CD4⁺ T cells infiltrating 21 breast tumors are 97 ± 2% CD45RO⁺

consistent with our microarray data



The number of CD4⁺CD45RO⁺ T cells is lowest in PB with a slight increase detected in the lymph node (LN) but both well below that observed in the TIL.

Definition of surface expression levels
(TIL = $CD4^+$ TIL; D-PB = $CD4^+$ D-PB)

lo = positive on TIL and D-PB and downregulated on TIL versus D-PB

hi = significantly upregulated on TIL compared to D-PB

± = TIL from some tumors were found to have increased expression while others had decreased expression levels relative to D-PB

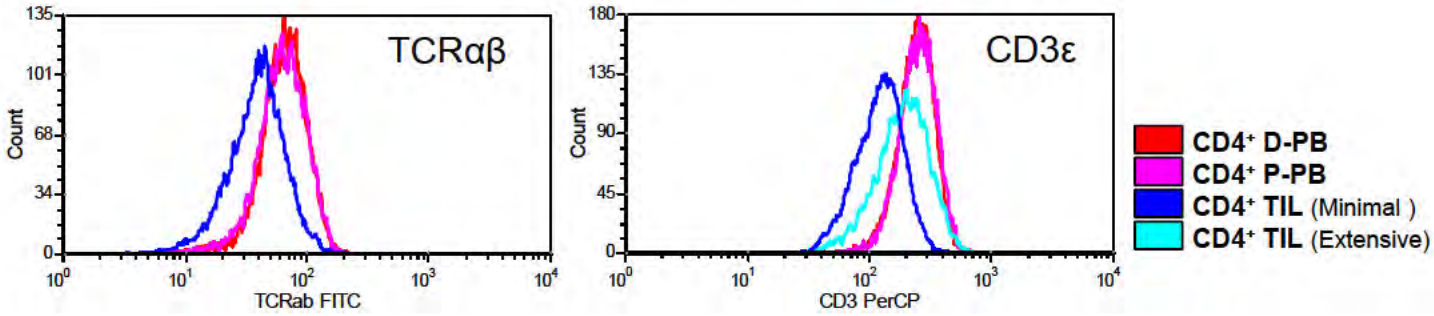
– = negative on TIL

+ = 1. upregulated on TIL (low to medium increase) compared to D-PB and/or
2. positive on TIL with no obvious up- or down-regulation compared to D-PB

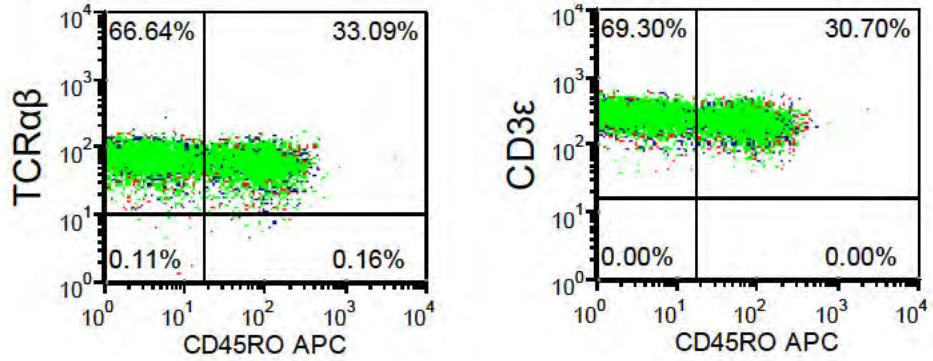
Flow Cytometric Analysis: TCR^{lo}/CD3^{lo}

Conclusion:
TCR/CD3 is downregulated on the CD4$^+$ TIL and to a greater extent on TIL from minimally infiltrated tumors.

CD3$^+$CD4$^+$ from Donor PB and Patient PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) CD3$^+$CD4$^+$ T cells from Donor PB



Array data from Table S2B:
Down *CD247* (CD3ζ) (Fc = -2.04)

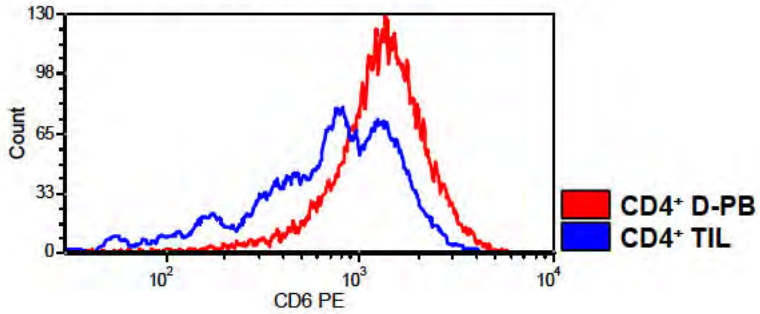
Array data from Table S2C:
Down CD3 (Fc = *CD247* -2.12; *CD3D* -2.06; *CD3E* -2.49; *CD3G* -2.32)
Down *TRA*@ (Fc = -2.44)

qRT-PCR confirmed TCR/CD3 pathway downregulation in minimally infiltrated relative to extensively infiltrated tumors and D-PB (Table S5B).

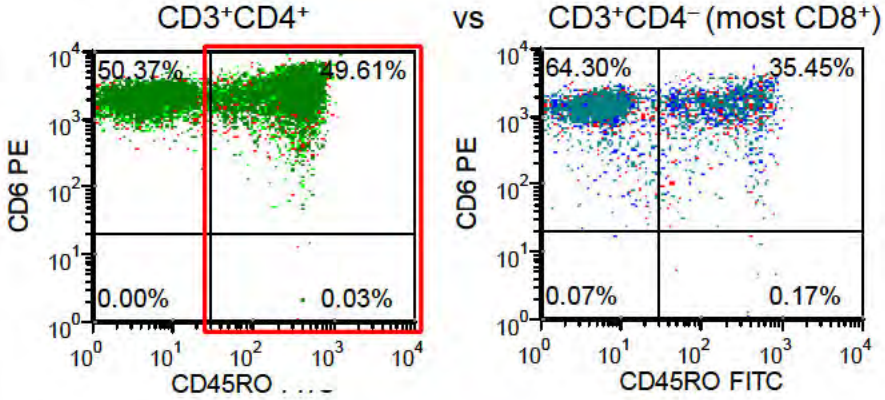
Flow Cytometric Analysis: CD6^{lo}

Conclusion:
CD6 is downregulated on the CD4$^+$ TIL.

CD3$^+$CD4$^+$ from Donor PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB

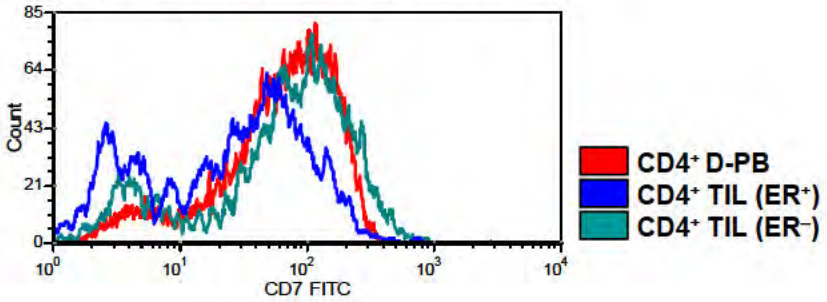


Array data from Table S2B:
Down (Fc = -2.18)

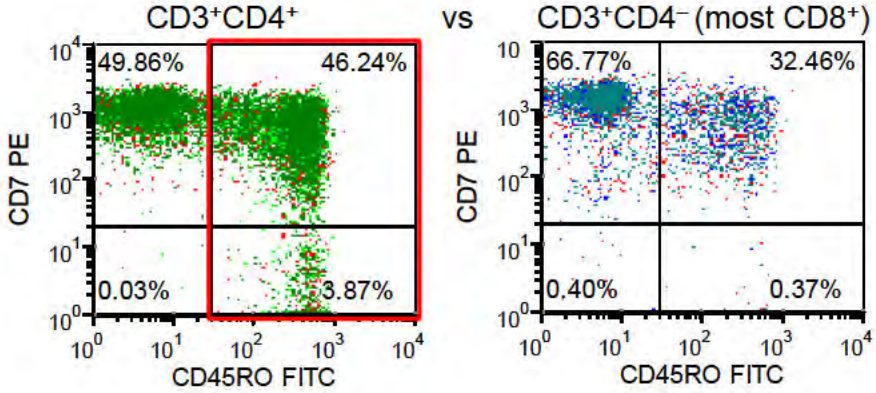
Flow Cytometric Analysis: CD7^{lo}

Conclusion:
CD7 expression is altered on the CD4$^+$ TIL, with changes variable between tumors.

CD3$^+$CD4$^+$ from Donor PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB

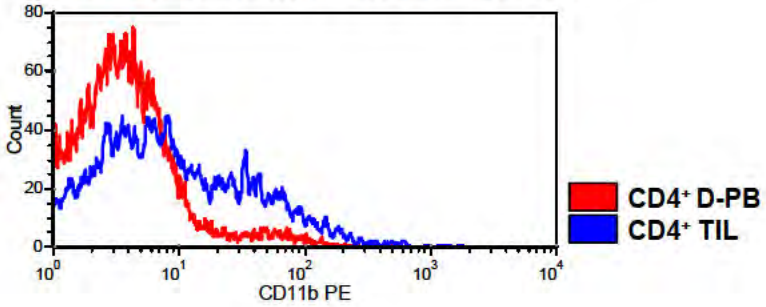


Array data from Table S2C:
Down (Fc = -2.05)

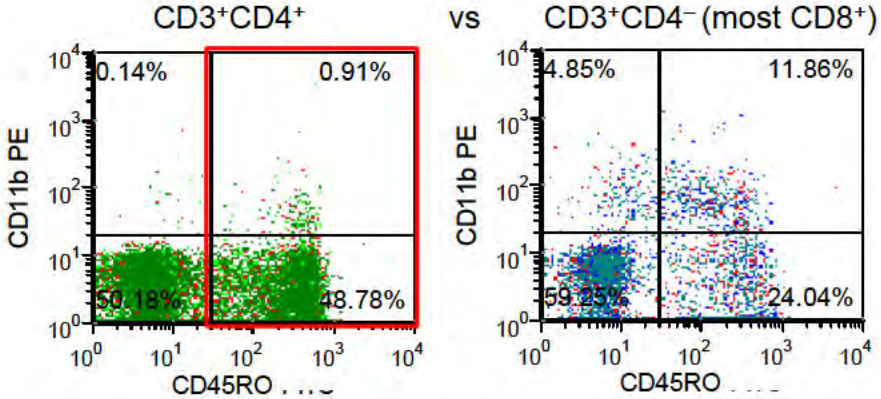
Flow Cytometric Analysis: CD11b$^+$ (*ITGAM*)

Conclusion:
CD11b expression is upregulated on a small subpopulation of CD4$^+$ T cells in healthy donors; In the TIL, this subpopulation increases along with overall CD11b expression.

CD3$^+$CD4$^+$ from Donor PB vs TIL



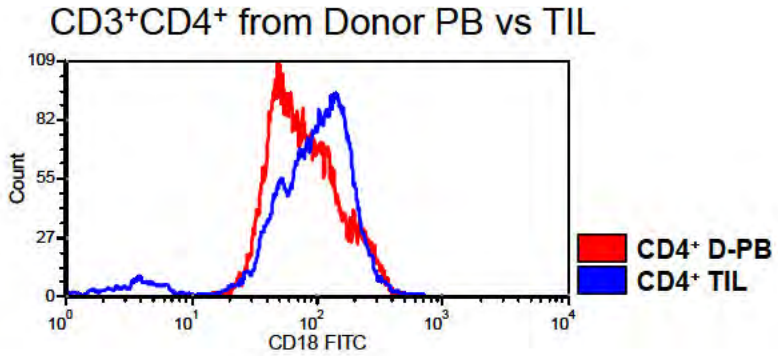
Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB



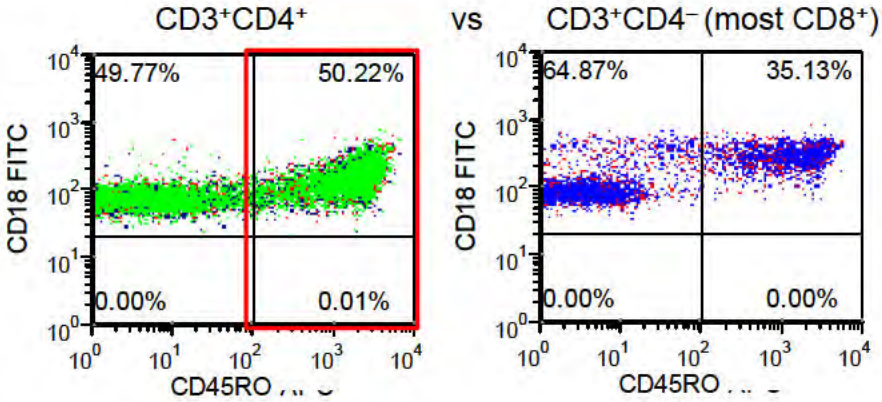
Array data from Table S2B:
Down (Fc = -2.37)

Flow Cytometric Analysis: CD18± (*ITGB2*)

Conclusion:
CD18 expression is up-regulated on some CD4⁺CD45RO⁺ T cells from healthy donors and this population slightly increases in the TIL



Expression on CD45RO⁺ vs CD45RA⁺(=RO⁻) T cells from Donor PB

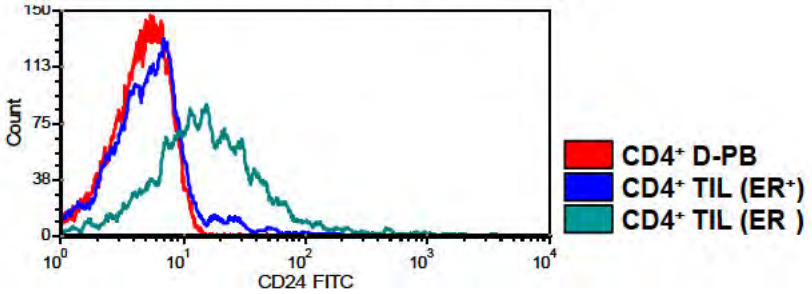


Array data from Table S2B:
Down (Fc = -2.94)

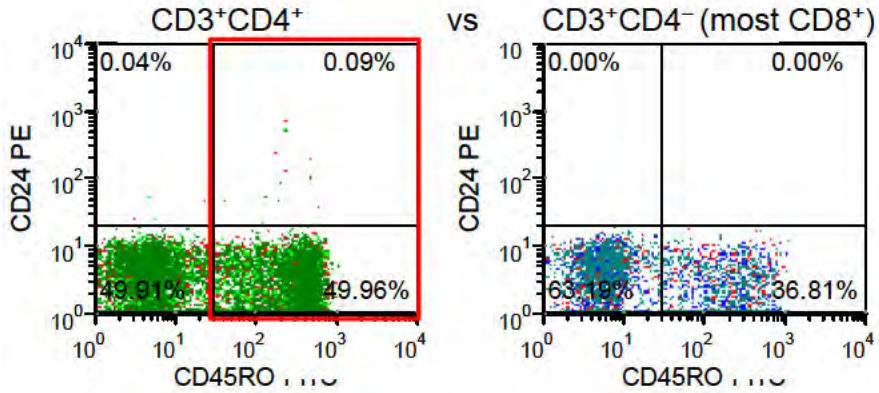
Flow Cytometric Analysis: CD24$^+$

Conclusion:
CD24 expression is upregulated on CD4$^+$ TIL, to a greater extent on ER$^-$ compared to ER$^+$ tumors.

CD3$^+$CD4$^+$ from Donor PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB

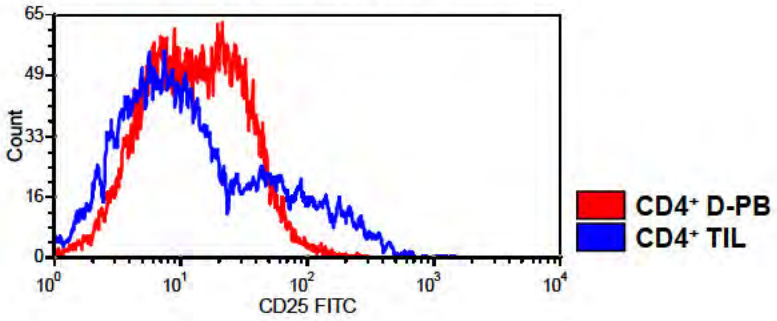


Array data from Table S2B:
Up (Fc = +86.48)

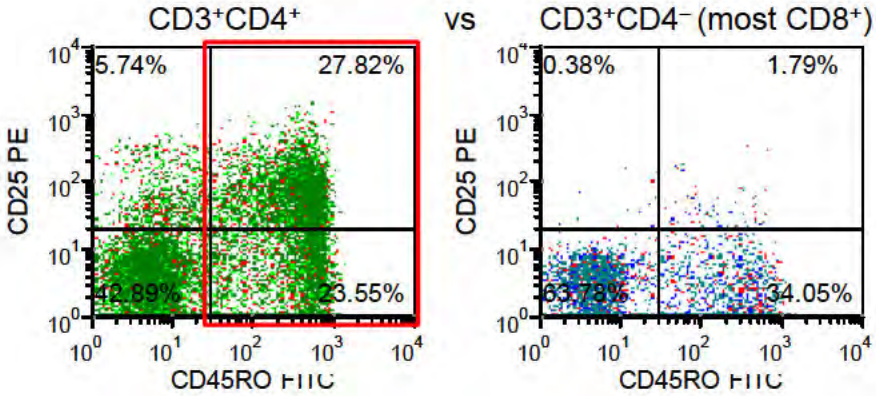
Flow Cytometric Analysis: CD25⁺ (*IL2RA*)

Conclusion:
CD25 expression is upregulated on a sub. o. ulation of CD4⁺ TIL.

CD3⁺CD4⁺ from Donor PB vs TIL



Expression on CD45RO⁺ vs CD45RA⁺(=RO⁻) T cells from Donor PB

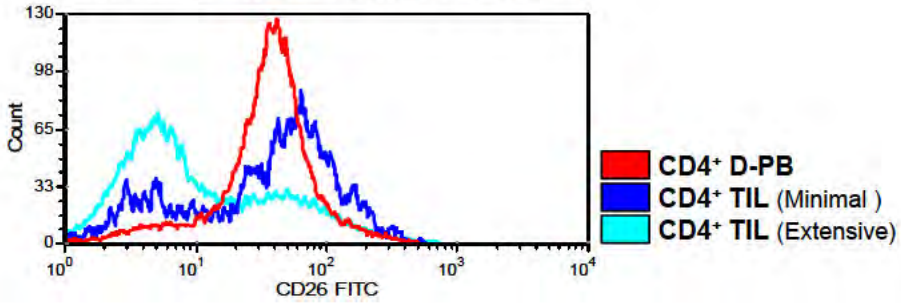


Array data from Table S2B:
Up (Fc = +3.36)

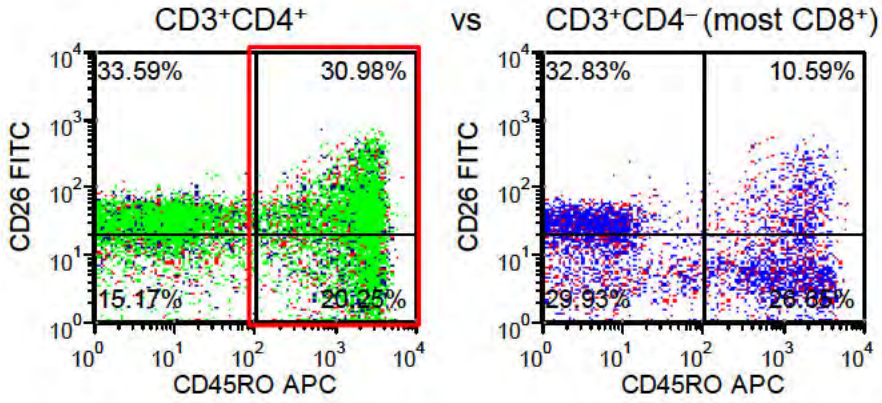
Flow Cytometric Analysis: CD26± (*DPP4*)

Conclusion:
CD26 expression is downregulated on the majority of TIL from an extensively infiltrated tumor but only a minority from a minimally infiltrated tumor. A subpopulation of CD4+ TIL in minimally infiltrated tumors display CD26 upregulation.

CD3+CD4+ from Donor PB vs TIL



Expression on CD45RO+ vs CD45RA+(=RO−) T cells from Donor PB



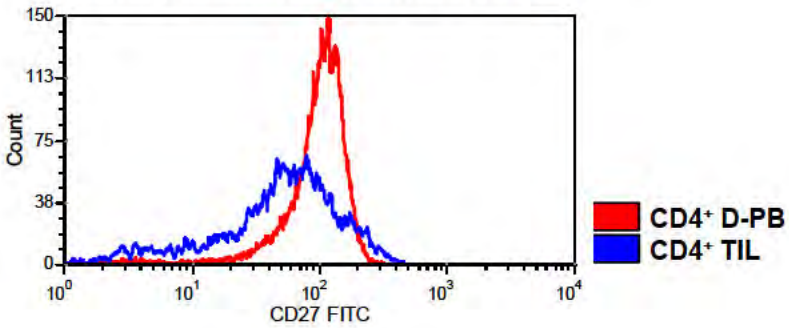
Array data from Table S2B:
Down (Fc = -3.94)

qRT-PCR did not detect a correlation between *DPP4* mRNA expression and the extent of lymphocyte infiltration (Figure 8; Table S5D).

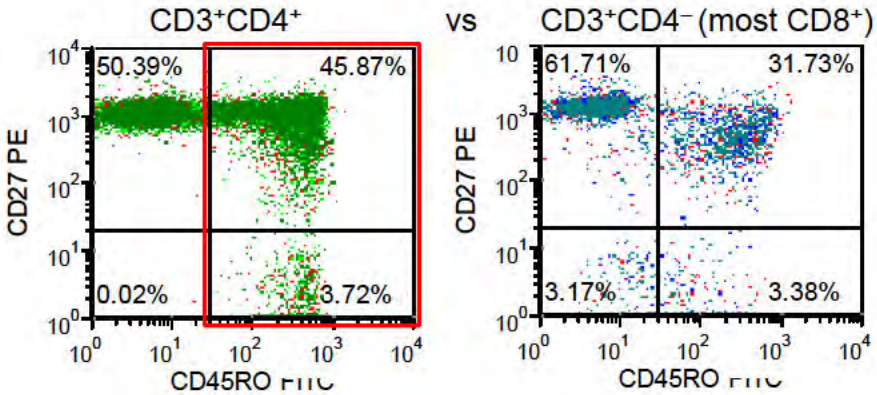
Flow Cytometric Analysis: CD27^{lo}

Conclusion:
CD27 expression is downregulated on the CD4$^+$ TIL.

CD3$^+$CD4$^+$ from Donor PB vs TIL



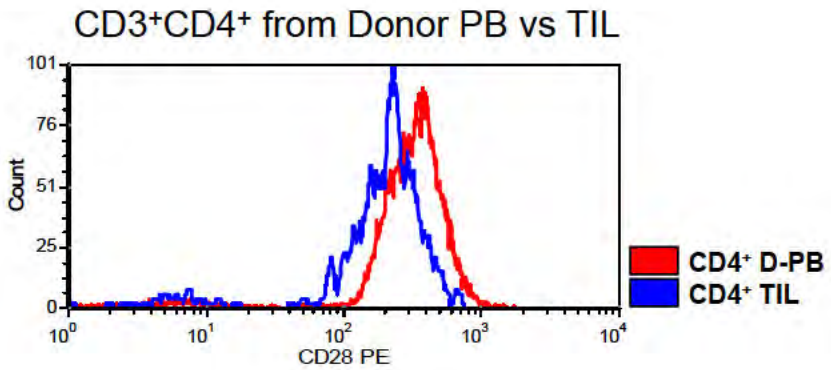
Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB



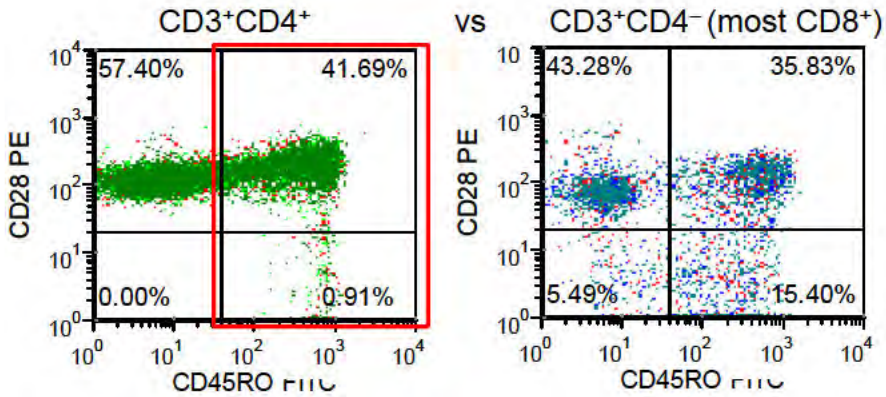
Array data from Table S2C:
Down (Fc = -2.50)

Flow Cytometric Analysis: CD28^{lo}

Conclusion:
CD28 expression is downregulated on CD4$^+$ TIL.

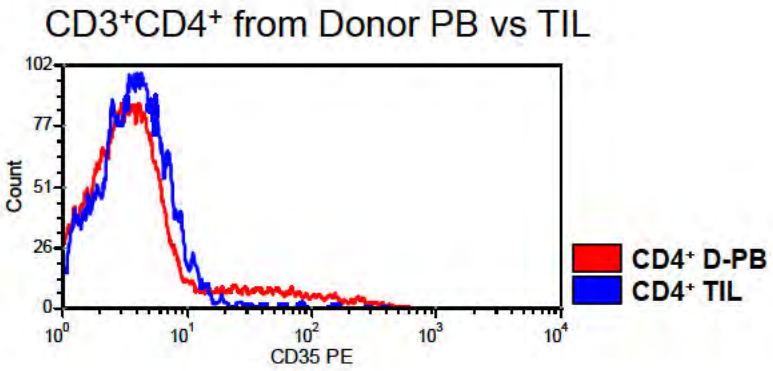


Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB

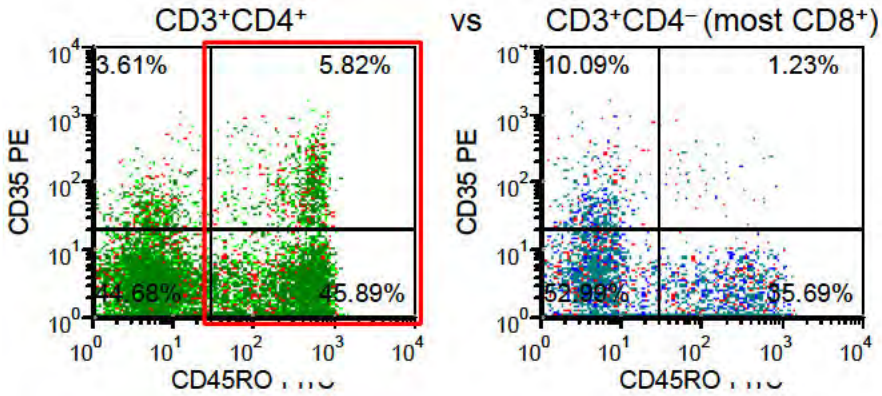


Flow Cytometric Analysis: CD35⁻ (*CR1*)

Conclusion:
CD35 expression is detected on a minority of CD4⁺ T cells in donor blood, and this CD35⁺ subpopulation appears to be lost in the TIL.



Expression on CD45RO⁺ vs CD45RA⁺(=RO⁻) T cells from Donor PB

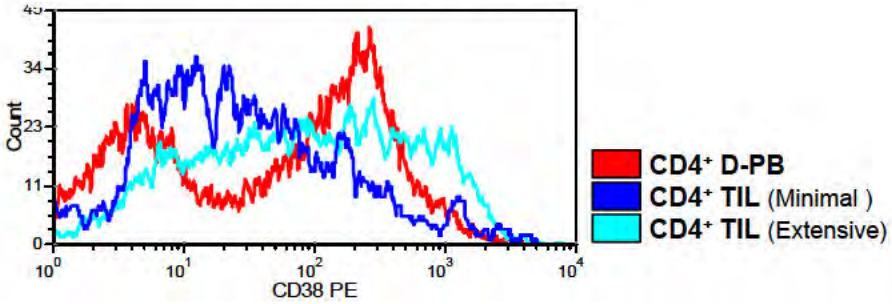


Array data from Table S2B:
Down (Fc = -4.17)

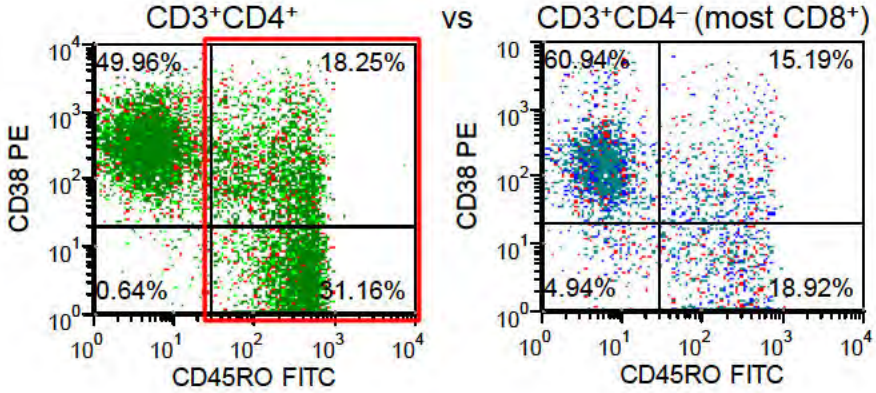
Flow Cytometric Analysis: CD38±

Conclusion:
CD38 expression is heterogeneous, with a significant proportion o. CD4+ TIL differentially displaying decreased expression in minimally infiltrated tumors and increased expression in extensively infiltrated tumors.

CD3+CD4+ from Donor PB vs TIL



Expression on CD45RO+ vs CD45RA+(=RO−) T cells from Donor PB

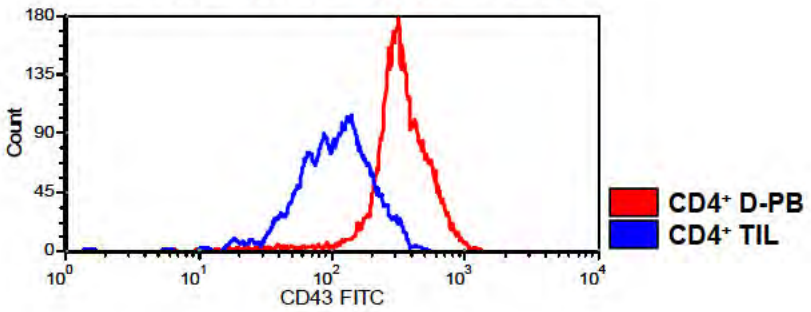


qRT-PCR confirmed *CD38* upregulation in extensively infiltrated relative to minimally infiltrated tumors (Figure 8; Table S5D).

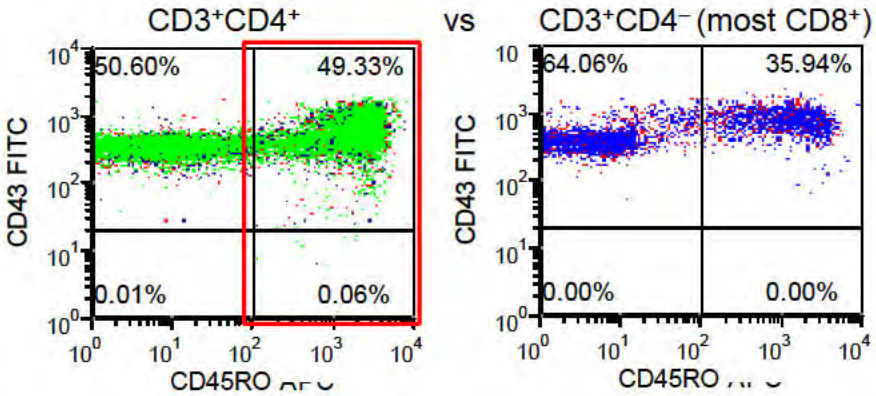
Flow Cytometric Analysis: CD43^{lo} (*SPN*)

Conclusion:
CD43 is downregulated on the CD4$^+$ TIL.

CD3$^+$CD4$^+$ from Donor PB vs TIL



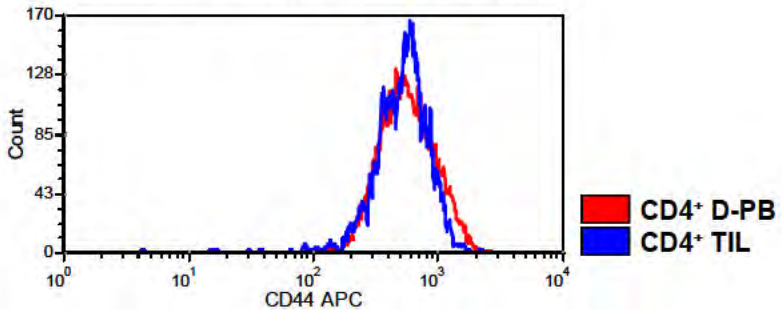
Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB



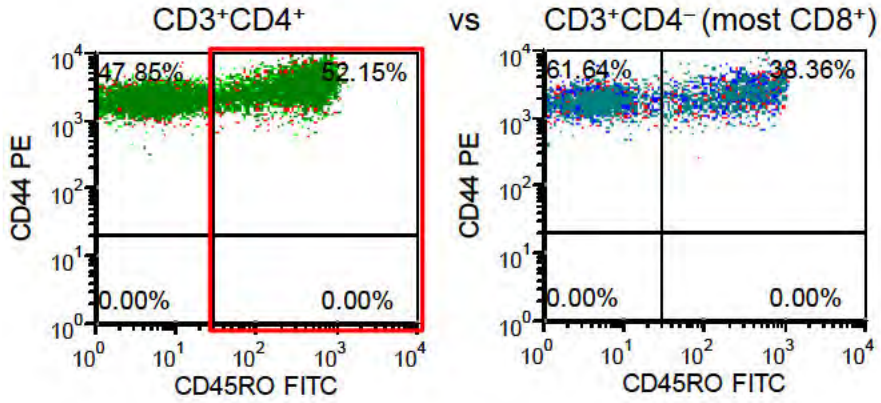
Flow Cytometric Analysis: CD44+

Conclusion:
CD44 expression is not significantly altered on the CD4+ TIL.

CD3+CD4+ from Donor PB vs TIL



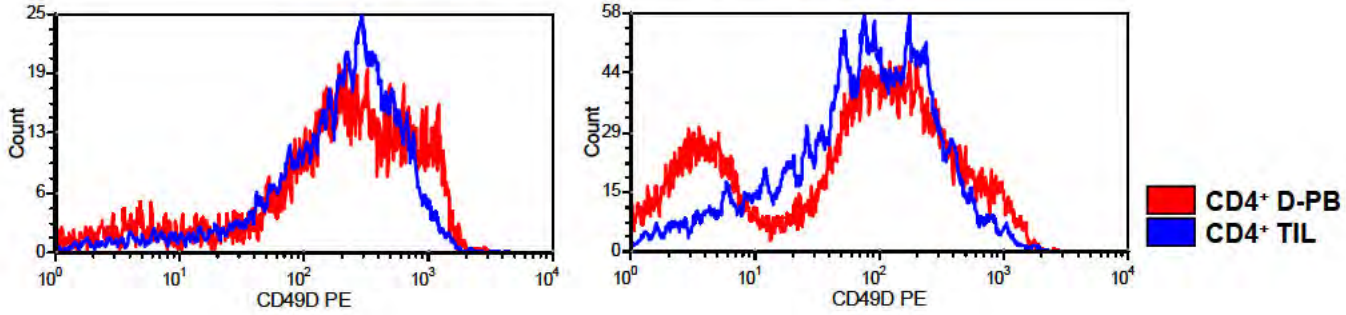
Expression on CD45RO+ vs CD45RA+(=RO−) T cells from Donor PB



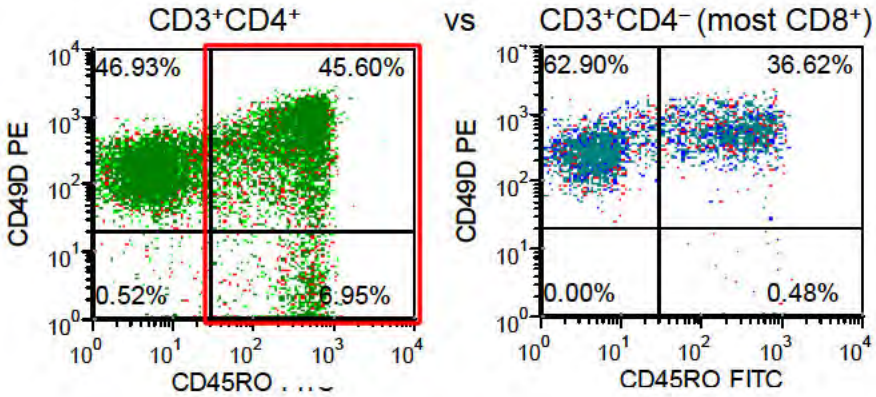
Flow Cytometric Analysis: CD49D± (*ITGA4*)

Conclusion:
Sub. o. ulations of CD49D^{hi} and CD49D^{lo} expressing cells are reduced in the CD4^{+} TIL.

CD3^{+}CD4^{+} from Donor PB vs TIL



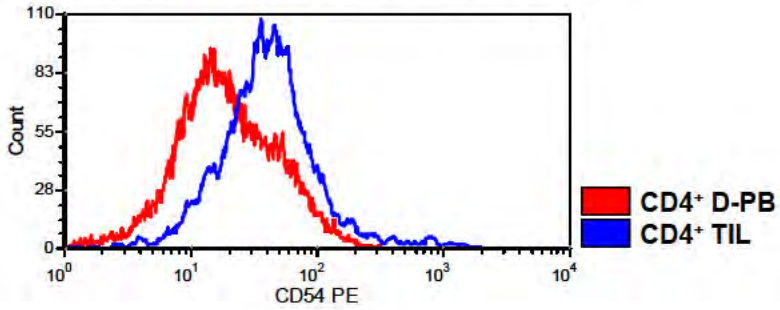
Expression on CD45RO^{+} vs CD45RA^{+}(=RO^{-}) T cells from Donor PB



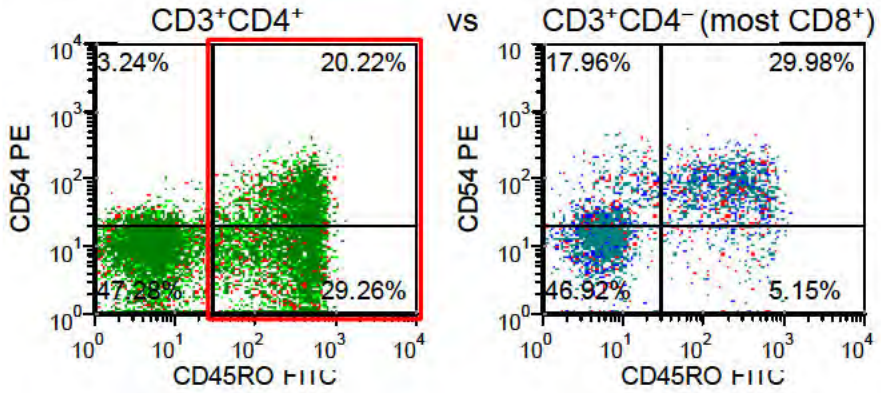
Flow Cytometric Analysis: CD54^{hi} (*ICAM1*)

Conclusion:
CD54 is upregulated on the CD4$^+$ TIL.

CD3$^+$CD4$^+$ from Donor PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB

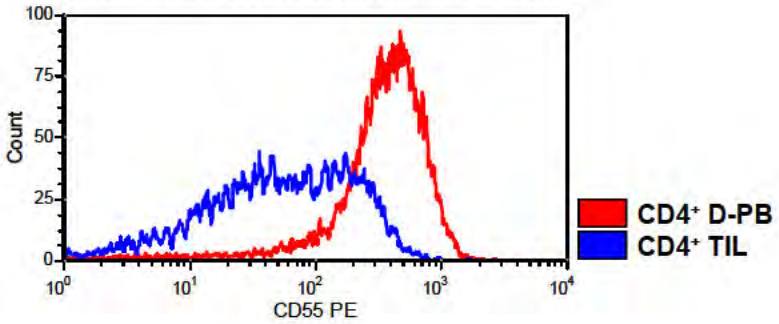


Array data from Table S2B:
Up (Fc = +3.42)

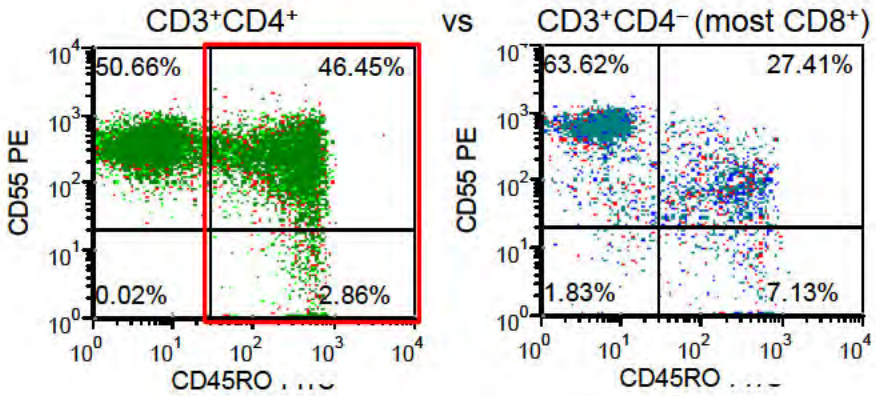
Flow Cytometric Analysis: CD55^{lo}

Conclusion:
CD55 is downregulated on the CD4$^+$ TIL.

CD3$^+$CD4$^+$ from Donor PB vs TIL



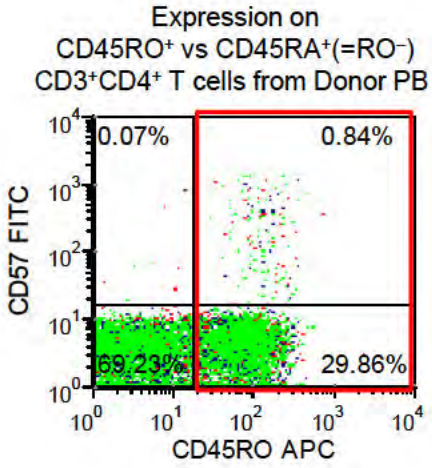
Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB



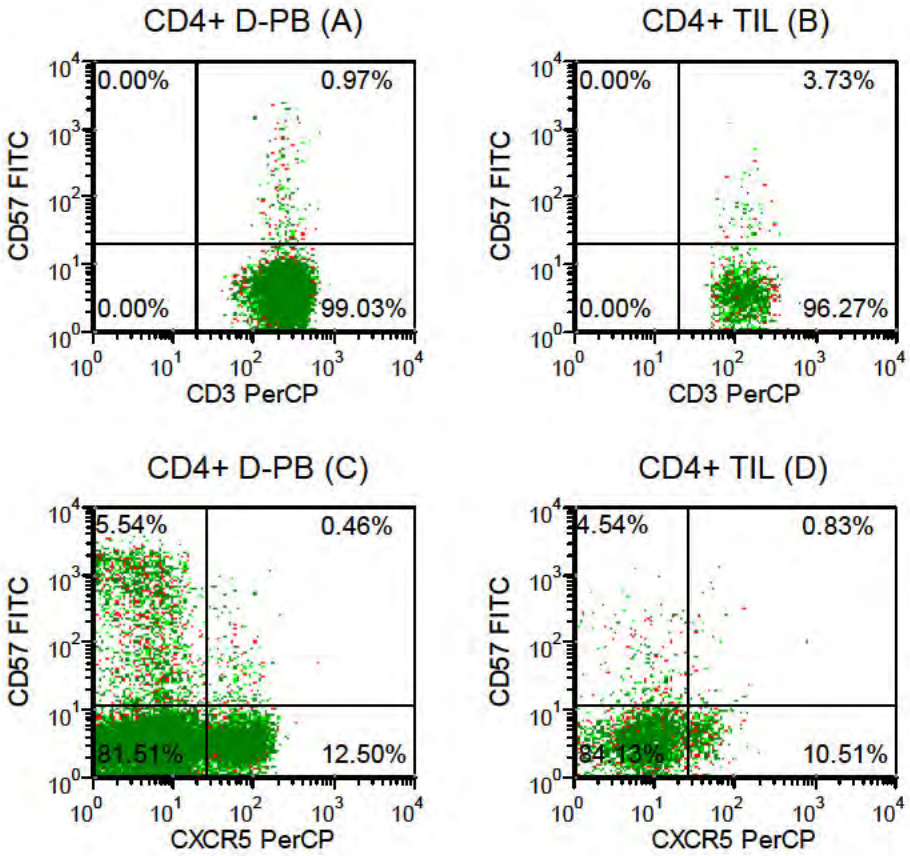
Array data from Table S2B:
Down (Fc = -2.71)

Flow Cytometric Analysis: CD57^{lo}

Conclusion:
The percentage of CD57$^+$ cells among CD4$^+$ TIL has a tendency to be higher than in donor blood, but cells expressing the highest CD57 levels are absent in the TIL.



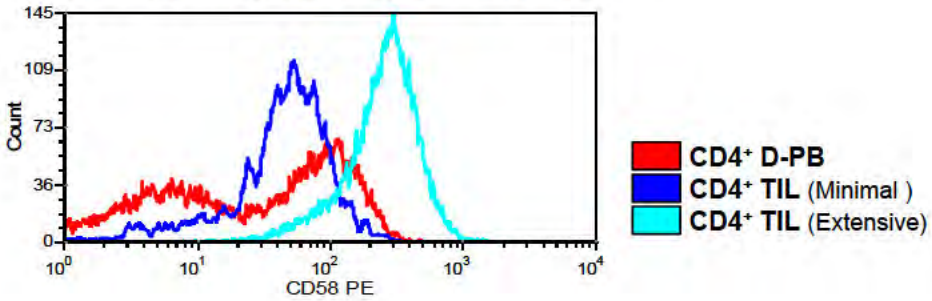
CD3$^+$CD4$^+$ from Donor PB vs TIL



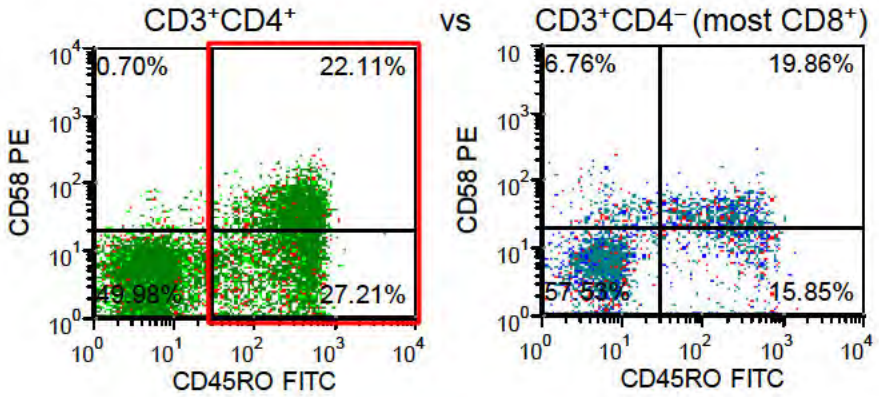
Flow Cytometric Analysis: CD58^{hi} (LFA-3)

Conclusion:
CD58 is upregulated on the CD4$^+$ TIL and to a greater extent in some extensively infiltrated tumors.

CD3$^+$CD4$^+$ from Donor PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB

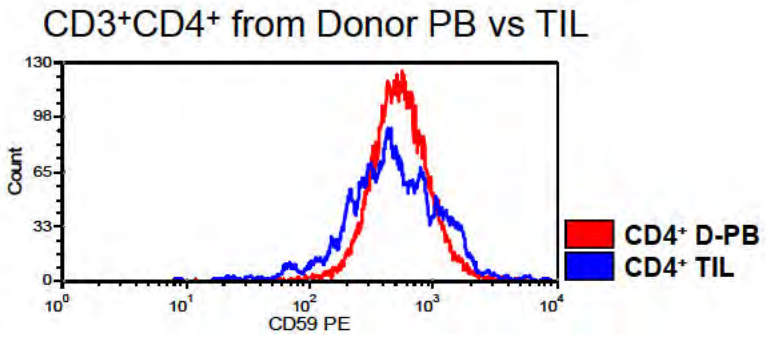


Array data from Table S2B:
Up (Fc = +3.03)

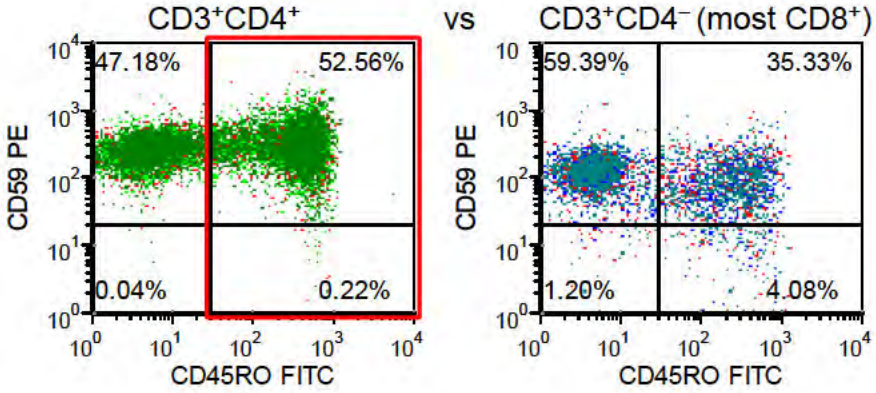
qRT-PCR confirmed *CD58* upregulation in extensively infiltrated relative to minimally infiltrated tumors (Figure 8; Table S5D).

Flow Cytometric Analysis: CD59$^+$

Conclusion:
CD59 expression is slightly altered on the CD4$^+$ TIL, with somewhat greater heterogeneity in surface density detected on the TIL.



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB

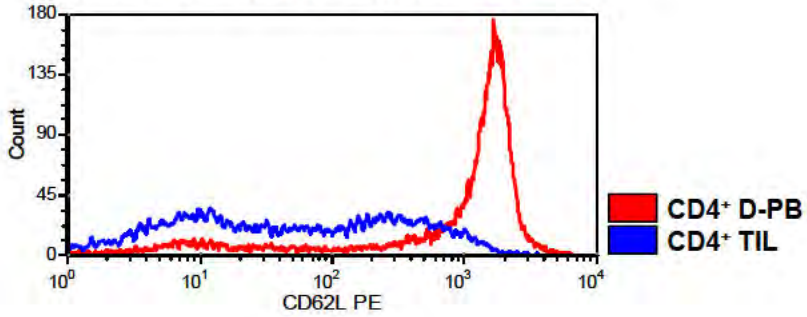


Array data from Table S2B:
Up (Fc = +4.53)

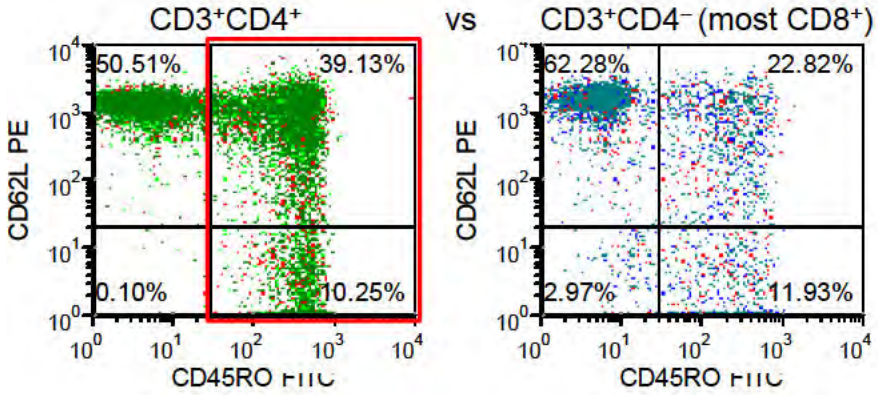
Flow Cytometric Analysis: CD62L^{lo} (*SELL*)

Conclusion:
CD62L is downregulated on the CD4$^+$ TIL.

CD3$^+$CD4$^+$ from Donor PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB

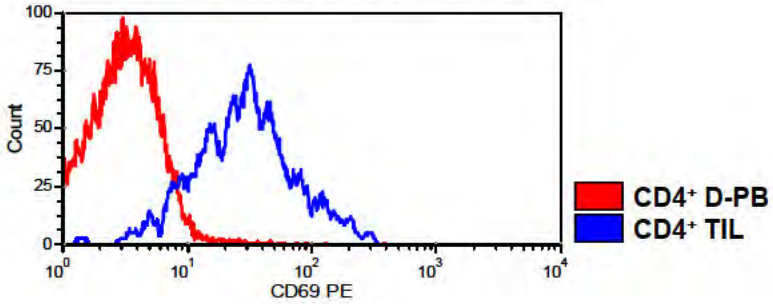


Array data from Table S2B:
Down (Fc = -2.80)

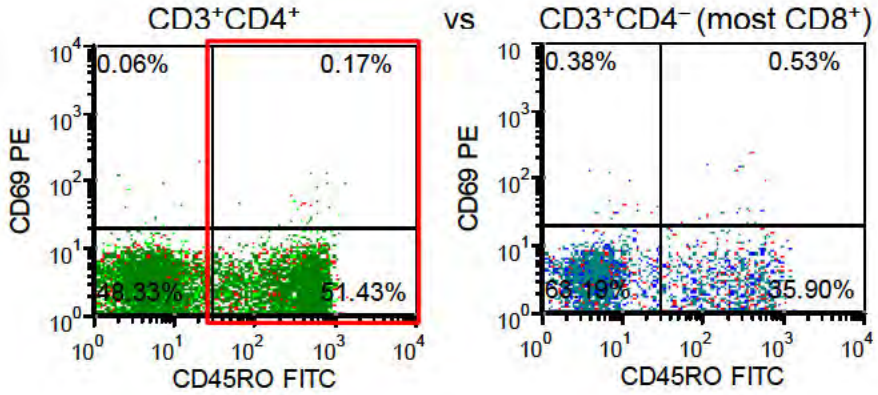
Flow Cytometric Analysis: CD69$^+$

Conclusion:
CD69 expression is upregulated on the surface of CD4$^+$ TIL.

CD3$^+$CD4$^+$ from Donor PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB



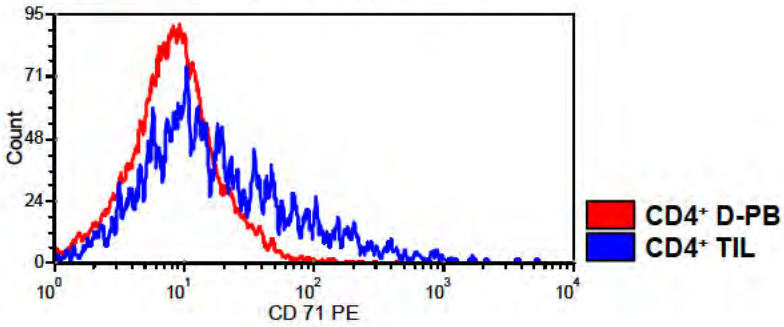
Array data from Table S2B:

Down (Fc = -2.04)

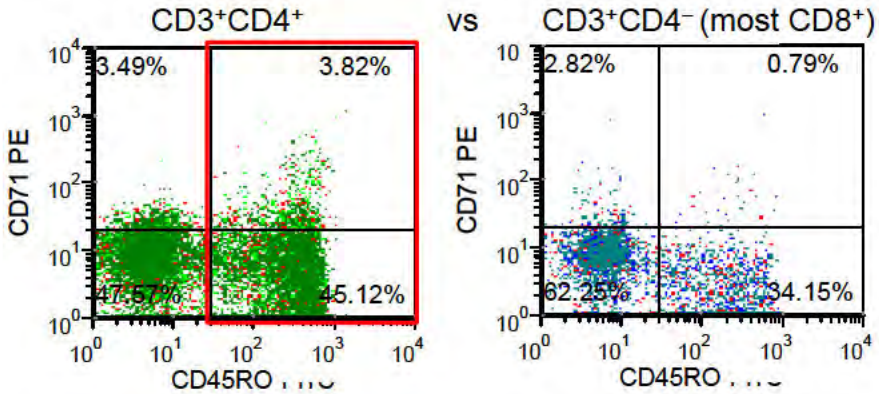
Flow Cytometric Analysis: CD71$^+$ (*TFRC*)

Conclusion:
CD71 is upregulated on the CD4$^+$ TIL.

CD3$^+$CD4$^+$ from Donor PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB

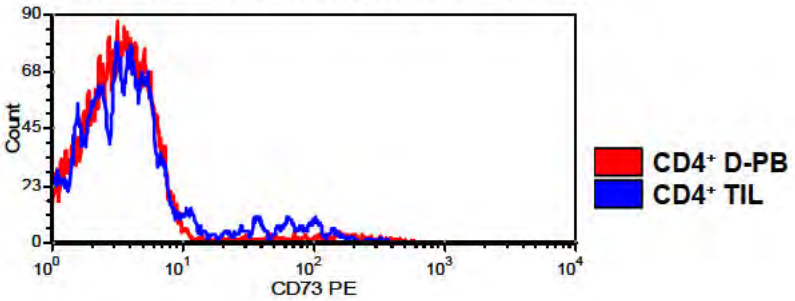


Array data from Table S2B:
Up (Fc = +2.46)

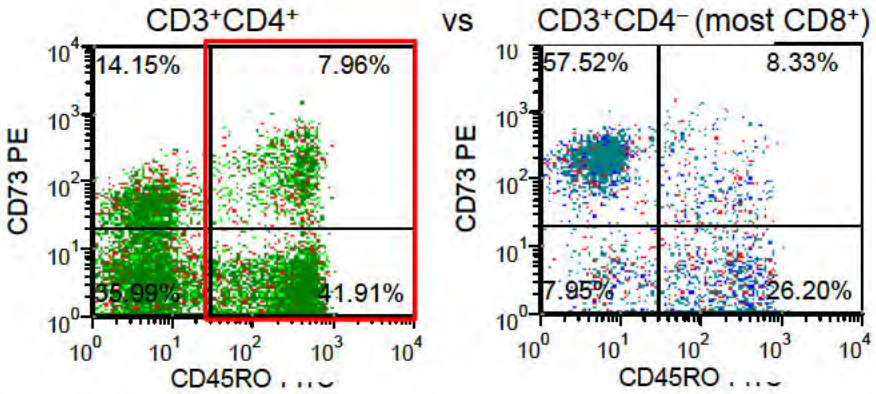
Flow Cytometric Analysis: CD73$^+$ (*NT5E*)

Conclusion:
The CD73$^+$ subpopulation slightly increases in the TIL.

CD3$^+$CD4$^+$ from Donor PB vs TIL



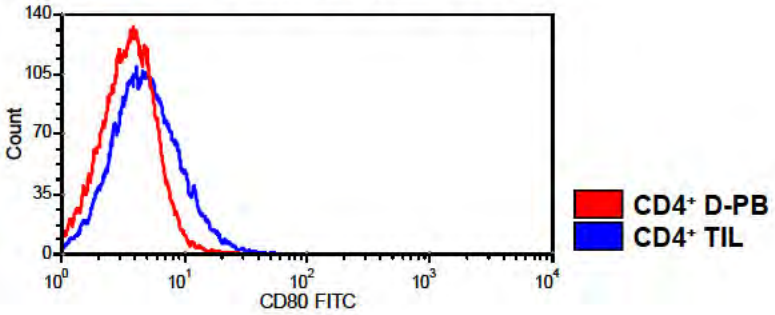
Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB



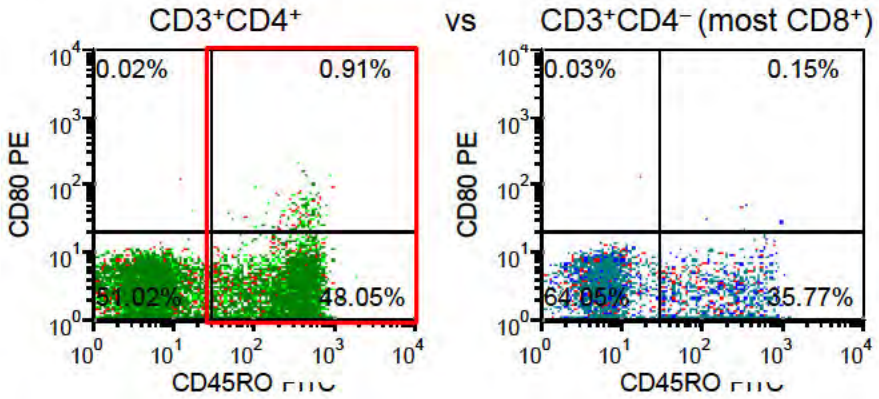
Flow Cytometric Analysis: CD80$^+$

Conclusion:
CD80 is slightly upregulated on the CD4$^+$ TIL.

CD3$^+$CD4$^+$ from Donor PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB

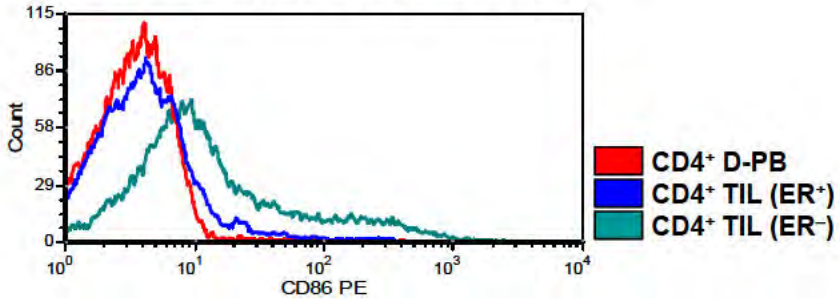


Array data from Table S2B:
Up (Fc = +5.05)

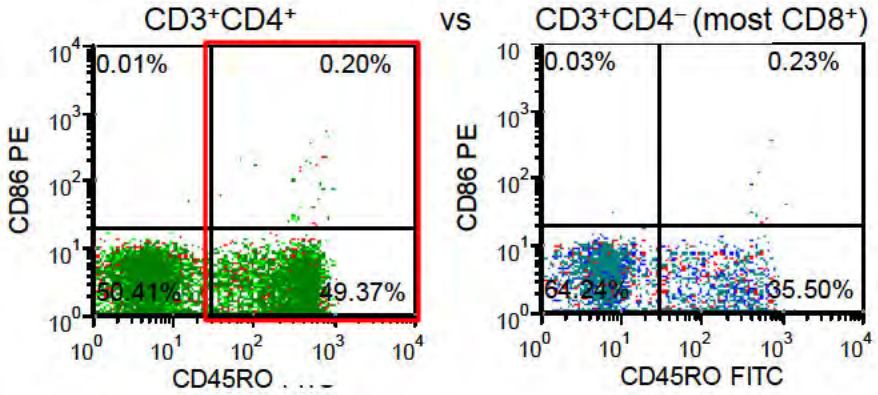
Flow Cytometric Analysis: CD86$^+$

Conclusion:
CD86 is upregulated on the CD4$^+$ TIL and to a significantly greater extent on TIL from ER$^-$ tumors.

CD3$^+$CD4$^+$ from Donor PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB



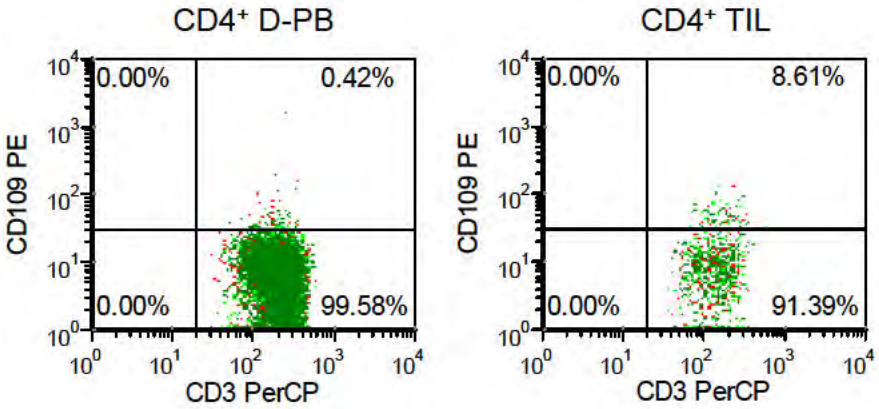
Array data from Table S2C:
Up (Fc = +4.43)

qRT-PCR confirmed the upregulation of *CD86* mRNA in 3 ER$^-$ compared to 18 ER$^+$ tumors; no correlation with the extent of lymphocyte infiltration was observed (Figure 8; Table S5D).

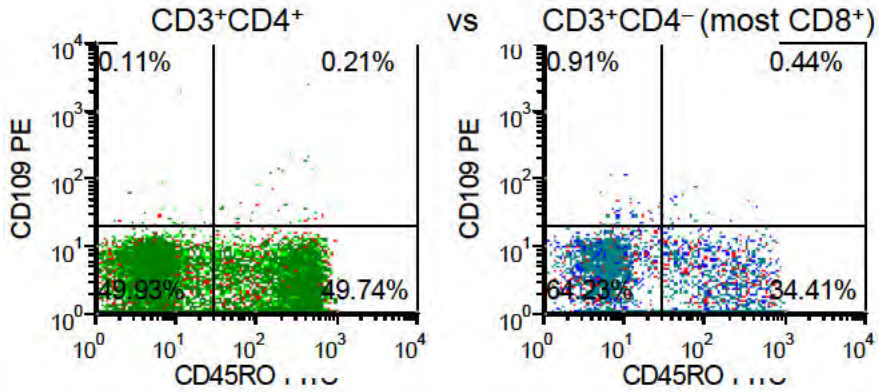
Flow Cytometric Analysis: CD109$^+$

Conclusion:
CD109 is slightly upregulated on the CD4$^+$ TIL.

CD3$^+$CD4$^+$ from Donor PB vs TIL



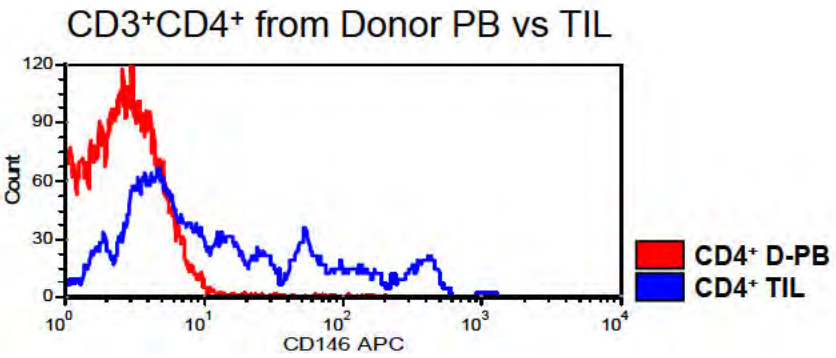
Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) from Donor PB



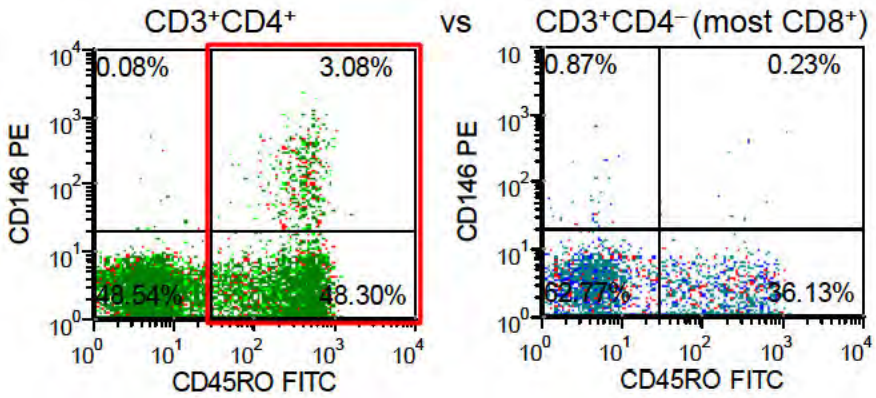
Array data from Table S2B:
Up (Fc = +6.52)

Flow Cytometric Analysis: CD146⁺ (*MCAM*)

Conclusion:
CD146 is upregulated on the CD4⁺ TIL.



Expression on CD45RO⁺ vs CD45RA⁺(=RO⁻) T cells from Donor PB

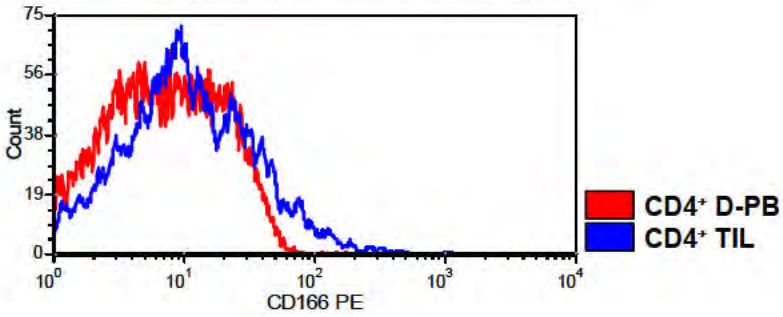


Array data from Table S2B:
Up (Fc = +6.60)

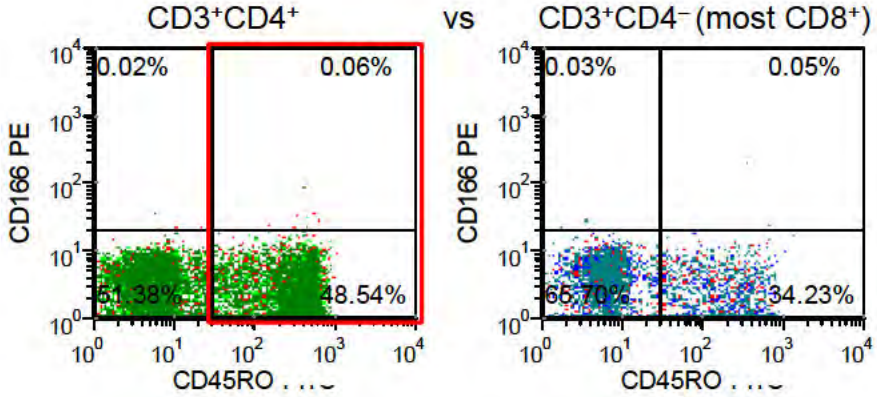
Flow Cytometric Analysis: CD166⁺ (*ALCAM*)

Conclusion:
CD166 is upregulated on the CD4⁺ TIL.

CD3⁺CD4⁺ from Donor PB vs TIL



Expression on CD45RO⁺ vs CD45RA⁺(=RO⁻) T cells from Donor PB

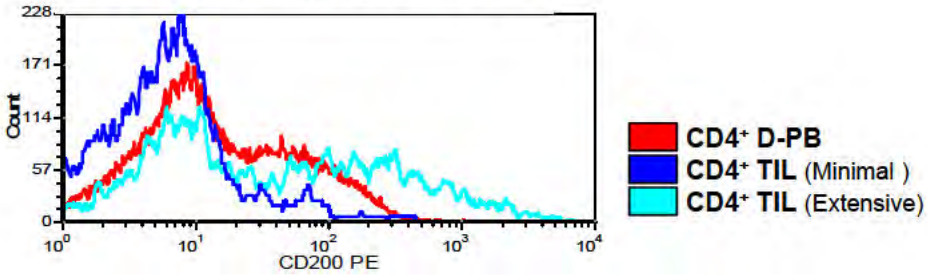


Array data from Table S2B:
Up (Fc = +7.22)

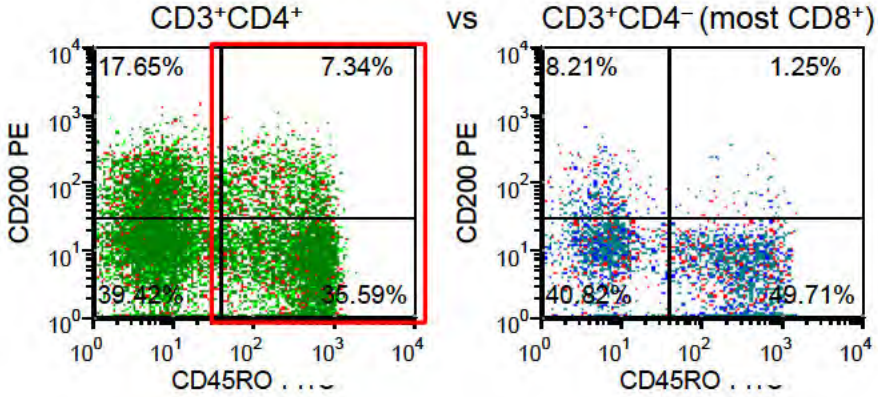
Flow Cytometric Analysis: CD200±

Conclusion:
CD200 expression is heterogeneous on CD4+ TIL with minimally-infiltrated tumors displaying decreased expression and extensively infiltrated tumors increased expression relative to blood from healthy controls.

CD3+CD4+ from Donor PB vs TIL



Expression on CD45RO+ vs CD45RA+(=RO−) T cells from Donor PB

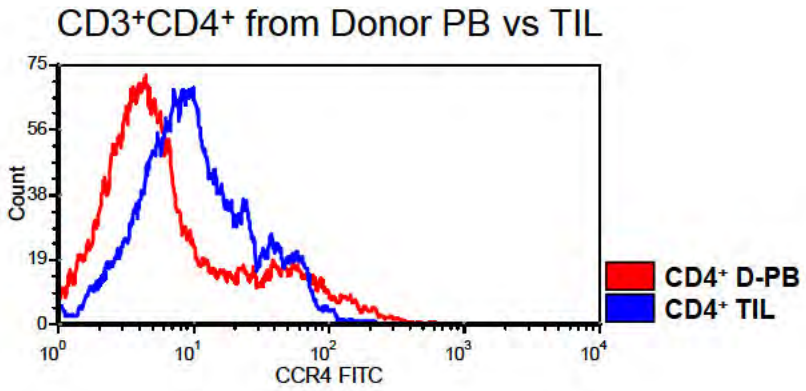


Array data from Table S2B:
Up (Fc = +11.24)

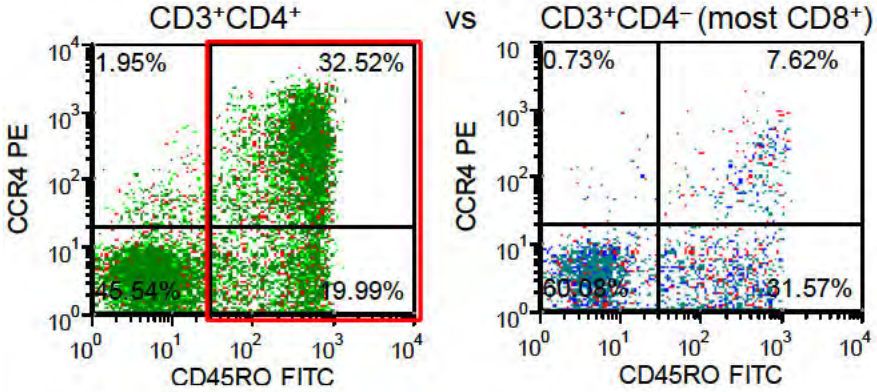
qRT-PCR confirmed *CD200* upregulation in extensively infiltrated relative to minimally infiltrated tumors (Figure 8; Table S5D).

Flow Cytometric Analysis: CCR4⁺ (CD194)

Conclusion:
CCR4 expression is upregulated on the majority of CD4⁺ TIL; however, the most positive cells display lower expression compared with their counterparts from healthy donor blood.



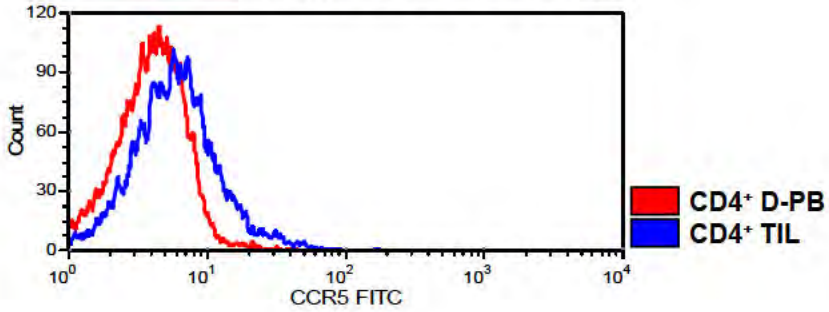
Expression on CD45RO⁺ vs CD45RA⁺(=RO⁻) T cells from Donor PB



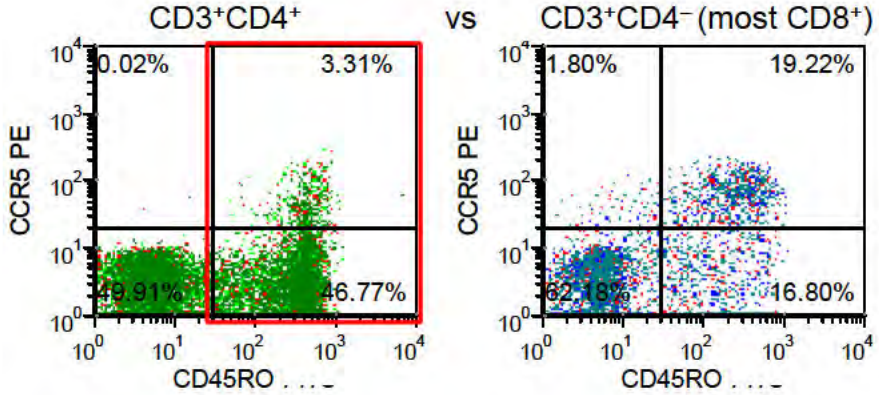
Flow Cytometric Analysis: CCR5⁺ (CD195)

Conclusion:
CCR5 expression is upregulated on the CD4⁺ TIL.

CD3⁺CD4⁺ from Donor PB vs TIL



Expression on CD45RO⁺ vs CD45RA⁺(=RO⁻) T cells from Donor PB

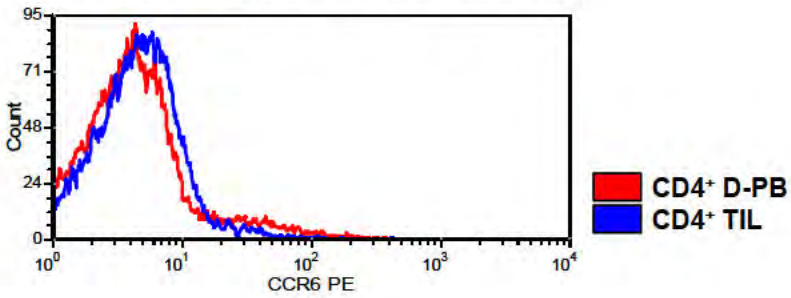


Array data from Table S2B:
Up (Fc = +3.79)

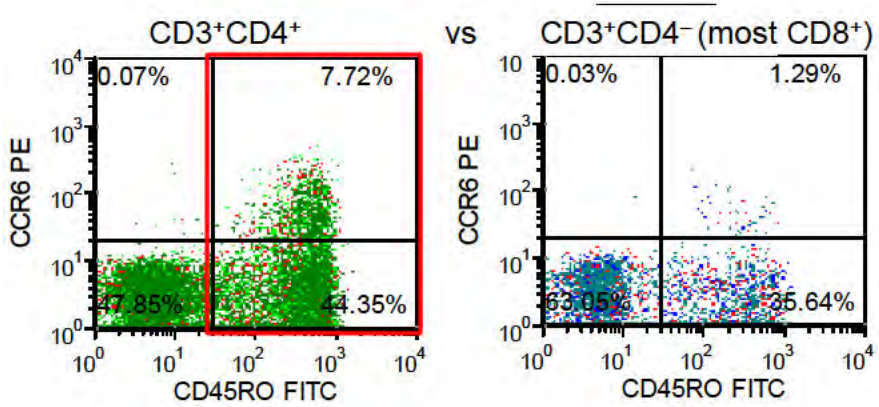
Flow Cytometric Analysis: CCR6^{lo} (CD196)

Conclusion:
The small subpopulation of CCR6 positive cells in the CD4$^+$ TIL have lower expression than their counterparts from healthy donors.

CD3$^+$CD4$^+$ from Donor PB vs TIL



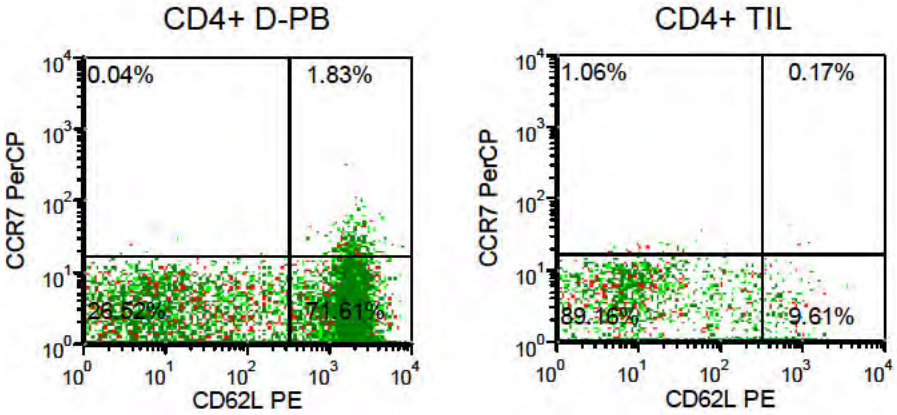
Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB



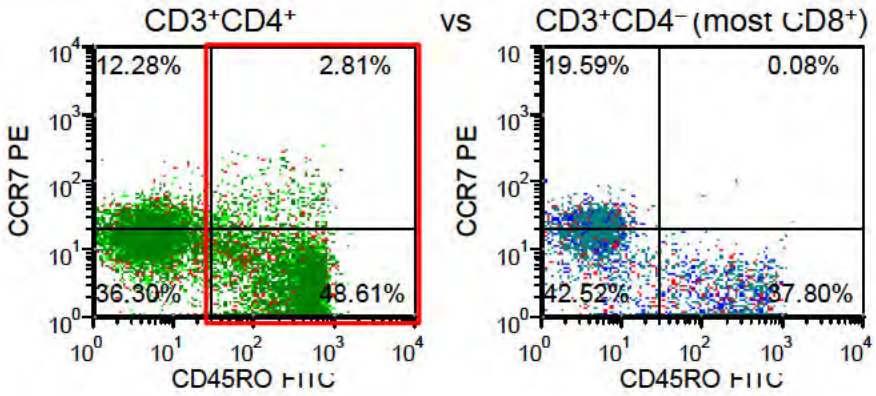
Flow Cytometric Analysis: CCR7^{lo} (CD197)

Conclusion:
Fewer CCR7 positive cells were detected in the CD4$^+$ TIL compared to CD4$^+$ T cells from healthy donor blood.

CD3$^+$CD4$^+$ from Donor PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB

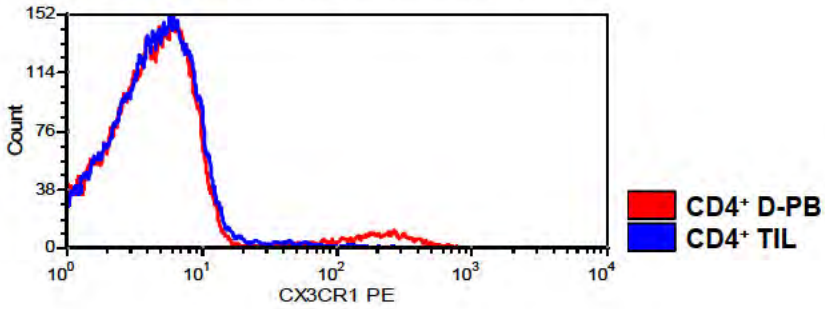


Array data from Table S2C:
Down (Fc = -2.53)

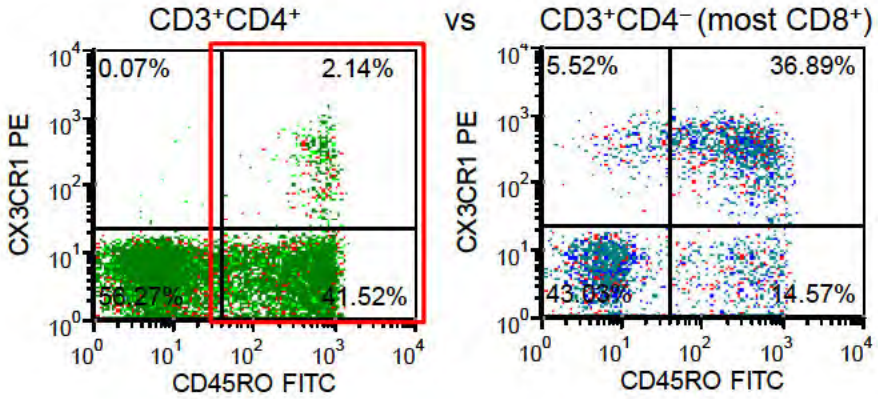
Flow Cytometric Analysis: CX3CR1^{lo}

Conclusion:
CX3CR1 expression is not expressed on the majority of CD4$^+$ TIL; however, the minor subpopulation of positive cells has reduced expression on the TIL compared with their counterparts from healthy donor blood.

CD3$^+$CD4$^+$ from Donor PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB

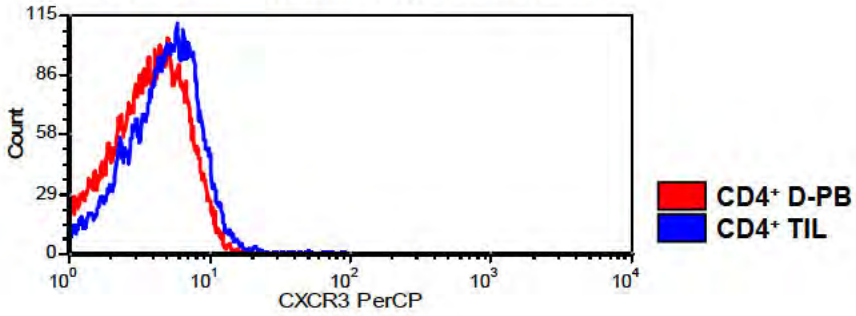


Array data from Table S2B:
Down (Fc = -4.13)

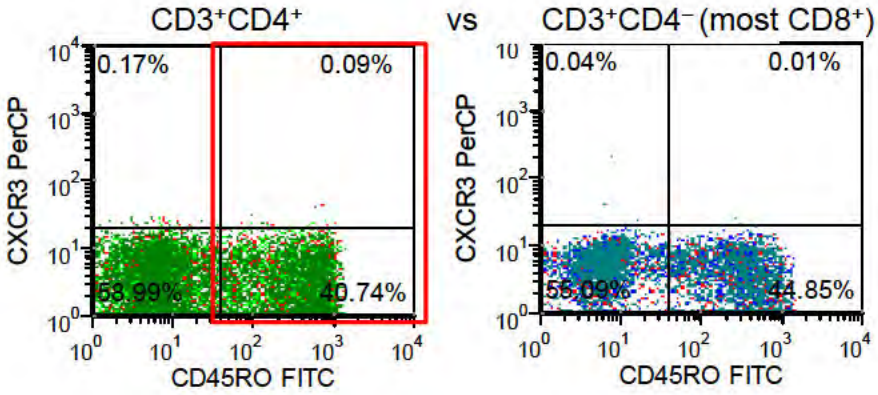
Flow Cytometric Analysis: CXCR3⁺ (CD183)

Conclusion:
CXCR3 expression is slightly upregulated on the CD4⁺ TIL.

CD3⁺CD4⁺ from Donor PB vs TIL



Expression on CD45RO⁺ vs CD45RA⁺(=RO⁻) T cells from Donor PB

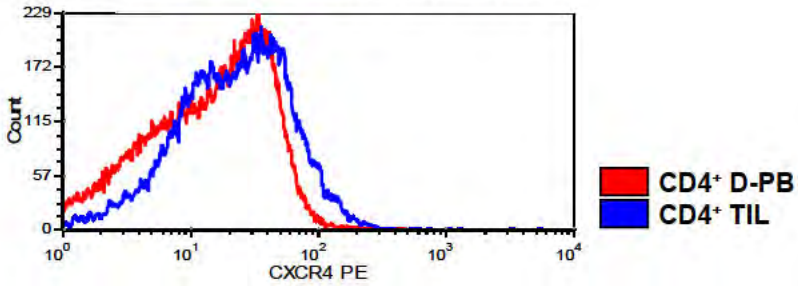


Array data from Table S2B:
Up (Fc = +2.80)

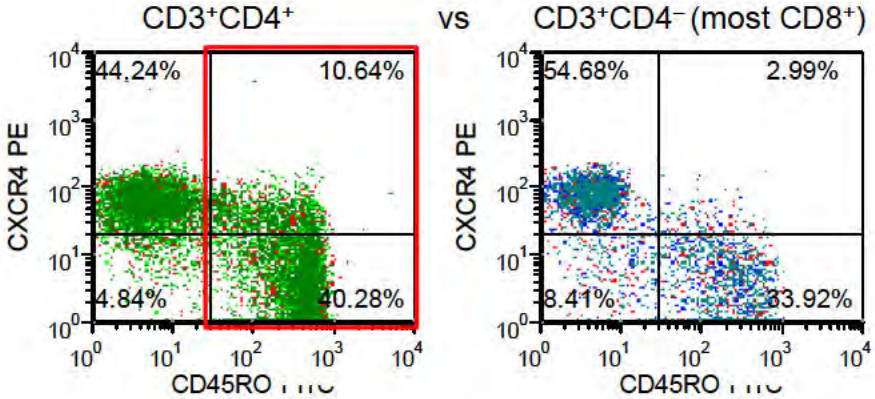
Flow Cytometric Analysis: CXCR4⁺ (CD184)

Conclusion:
CXCR4 expression is slightly upregulated on the CD4⁺ TIL.

CD3⁺CD4⁺ from Donor PB vs TIL

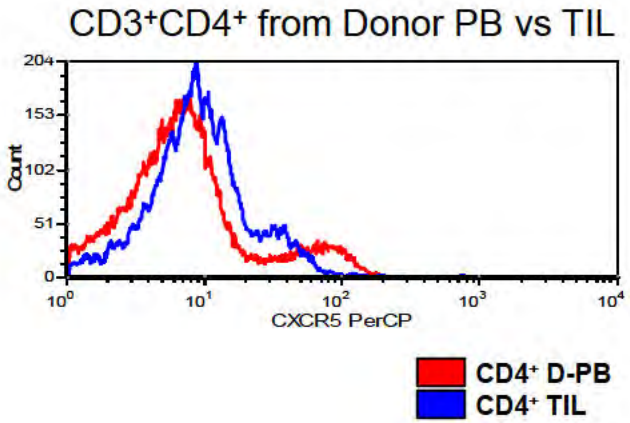


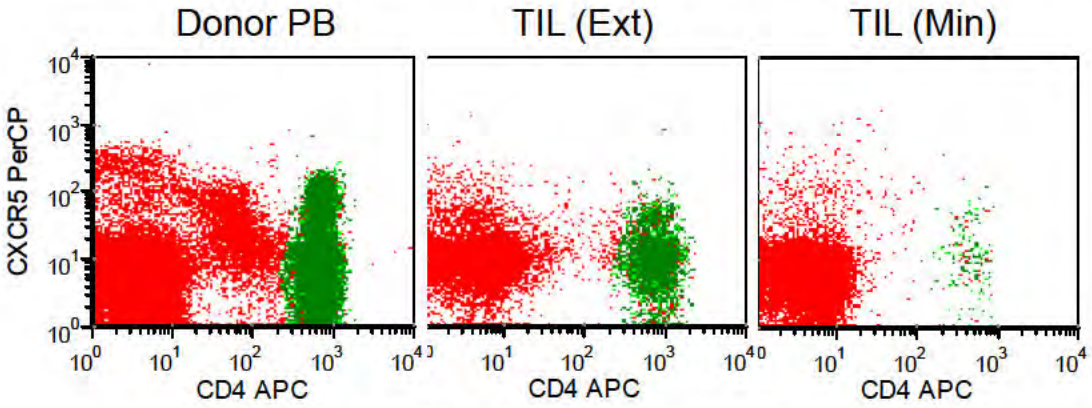
Expression on CD45RO⁺ vs CD45RA⁺(=RO⁻) T cells from Donor PB



Flow Cytometric Analysis: CXCR5⁺ (CD185)

Conclusion:
CXCR5 is slightly upregulated on the majority of CD4⁺ TIL; however, the most positive cells have lower expression than CD4⁺ T cells from healthy donors.

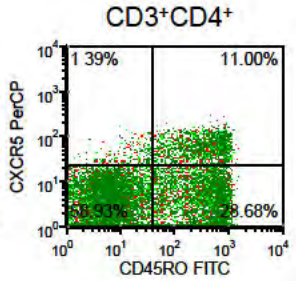




Gate 1 (page 3; green = Gate 1+ Gate 2)

Dot plots confirm that the CXCR5 antibody labeling is functional despite the low levels of expression observed on the TIL.

Expression on CD45RO⁺ vs CD45RA⁺(=RO⁻) $CD3^+CD4^+$ T cells from Donor PB



Possibly, the $CXCR5^{lo}$ cells reflect CXCL13-mediated surface downregulation on GC-resident TIL. The cells with increased surface CXCR5 might represent cells migrating to the T:B cell interface (Kerfoot *et al.*, 2011) within tertiary lymphoid structures in extensively infiltrated tumors.

Downregulation of surface CXCR5 by tumor SN treatment supports this hypothesis (Figure 6).

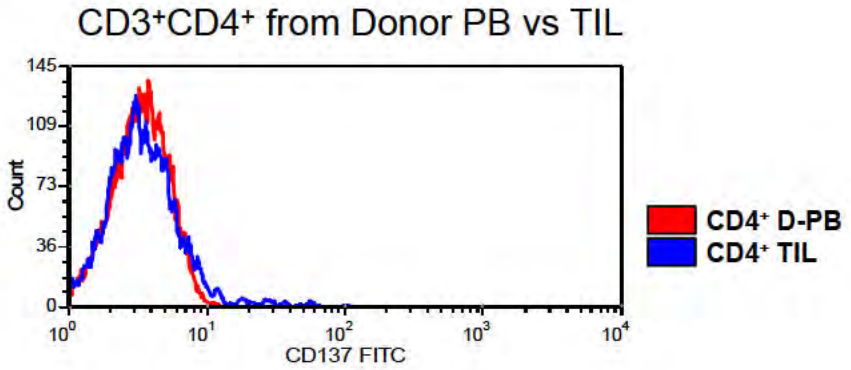
Note: CXCR5 has been shown to be differentially expressed on Tfh subpopulations in secondary lymphoid organs (Bentebibel *et al.*; Elsner *et al.*).

Kerfoot *et al.*, Immunity. (2011) Jun 24;34(6):947-60.
Bentebibel *et al.*, Proc Natl Acad Sci USA. (2011) 33:E488-97
Elsner *et al.* J Virol. (2012) Apr 24 (Epub ahead of print)

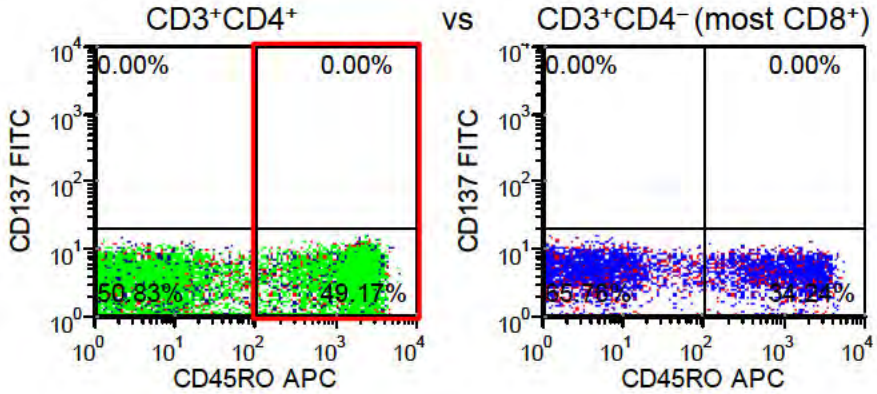
Flow Cytometric Analysis: 4-1BB$^+$ (*TNFRSF9*; CD137)

Conclusion:
A small subpopulation of CD4$^+$ TIL upregulate 4-1BB surface expression.

Note: This receptor is weakly expressed on unstimulated CD4$^+$ and CD8$^+$ TIL (Ju *et al.*) but can be induced on CD8$^+$ TIL stimulated *in vitro* (Hernandez-Chacon *et al.).*



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB



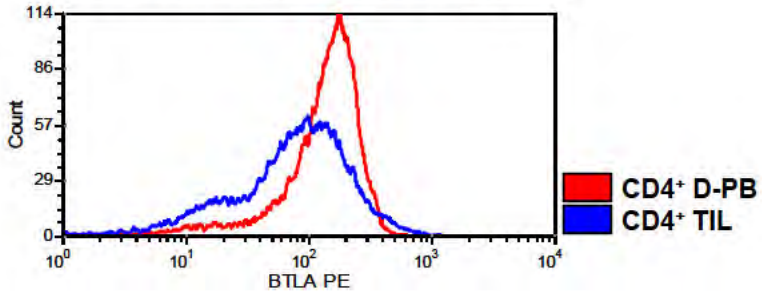
Array data from Table S2B:
Up (Fc = +9.56)

Ju *et al.*, Immunology and Cell Biology (2005) 83: 344–351.
Hernandez-Chacon *et al.*, J Immunother. (2011) 34: 236–250.

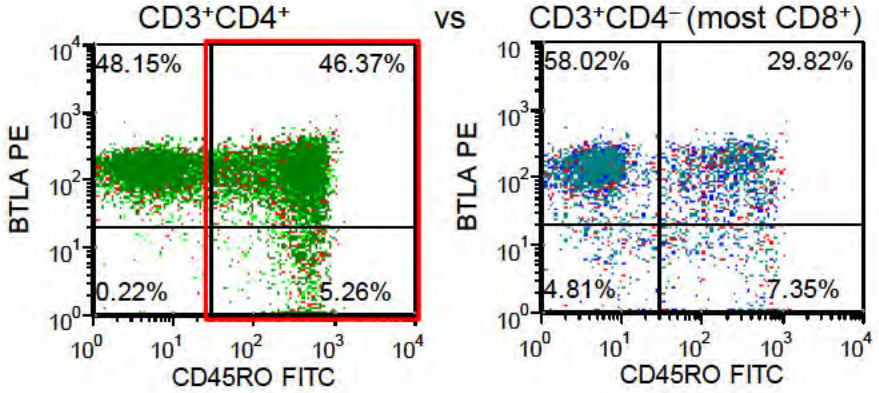
Flow Cytometric Analysis: BTLA± (CD272)

Conclusion:
BTLA expression is downregulated on the CD4⁺ TIL.

CD3⁺CD4⁺ from Donor PB vs TIL



Expression on CD45RO⁺ vs CD45RA⁺(=RO⁻) T cells from Donor PB

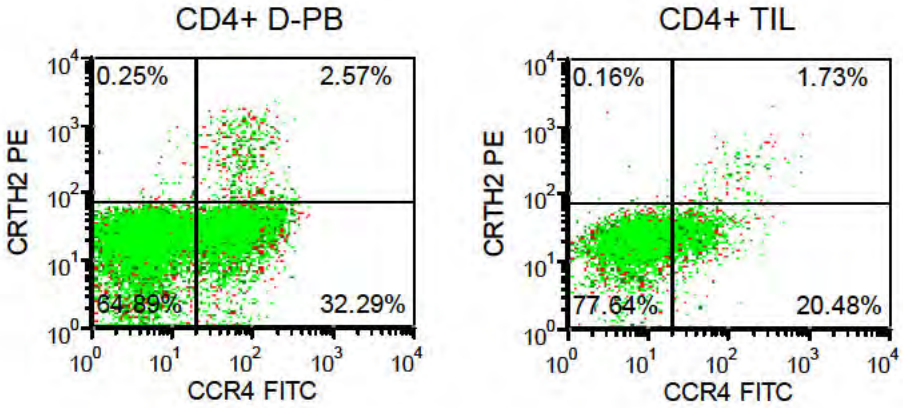


Array data from Table S2B:
Up (Fc = +2.51)

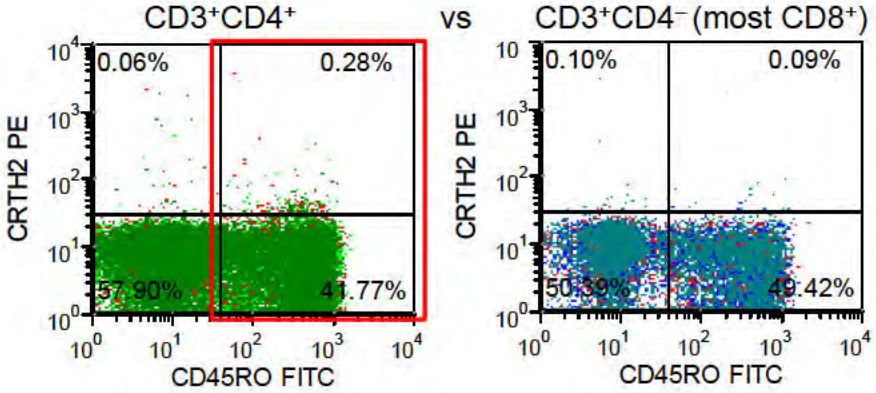
Flow Cytometric Analysis: CRTH2^{+} (*GPR44*; CD294)

Conclusion:
CRTH2 expression is not significantly different on the CD4^{+} TIL including a small subpopulation of (CCR4^{hi}) positive cells.

CD3^{+}CD4^{+} from Donor PB vs TIL



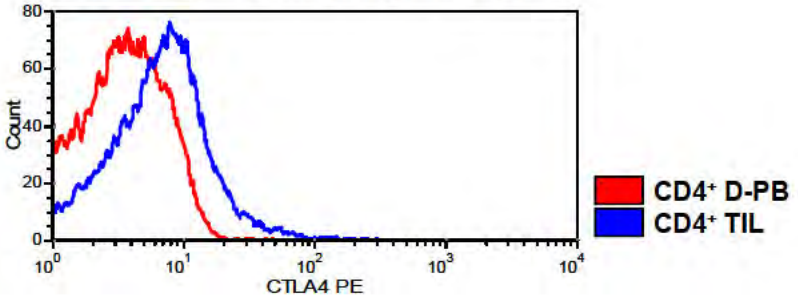
Expression on CD45RO^{+} vs CD45RA^{+}(=RO^{-}) T cells from Donor PB



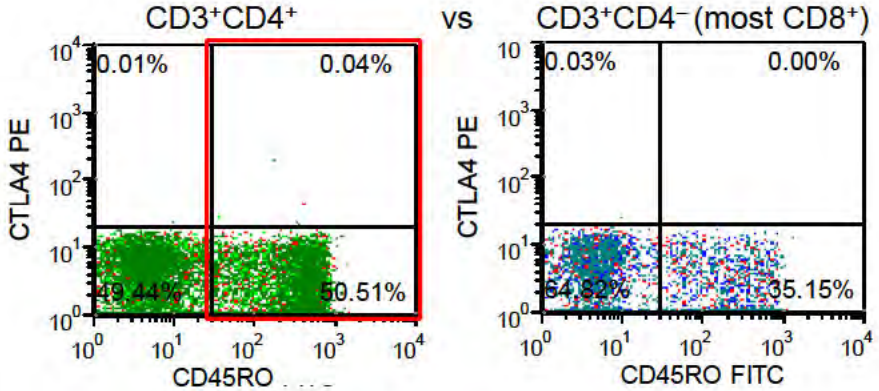
Flow Cytometric Analysis: CTLA4⁺ (CD152)

Conclusion:
CTLA4 is upregulated on the CD4⁺ TIL.

CD3⁺CD4⁺ from Donor PB vs TIL



Expression on CD45RO⁺ vs CD45RA⁺(=RO⁻) T cells from Donor PB

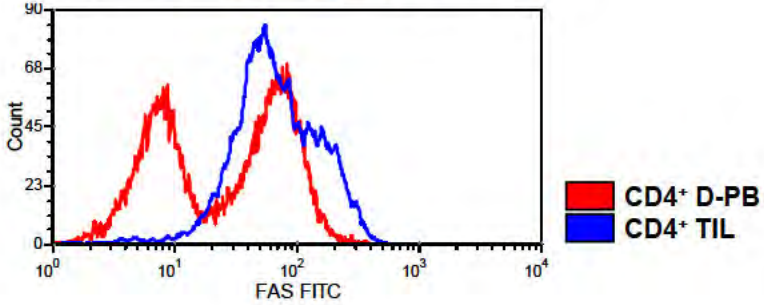


Array data from Table S2B:
Up (Fc = +4.87)

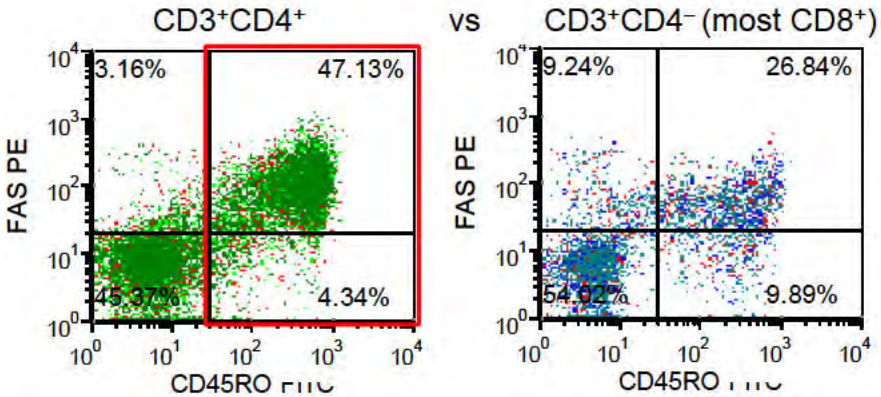
Flow Cytometric Analysis: FAS^{hi} (CD95)

Conclusion:
FAS is upregulated on the CD4$^+$ TIL;
FAS is also upregulated on CD45RO$^+$ T cells from healthy donors (below).

CD3$^+$CD4$^+$ from Donor PB vs TIL



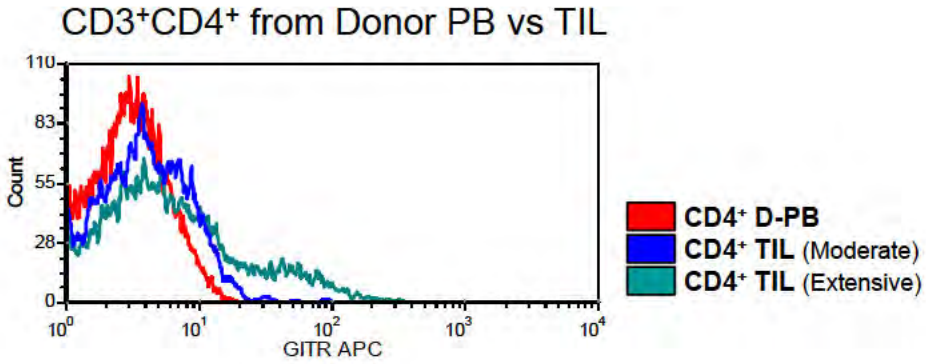
Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB



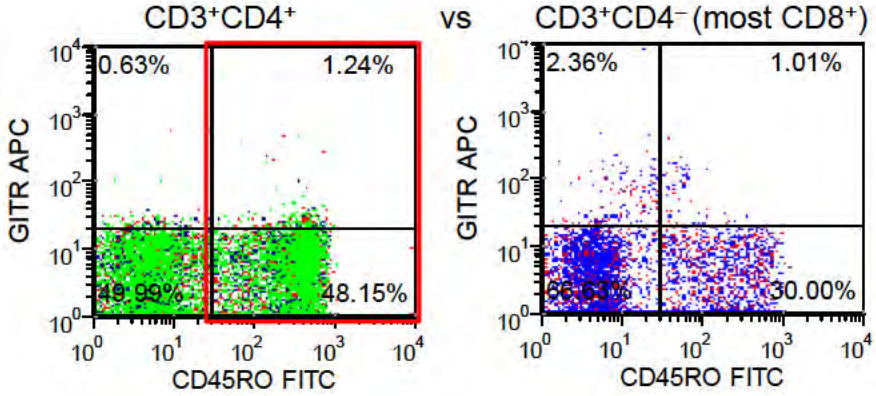
Array data from Table S2B:
Up (Fc = +2.07)

Flow Cytometric Analysis: GITR⁺ (*TNFRSF18*)

Conclusion:
GITR expression is upregulated on the CD4⁺ TIL, there is an increased frequency of higher expression on extensively infiltrated compared to moderately infiltrated tumors.



Expression on CD45RO⁺ vs CD45RA⁺(=RO⁻) T cells from Donor PB

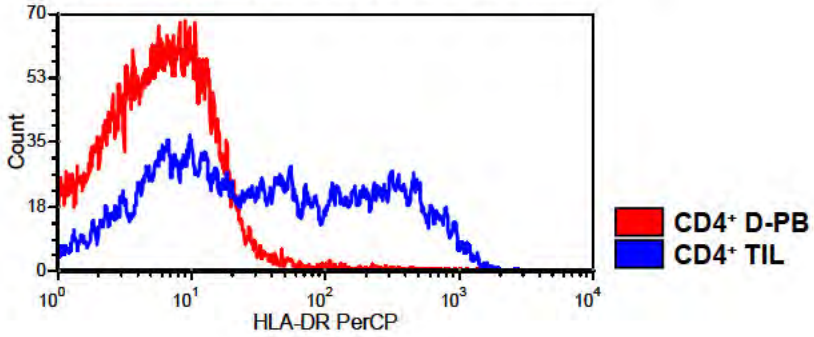


qRT-PCR confirmed a tendency to *TNFRSF18* upregulation associated with increased CD4+ TIL infiltration (Figure 8; Table S5D).

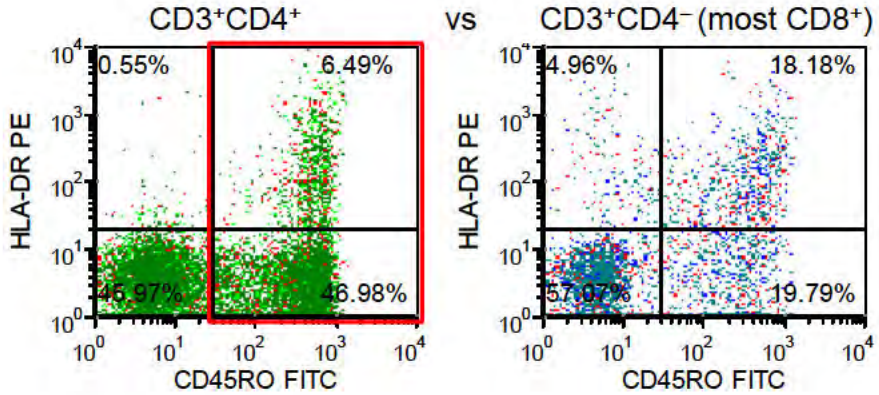
Flow Cytometric Analysis: HLA-DR$^+$

Conclusion:
HLA-DR expression is significantly upregulated on the CD4$^+$ TIL.

CD3$^+$CD4$^+$ from Donor PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB

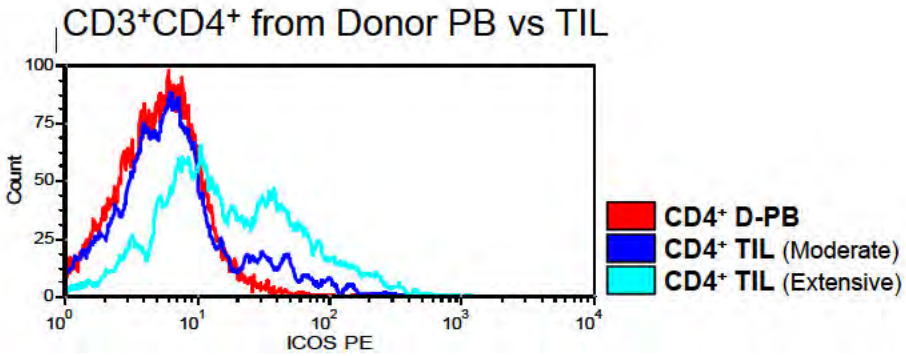


Array data from Table S2B:

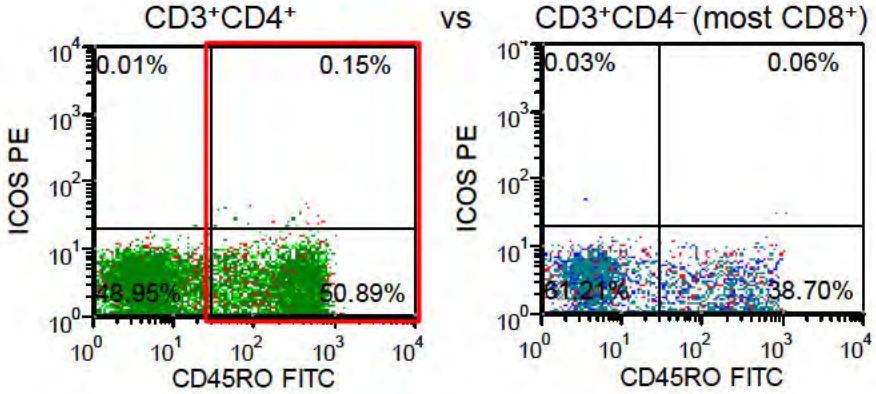
Up (Fc = *HLA-DRA* +3.77; *HLA-DRB1* +3.23; *HLA-DRB5* +2.85)

Flow Cytometric Analysis: ICOS⁺ (CD278)

Conclusion:
ICOS expression is heterogeneous, with the majority of CD4⁺ TIL showing increased expression in extensively infiltrated tumors while only a minor equivalent population is present in moderately infiltrated tumors.



Expression on CD45RO⁺ vs CD45RA⁺(=RO⁻) T cells from Donor PB



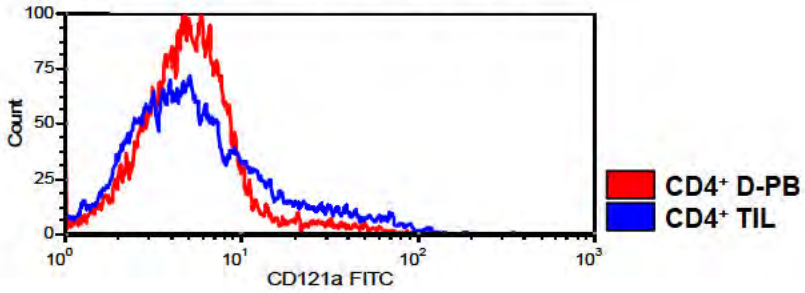
Array data from Table S2B:
Up (Fc = +2.83)

qRT-PCR confirmed *ICOS* upregulation in extensively infiltrated relative to minimally infiltrated tumors (Figure 8; Table S5D).

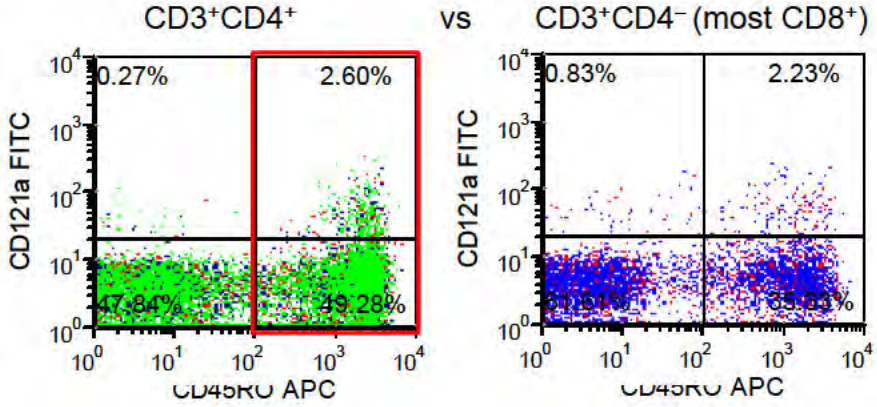
Flow Cytometric Analysis: IL-1R1^{+} (CD121α)

Conclusion:
IL1R1 is slightly upregulated on the CD4^{+} TIL.

CD3^{+}CD4^{+} from Donor PB vs TIL



Expression on CD45RO^{+} vs CD45RA^{+}(=RO^{-}) T cells from Donor PB

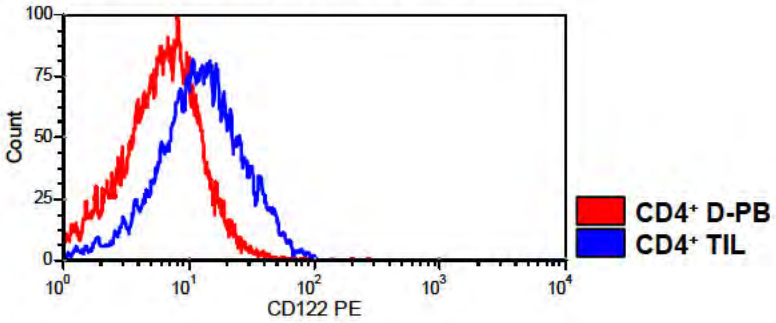


Array data from Table S2B:
Up (Fc = +9.59)

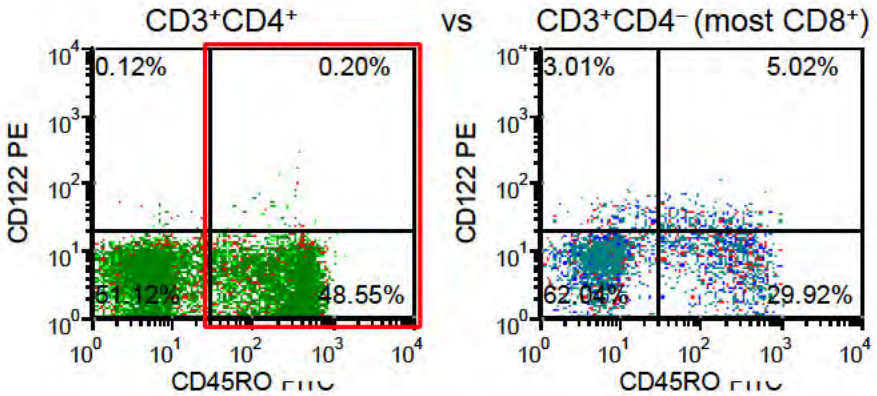
Flow Cytometric Analysis: IL-2RB$^+$ (CD122)

Conclusion:
IL2RB expression increases on the CD4$^+$ TIL.

CD3$^+$CD4$^+$ from Donor PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB

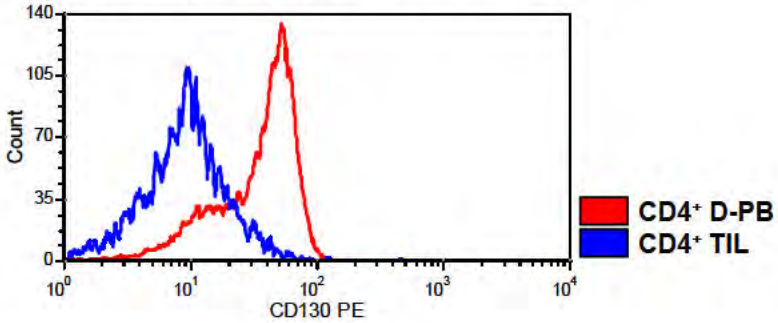


Array data from Table S2B:
Up (Fc = +2.50)

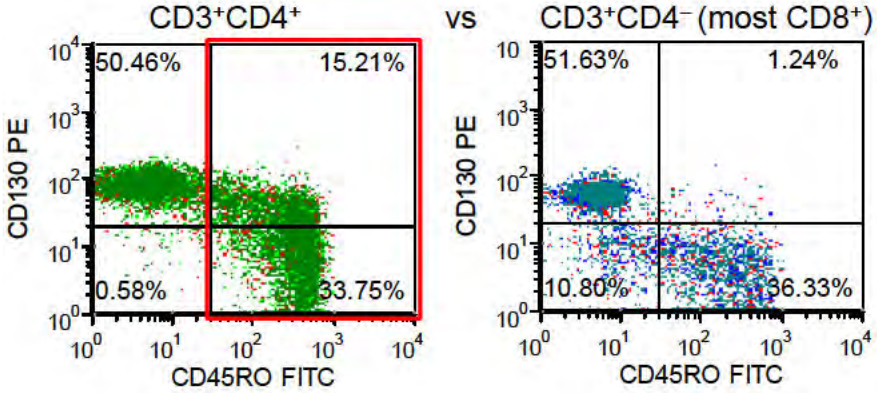
Flow Cytometric Analysis: IL-6ST^{lo} (CD130)

Conclusion:
IL6ST is downregulated on the CD4$^+$ TIL.

CD3$^+$CD4$^+$ from Donor PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB

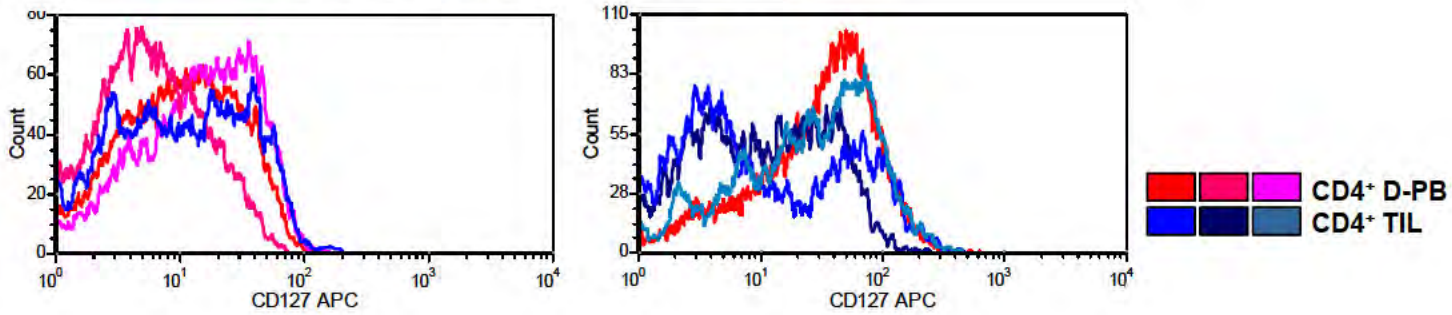


Array data from Table S2C:
Down (Fc = -1.62)

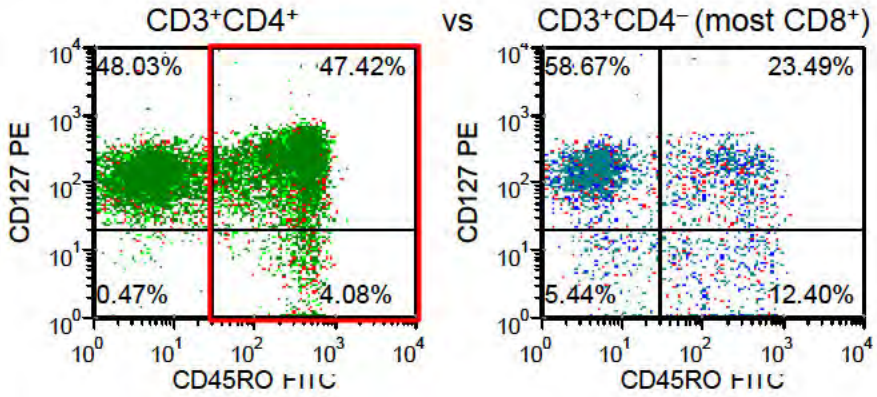
Flow Cytometric Analysis: IL-7R± (CD127)

Conclusion:
IL7R expression varies considerably on both CD4+ TIL and CD4+ T cells from donor PB.

CD3+CD4+ from Donor PB vs TIL



Expression on CD45RO+ vs CD45RA+(=RO−) T cells from Donor PB

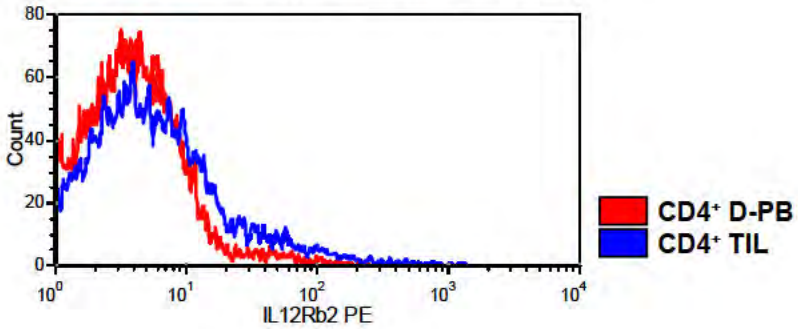


Array data from Table S2C:
Down (Fc = -1.59)

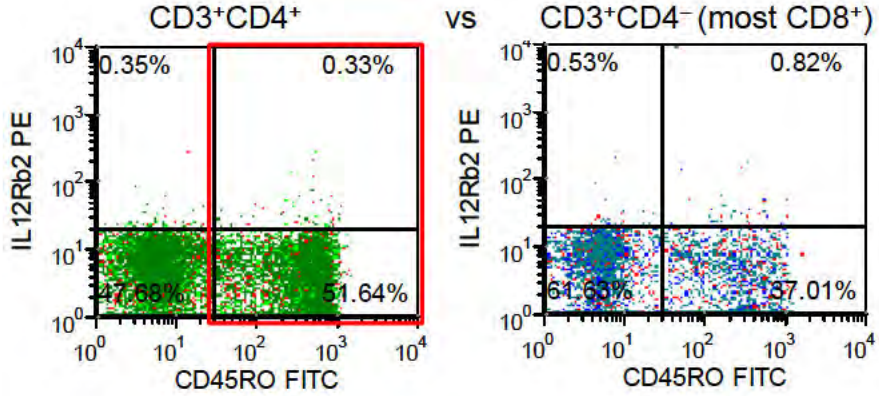
Flow Cytometric Analysis: IL-12Rβ2⁺

Conclusion:
IL12RB2 expression increases on the CD4⁺ TIL.

CD3⁺CD4⁺ from Donor PB vs TIL



Expression on CD45RO⁺ vs CD45RA⁺(=RO⁻) T cells from Donor PB

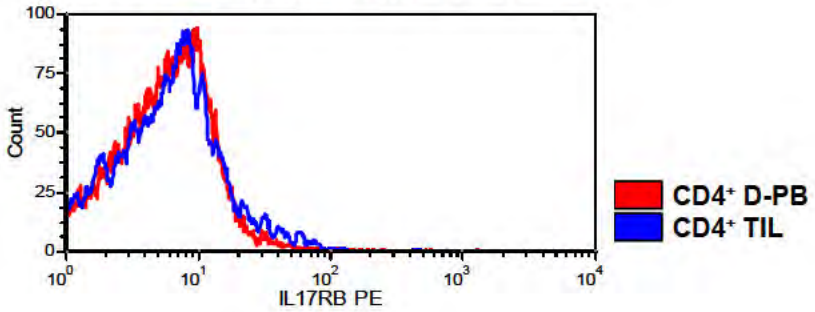


Array data from Table S2B:
Up (Fc = +5.84)

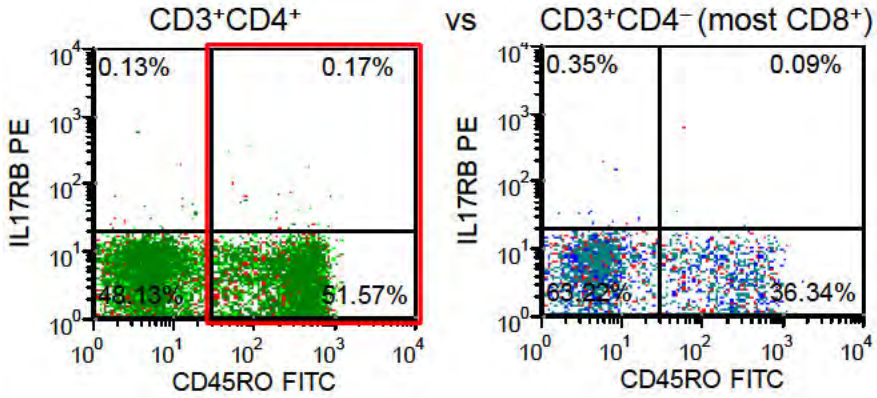
Flow Cytometric Analysis: IL-17RB$^+$

Conclusion:
IL17RB expression is essentially unchanged on the CD4$^+$ TIL.

CD3$^+$CD4$^+$ from Donor PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB

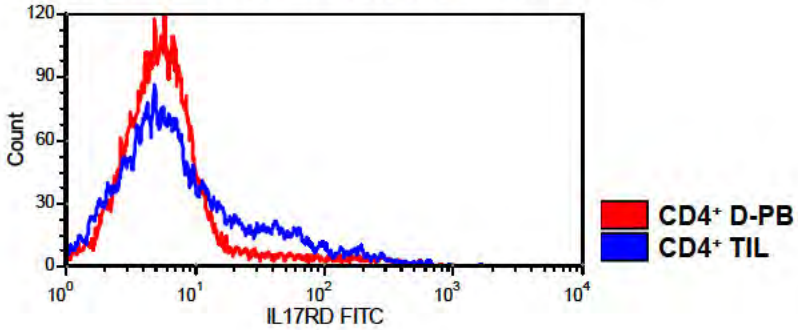


Array data from Table S2B:
Up (Fc = +2.61)

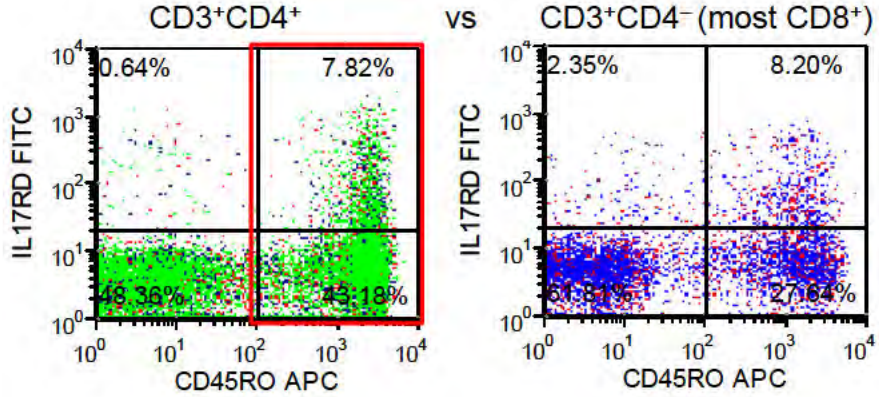
Flow Cytometric Analysis: IL-17RD$^+$

Conclusion:
The IL17RD$^+$ subset increases in the CD4$^+$ TIL.

CD3$^+$CD4$^+$ from Donor PB vs TIL



Expression on CD45RO$^+$ vs CD45RA$^+$(=RO$^-$) T cells from Donor PB

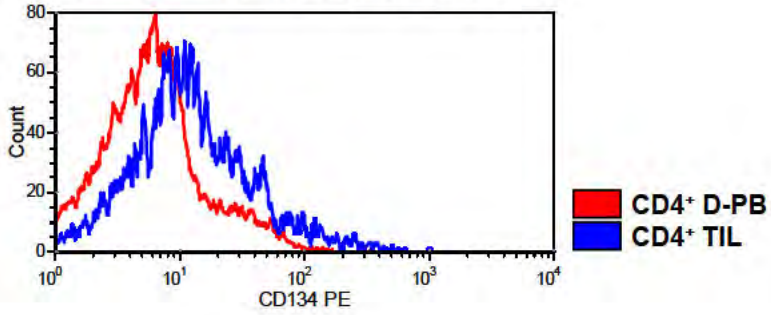


Array data from Table S2B:
Up (Fc = +7.49)

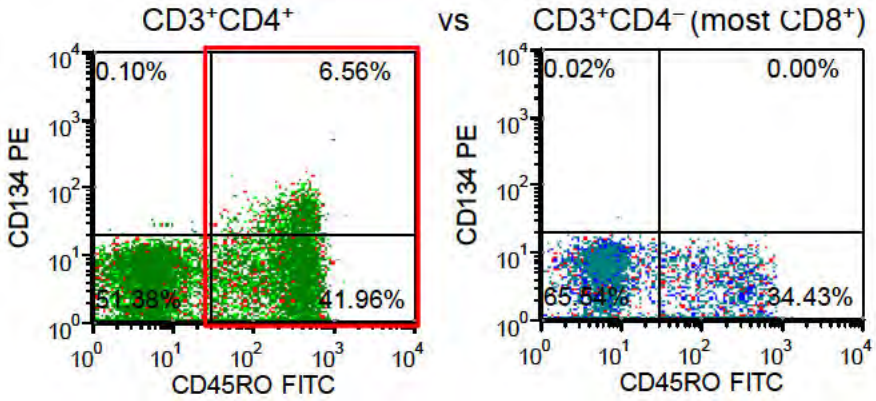
Flow Cytometric Analysis: OX40⁺ (*TNFRSF4*; CD134)

Conclusion:
OX40 is upregulated on the CD4⁺ TIL.

CD3⁺CD4⁺ from Donor PB vs TIL



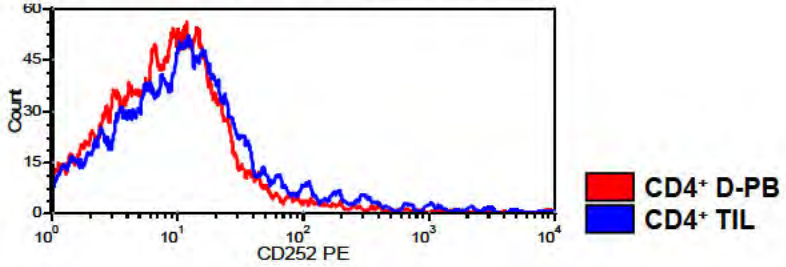
Expression on CD45RO⁺ vs CD45RA⁺(=RO⁻) T cells from Donor PB

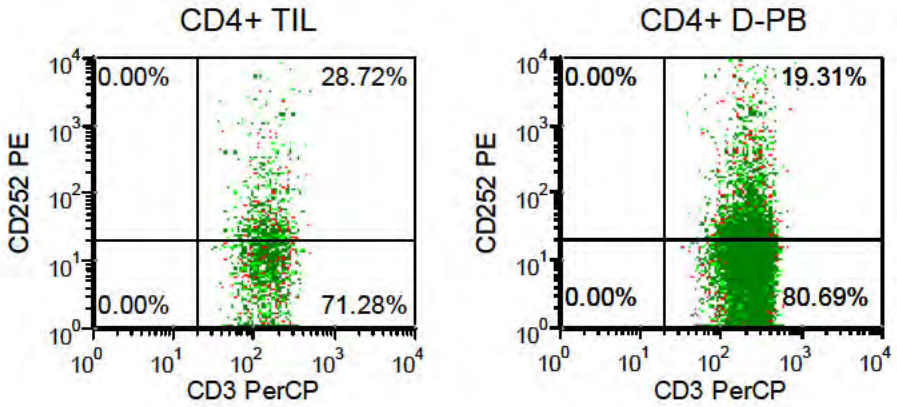


Flow Cytometric Analysis: OX40L$^+$ (*TNFSF4*; CD252)

Conclusion:
OX40L expression is slightly upregulated on the TIL compared to healthy donor blood.

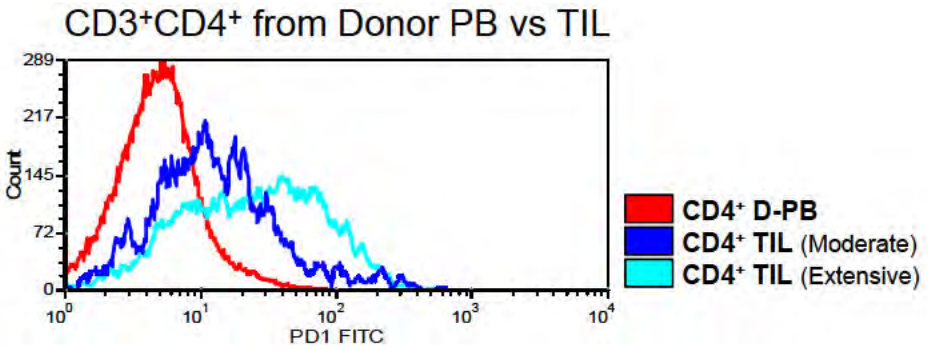
CD3$^+$CD4$^+$ from Donor PB vs TIL



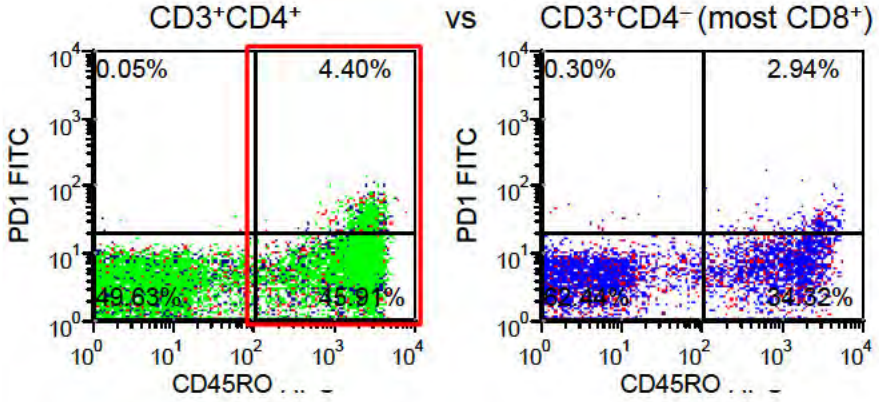


Flow Cytometric Analysis: PD-1⁺ (*PDCD1*; CD279)

Conclusion:
PD-1 expression is heterogeneous, with a fraction of CD4⁺ TIL showing increased expression. Higher levels are detected on TIL from extensively infiltrated compared to moderately infiltrated tumors.



Expression on CD45RO⁺ vs CD45RA⁺(=RO⁻) T cells from Donor PB

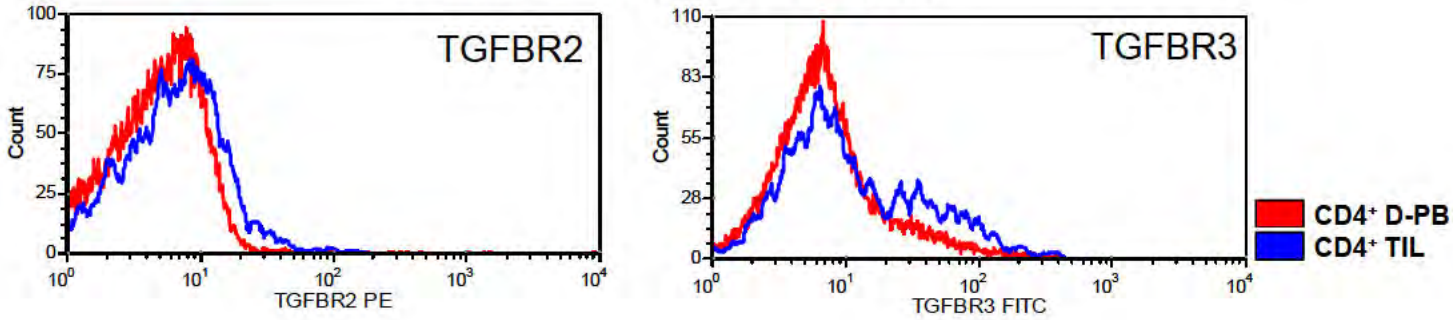


qRT-PCR confirmed *PDCD1* upregulation in extensively infiltrated relative to minimally infiltrated tumors (Figure 8; Table S5D).

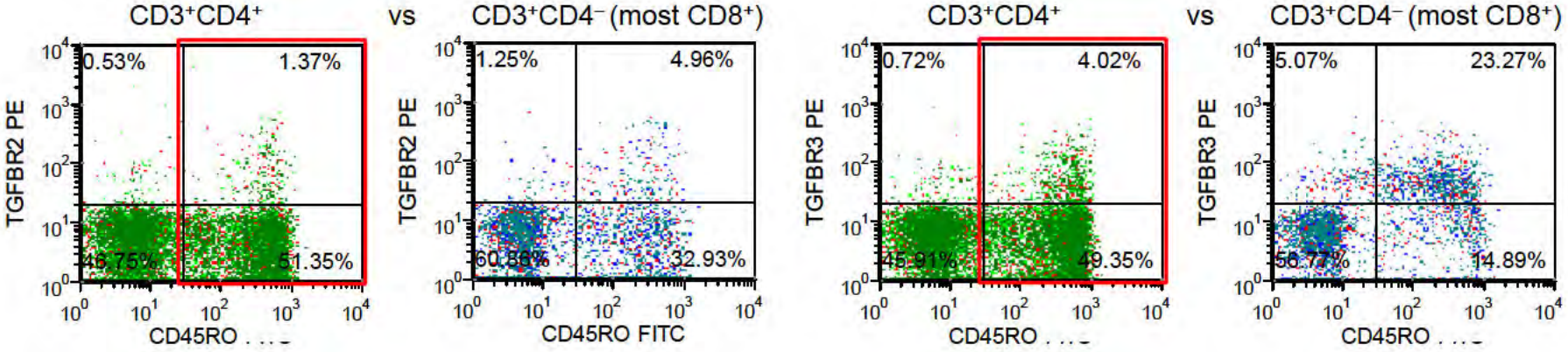
Flow Cytometric Analysis: TGFβ receptor

Conclusion:
TGFβ receptor expression is upregulated on a subpopulation of $CD4^+$ TIL.

$CD3^+CD4^+$ from Donor PB vs TIL



Expression on $CD45RO^+$ vs $CD45RA^+(=RO^-)$ T cells from Donor PB

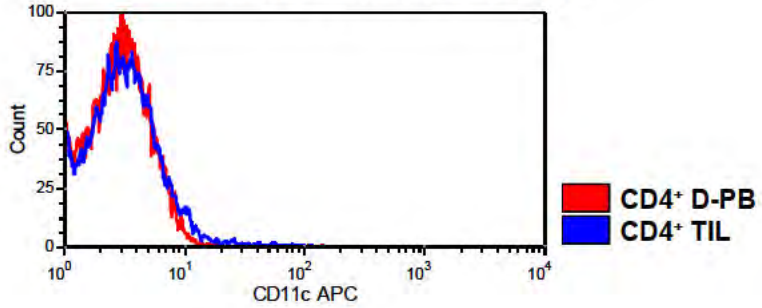


Array data from Table S2B:
Down (Fc = -2.88 *TGFBR2*)

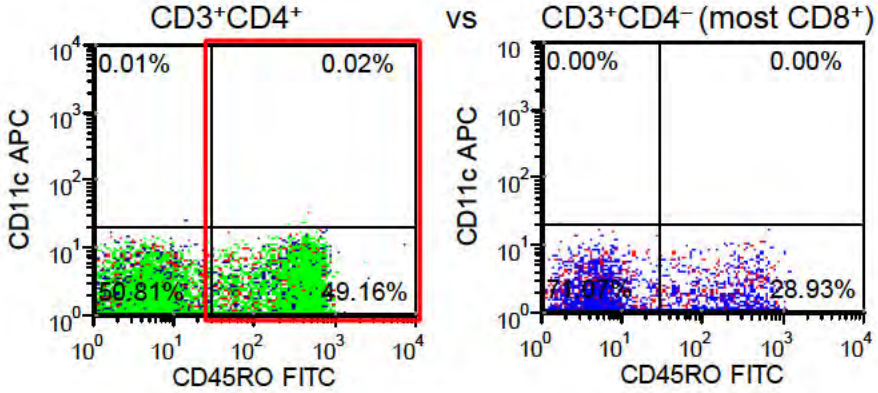
Flow Cytometric Analysis: CD11c⁻ (*ITGAX*)

Conclusion:
Used as a negative control to exclude the possibility that the differences detected were (in part) due to the presence of $CD4^{lo}$ monocytes / dendritic cells. CD11c expression was not detected on $CD4^+$ T cells from healthy donors and was not upregulated on $CD4^+$ TIL.

$CD3^+CD4^+$ from Donor PB vs TIL



Expression on $CD45RO^+$ vs $CD45RA^+(=RO^-)$ T cells from Donor PB

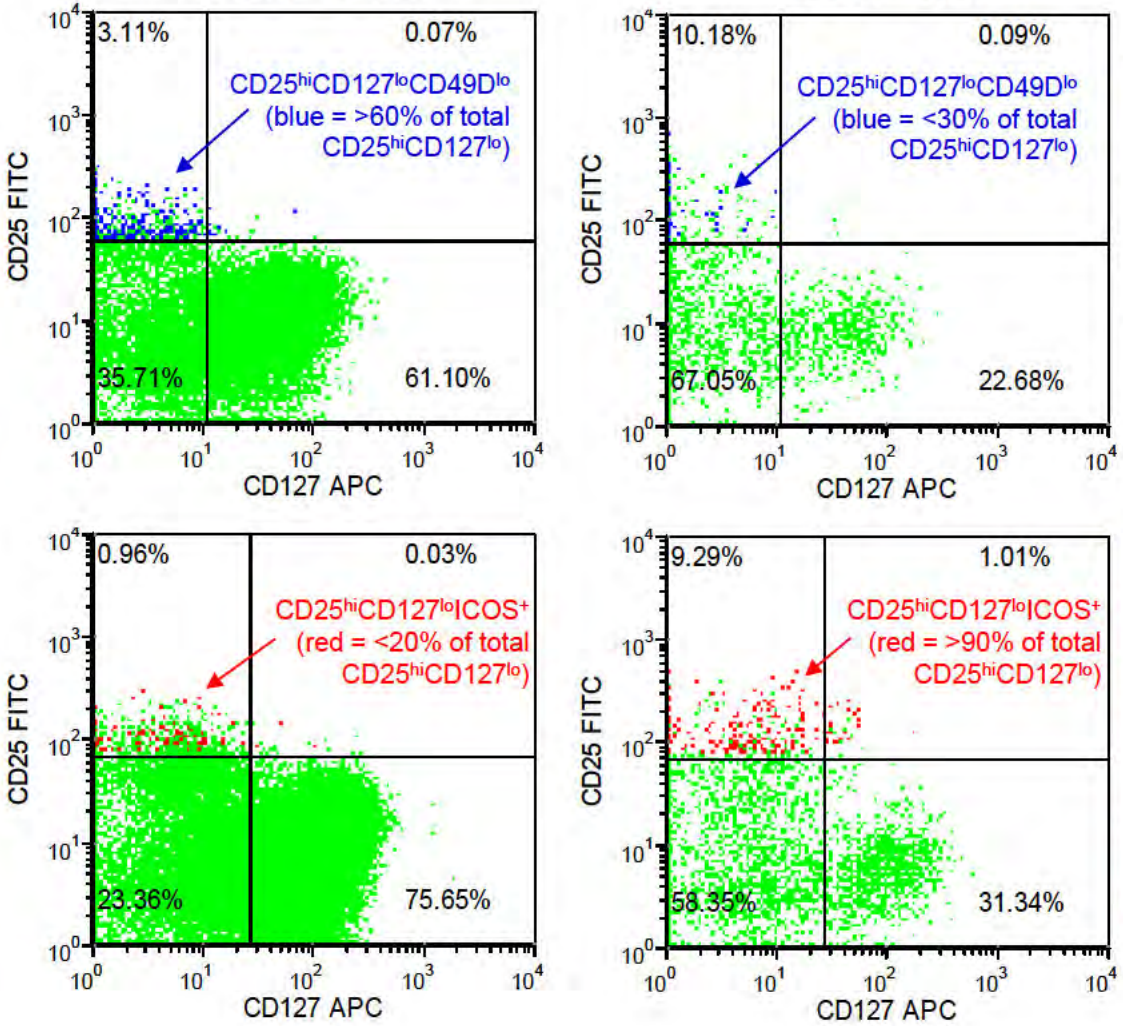


Flow Cytometric Analysis: Treg surface markers

CD3⁺CD4⁺ from Donor PB vs TIL

Treg
[Kleinewietfeld *et al.*, Blood. 2009 ;113(4):827-36]

activated "Treg"



Conclusion:

Conventional $CD25^{hi}CD127^{lo}$ Treg increase in the $CD4^+$ TIL compared to their donor PB counterparts.

However in the TIL, <30% of these cells are "true Treg" $CD25^{hi}CD127^{lo}CD49D^{lo}$ while >90% are activated T cells: $CD25^{hi}CD127^{lo}ICOS^+$.

The inverse is observed for $CD4^+$ T cells from donor PB.

Antibodies used in the flow cytometry experiments to examine T cell subpopulation markers in homogenates of human breast tumor fragments

Antibodies used to investigate surface molecule expression on $CD4^+$ TIL using flow cytometry

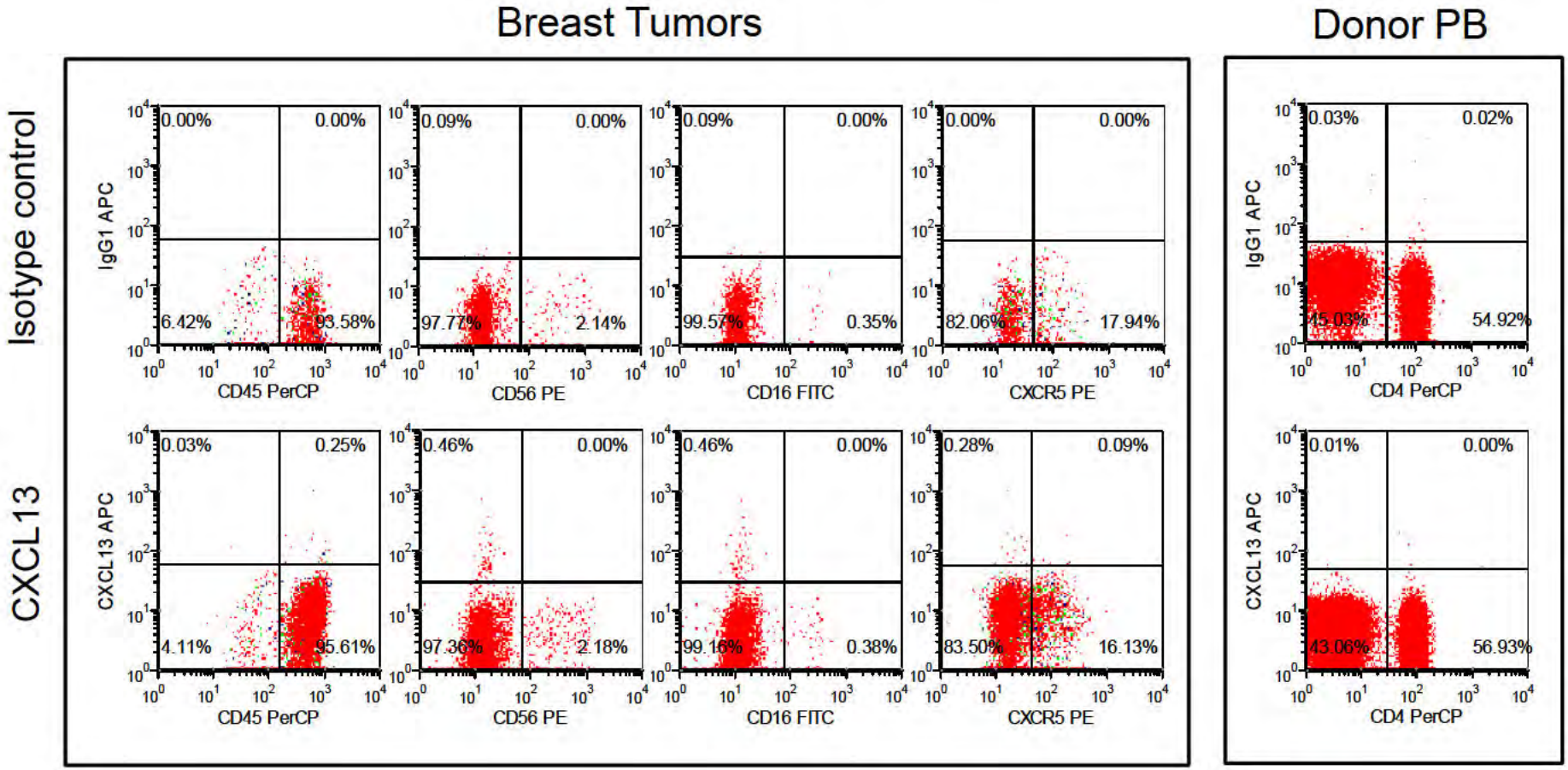
Antibody	Alternate name	Gene symbol	Fluorochrome	Source	Reference	Antibody	Alternate name	Gene symbol	Fluorochrome	Source	Reference
CD3		CD3E	PerCP	BD	345766	CD146	MUC18	MCAM	APC	Myltenyi Biotec	130-092-849
CD4		CD4	APC	BD Pharmingen	555349	CD166		ALCAM	PE	BD Pharmingen	559263
CD45RA		PTPRC	FITC	BD Pharmingen	555488	CD200	OX2	CD200	PE	BioLegend	329205
CD45RO		PTPRC	FITC	BD	555492	CCR4	CD194	CCR4	FITC	R&D	FAB1567F
CD45RO		PTPRC	APC	BD	340438	CCR4	CD194	CCR4	PE	R&D	FAB1567P
TCRαβ		TRA@	FITC	BD Pharmingen	555547	CCR5	CD195	CCR5	FITC	R&D	FAB182F
CD6		CD6	PE	BD Pharmingen	555358	CCR5	CD195	CCR5	PE	BD Pharmingen	555993
CD7		CD7	FITC	BD	347483	CCR6	CD196	CCR6	PE	eBioscience	12-1969
CD7		CD7	PE	BD	332774	CCR7	CD197	CCR7	PE	BD Pharmingen	552176
CD11b		ITGAM	PE	BD Pharmingen	555388	CCR7	CD197	CCR7	PerCP	R&D	FAB197C
CD18	LFA-1β	ITGB2	FITC	eBioscience	11-0187	CX3CR1		CX3CR1	PE	BioLegend	341603
CD24		CD24	FITC	BD Pharmingen	555427	CXCR3	CD183		PerCP	R&D	FAB160C
CD24		CD24	PE	BD Pharmingen	555428	CXCR4	CD184	CXCR4	PE	BD Pharmingen	555974
CD25	IL-2Rα	IL2RA	FITC	BD	345796	CXCR5	CD185	BLR1	PerCP-Cy5.5	BioLegend	335001
CD25	IL-2Rα	IL2RA	PE	BD	341011	4-1BB	CD137	TNFRSF9	FITC	eBioscience	11-1379
CD26		DPP4	FITC	BD	340426	BTLA	CD272	BTLA	PE	BD Pharmingen	558485
CD27		TNFRSF7	FITC	BD Pharmingen	555440	CRTH2	CD294	GPR44	PE	Myltenyi Biotec	130-091-238
CD27		TNFRSF7	PE	BD Pharmingen	555441	CTLA-4	CD152	CTLA4	PE	BD Pharmingen	555853
CD28		CD28	PE	eBioscience	12-0289	FAS	CD95	FAS	FITC	R&D	FAB142F
CD35		CR1	PE	BD Pharmingen	559872	FAS	CD95	FAS	PE	BD	340480
CD38		CD38	PE	BD Pharmingen	555460	GITR	AITR	TNFRSF18	APC	eBioscience	17-5875
CD43		SPN	FITC	BD Pharmingen	560198	HLA-DR		HLA-DRA/B1-6	PE	BD Pharmingen	555812
CD44		CD44	PE	BD Pharmingen	555479	HLA-DR		HLA-DRA/B1-6	PerCP	BD	347402
CD44		CD44	APC	BD Pharmingen	559942	ICOS	CD278	ICOS	PE	BD Pharmingen	557802
CD49D		ITGA4	PE	BD	340296	IL-1R1	CD121α	IL1R1	FITC	R&D	FAB269F
CD54		ICAM1	PE	eBioscience	12-0549	IL-2Rβ	CD122	IL2RB	PE	BD	340254
CD55	DAF	CD55	PE	iQP Prosan	IQP-520R	IL-6ST	CD130	IL6ST	PE	R&D	FAB228P
CD57		B3GAT1	FITC	BD	333169	IL-7R	CD127	IL7R	PE	BD Pharmingen	557938
CD58		CD58	PE	BD	340295	IL-7R	CD127	IL7R	Alexa Fluor 647	BD Pharmingen	558598
CD59	MIRL	CD59	PE	iQP Prosan	IQP-521R	IL-12Rβ2		IL12RB2	PE	R&D	FAB1959P
CD62L		SELL	PE	BD	341012	IL-17RB		IL17RB	PE	R&D	FAB1207P
CD69		CD69	PE	BD	341652	IL-17RD		IL17RD	FITC	R&D	FAB2275F
CD71		TFRC	PE	BD Pharmingen	555537	OX-40	CD134	TNFRSF4	PE	BD Pharmingen	555838
CD73		NT5E	PE	BD Pharmingen	550257	OX-40L	CD252	TNFSF4	PE	BD Pharmingen	558164
CD80	B7-1	CD80	FITC	BD Pharmingen	555683	PD1	CD279	PDCD1	FITC	eBioscience	11-9969
CD80	B7-1	CD80	PE	BD	340294	TGFβR2		TGFBR2	PE	R&D	FAB241P
CD86		CD86	PE	BD Pharmingen	555658	TGFβR3		TGFBR3	FITC	R&D	FAB242F
CD109	CPAMD7	CD109	PE	eBioscience	12-1099	TGFβR3		TGFBR3	PE	R&D	FAB242P
CD146	MUC18	MCAM	PE	BD Pharmingen	550315	CD11c		ITGAX	APC	BD	333144

Figure S3:

Flow cytometric analysis of cellular subpopulations expressing CXCL13 in human breast tumors

(Patients from the Confirmation Set; Table S1C)

Figure S3A:
Lymphocytes Infiltrating Breast Carcinomas
(Lymphocyte Gate)



Within the infiltrating lymphocyte population, CXCL13 is not produced by CD56⁺ or CD16⁺ NK cells. CXCL13 producer cells are CXCR5⁻ lymphocytes. We also found that CXCR5⁺CD4⁺ TIL are PD1^{lo}CD200^{lo} (data not shown).

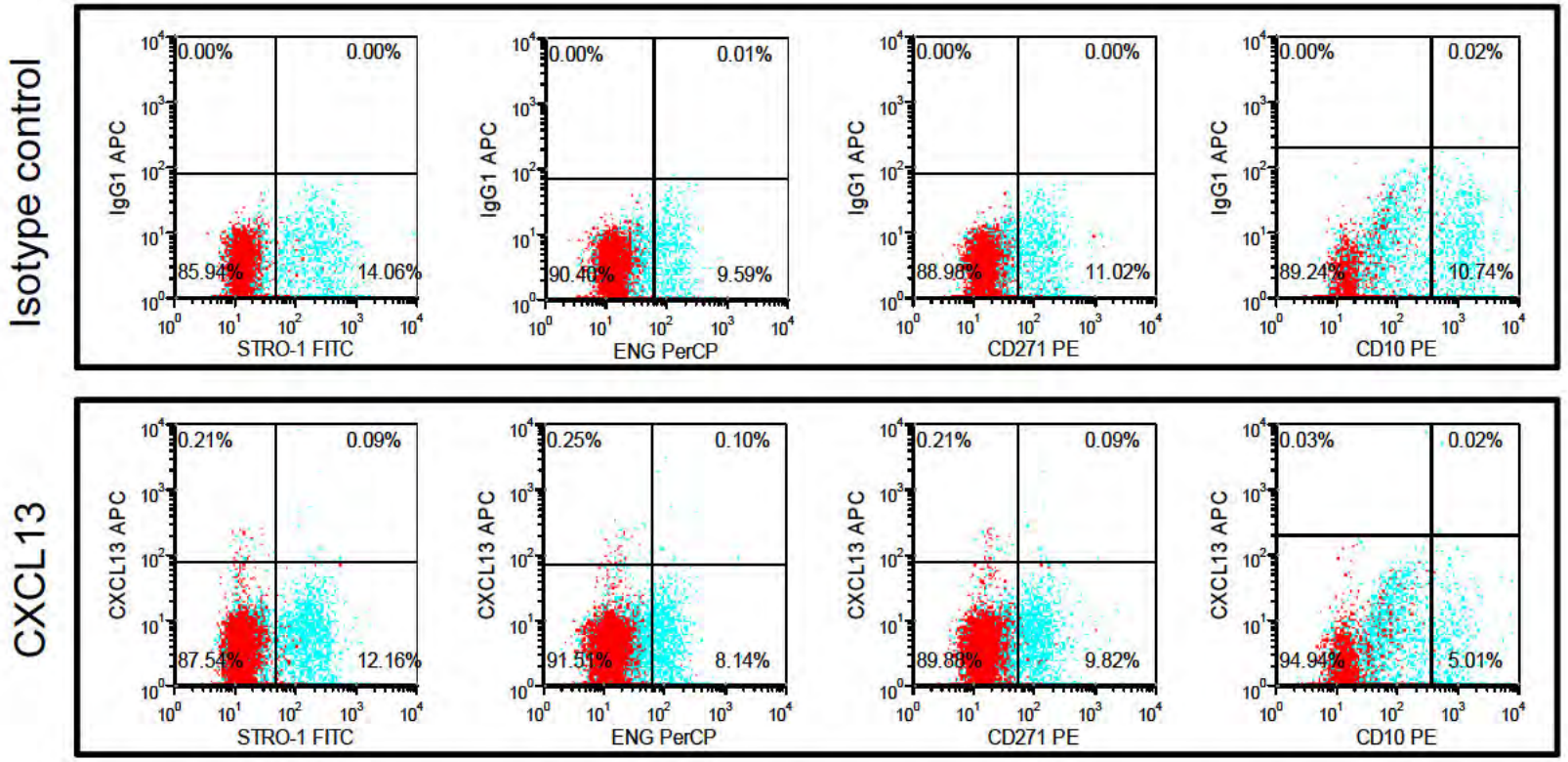
D-PB lymphocytes do not produce CXCL13

Note: Although commonly accepted as a Tfh cell surface marker, CXCR5 was recently shown to be differentially expressed on Tfh subpopulations in secondary lymphoid organs [Bentebibel *et al.*, Proc Natl Acad Sci USA. (2011) 33:E488-97; Elsner *et al.* J Virol. (2012) Apr 24 (Epub ahead of print)].

Figure S3B:

Total Tumor Homogenates

Viable cell gate (Lymphocyte gate in red)



Bone marrow stromal cells

Endothelial cells

Epithelial cells
Mesenchymal stem cells
Lymphocytes
Follicular DCs

Myoepithelial cells
B cell progenitors
Germinal center B cells
Granulocytes
Fibroblasts

In breast tumors, low levels of CXCL13 is also produced by stromal cells, endothelial cells and/or FDCs.

Figure S3C:

Antibodies used ın the flow cytometry experiments to examine cell surface markers and intracellular CXCL13 protein expression in homogenates of human breast tumor fragments

Antibody	Alternate name	Gene symbol	Fluorochrome	Source	Reference
CD3		CD3E	PerCP	BD	347344
CD4		CD4	PE	BD	345769
CD4		CD4	PerCP	BD	345770
CD8		CD8A/B	FITC	Beckman Coulter	A07756
CD10		MME	PE	BD Pharmingen	555375
CD14		CD14	FITC	BD Pharmingen	555397
CD16		FCGR3A	FITC	Beckman Coulter	IM0814
CD19		CD19	PE	BD	345777
CD38		CD38	PE	BD Pharmingen	555460
CD45		PTPRC	PerCP	BD	345809
CD56		NCAM1	PE	Beckman Coulter	A07788
CD200	OX2	CD200	PE	BioLegend	329205
CD271		NGFR	PE	BD Pharmingen	557196
control IgG1			APC	R&D	IC002A
CXCL13		CXCL13	APC	R&D	IC801A
CXCR5	CD185	BLR1	PE	R&D	FAB190P
ENG	CD105	ENG	PerCP	R&D	FAB10971C
EpCAM	CD326, TACSTD1	EPCAM	FITC	Myltenyi Biotec	130-080-301
PD-1	CD279	PDCD1	FITC	eBioscience	11-9969
STRO-1			FITC	Santa Cruz Biotechnology	sc-47733

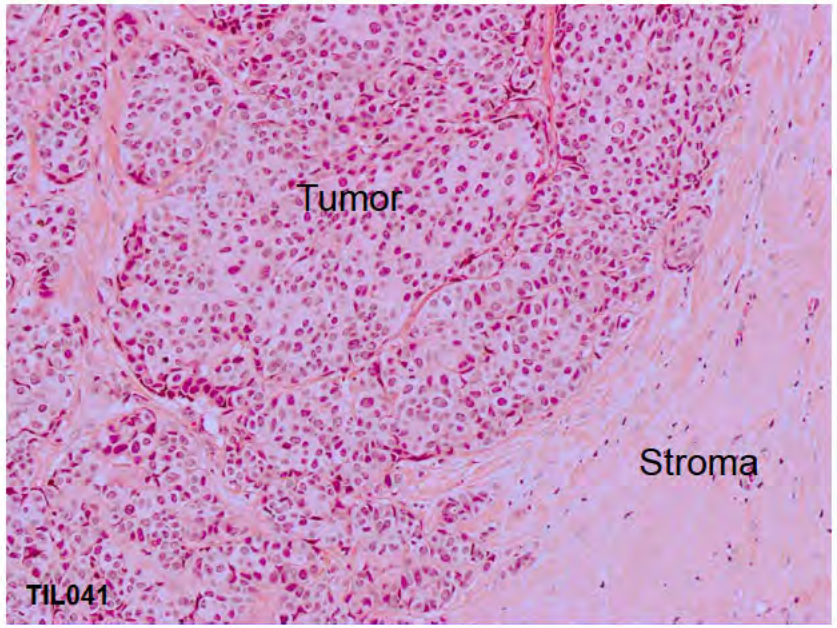
Gu-Trantien, *et al*.
Supplemental
Figure 4:

Hematoxylin & Eosin and Immunohistochemical Staining of Human Breast Tumors

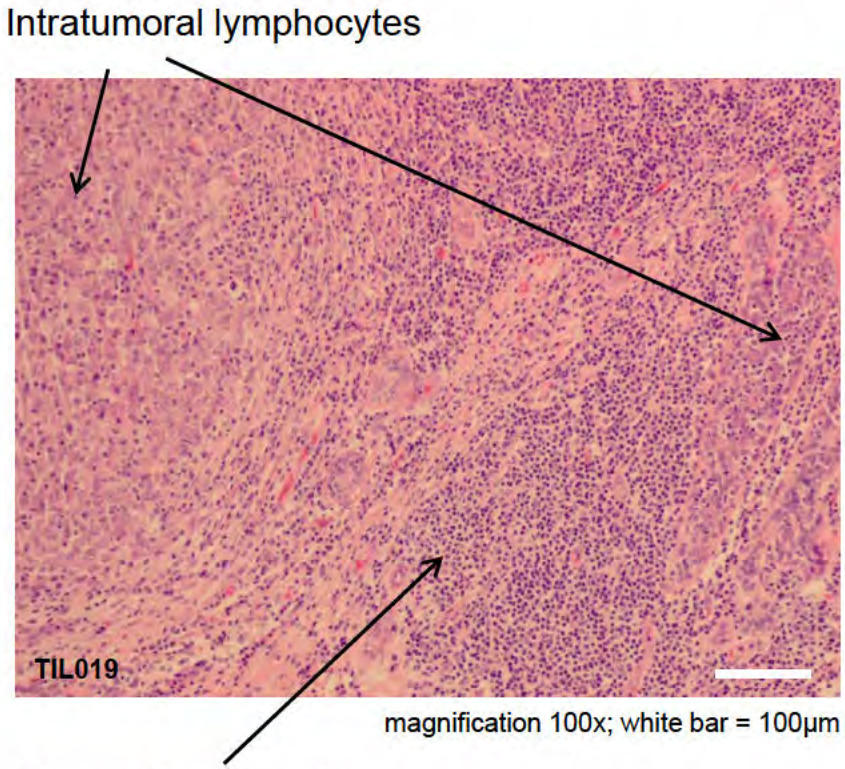
(Patients from the Confirmation Set; Table S1C)

Figure S4A:
Lymphocytes Infiltrating Breast Carcinomas
H&E staining

Minimally infiltrated tumor (Min)



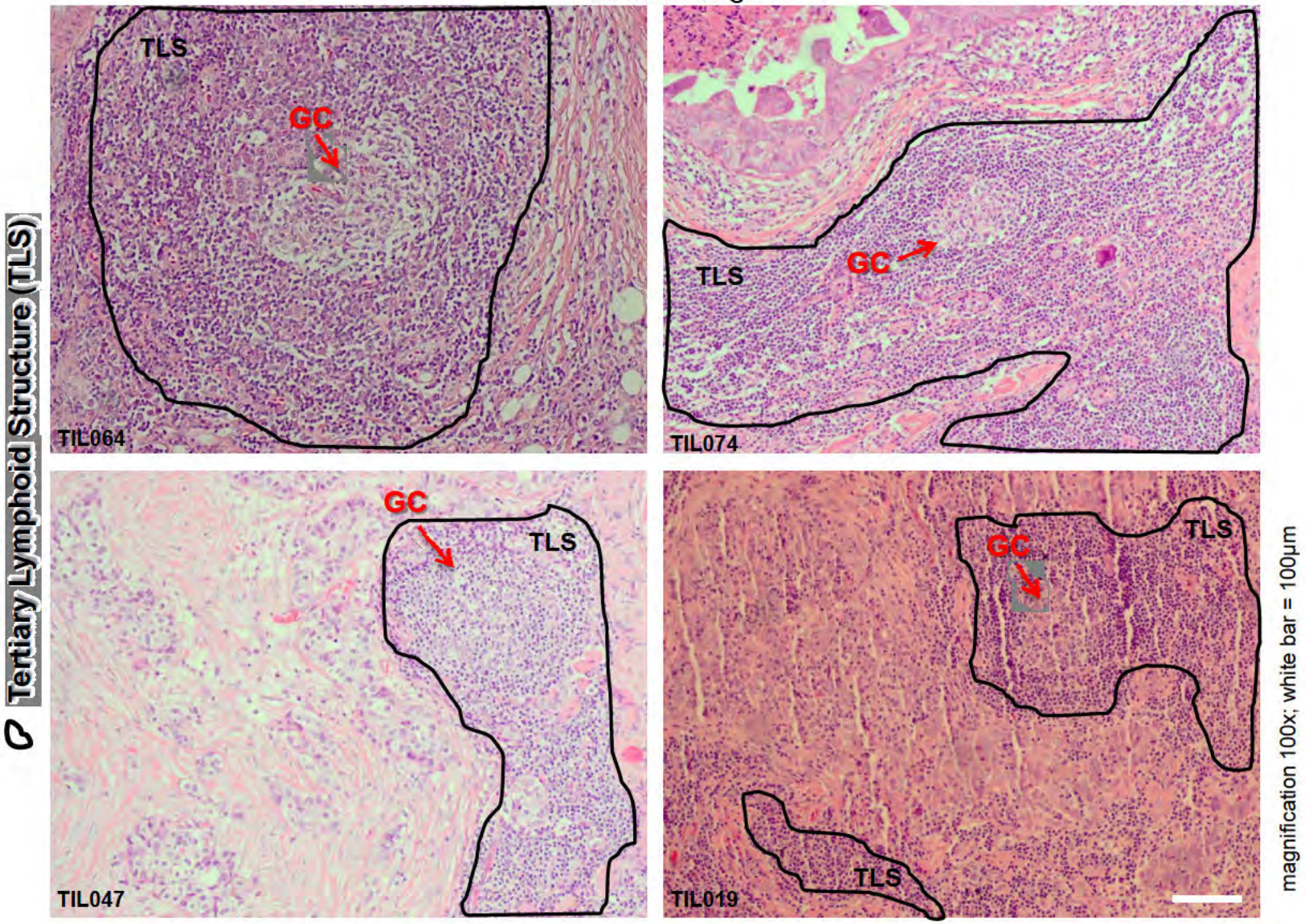
Extensively infiltrated tumor (Ext)



Lymphocytes infiltrates in human breast tumors vary between minimal to extensive (Table S1, C and D). Examples of a minimally infiltrated and an extensively infiltrated tumor are shown above.

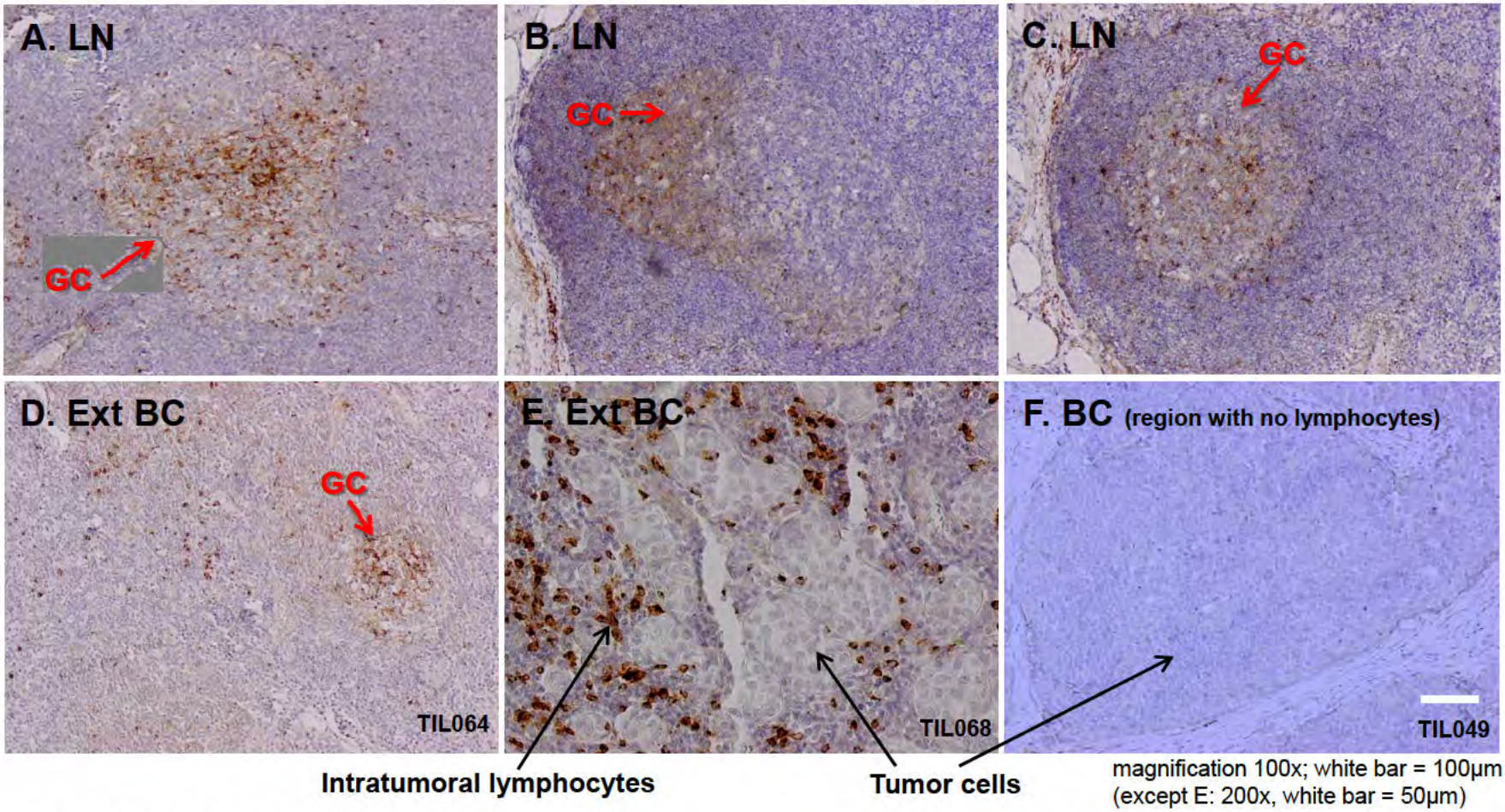
Figure S4B:
Lymphocytes Infiltrating Breast Carcinomas

H&E staining



Well organized tertiary lymphoid structures (TLS) with reactive germinal centers (GC) were detected adjacent to the tumor bed at a higher incidence in extensively compared to minimally infiltrated tumors (Table S1, C and D).

Figure S4C:
CXCL13 expression in Lymph Nodes (LN) and Breast Tumors (BC)
IHC staining



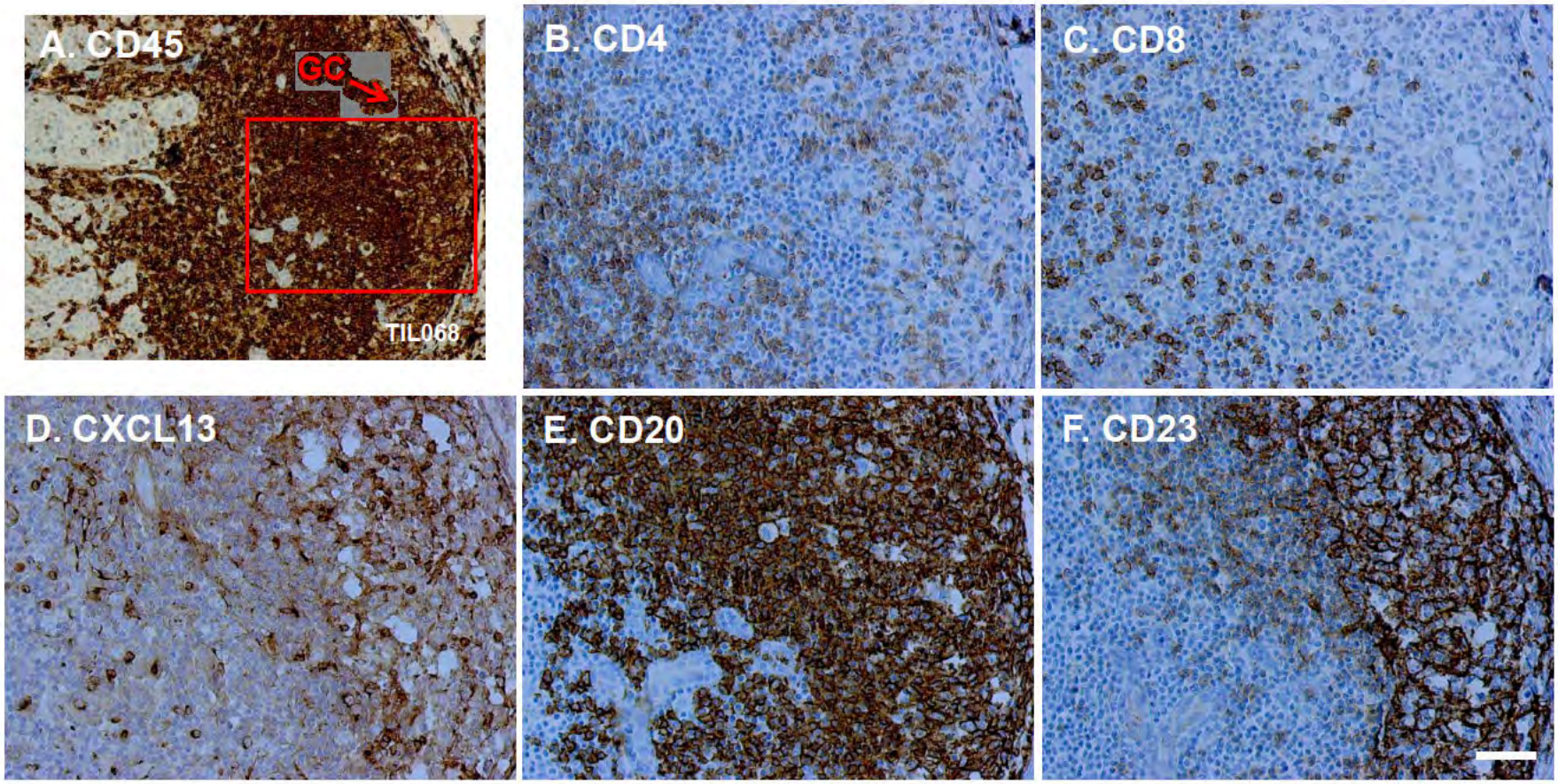
A-C. LN from BC patients:

Consistent with published data [Wang C, Hillsamer P, Kim CH. BMC Immunol. (2011) 12:53], CXCL13 is essentially expressed in the light zone of LN GC, where $CD4^+$ follicular helper T cells (Tfh) and follicular dendritic cells (FDCs) are located.

D-F. Breast tumors:

(D) CXCL13 is expressed in BC TLS, particularly in the GC. (E) CXCL13 expression is associated with intratumoral lymphocytes but not tumor cells. (F) CXCL13 is absent from tumor regions with no infiltrating lymphocytes.

Figure S4D:
Leukocytes Infiltrating Breast Carcinomas
TLS in TIL068; IHC staining

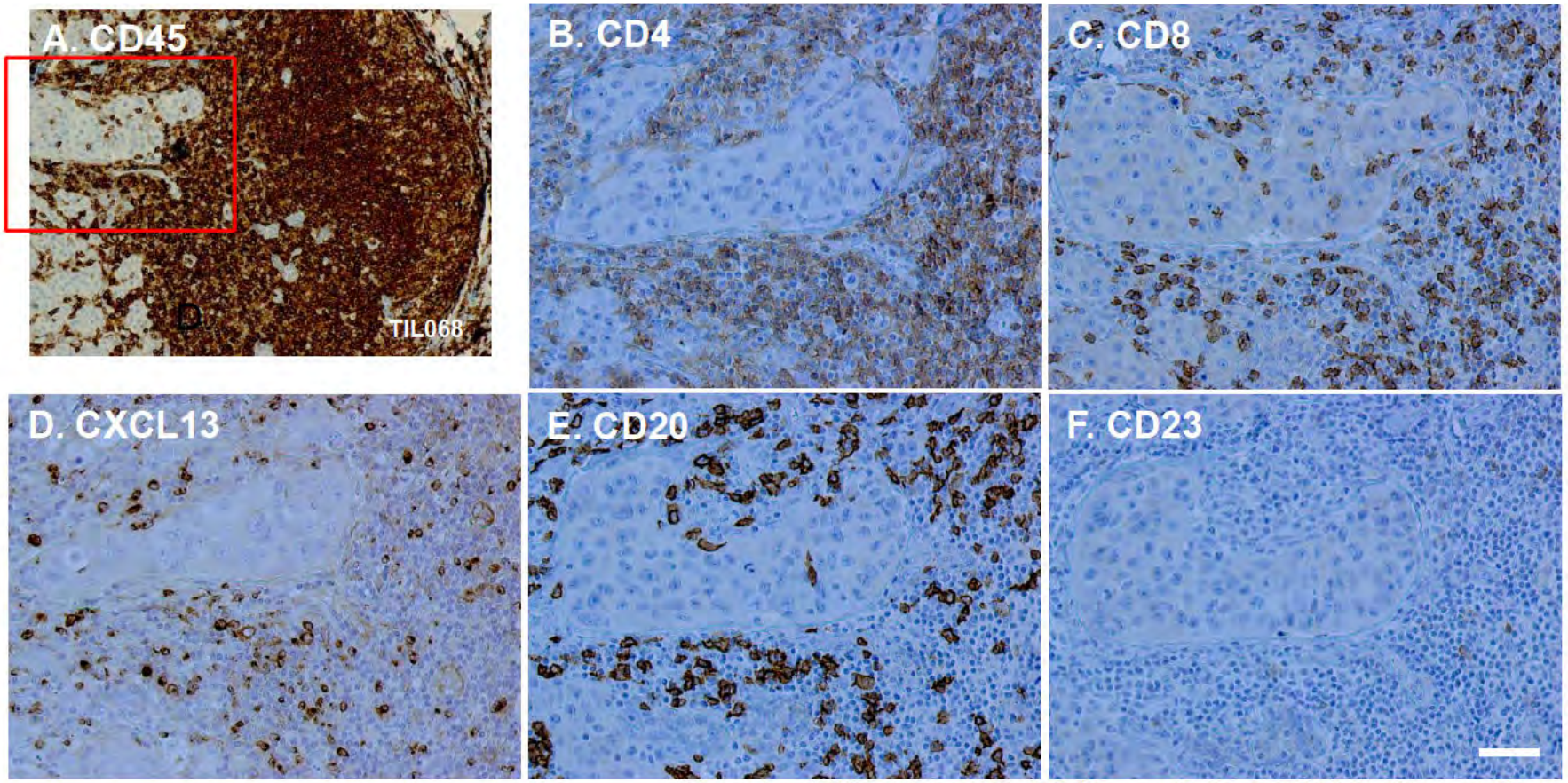


The region highlighted in red (A; 100x) is magnified in B-F (200x); white bar = 50µm

Among the lymphocytes infiltrating the center of this TLS, CXCL13 is expressed by cells that co-localize with $CD20^+$ B cells. Cytoplasmic CXCL13 staining is essentially limited to small lymphocytes and a few FDCs in the GC.

*This supplemental figure shows higher magnifications of subregions for some of the markers shown in Figure 10C.

Figure S4E:
Leukocytes Infiltrating Breast Carcinomas
TLS in TIL068; IHC staining

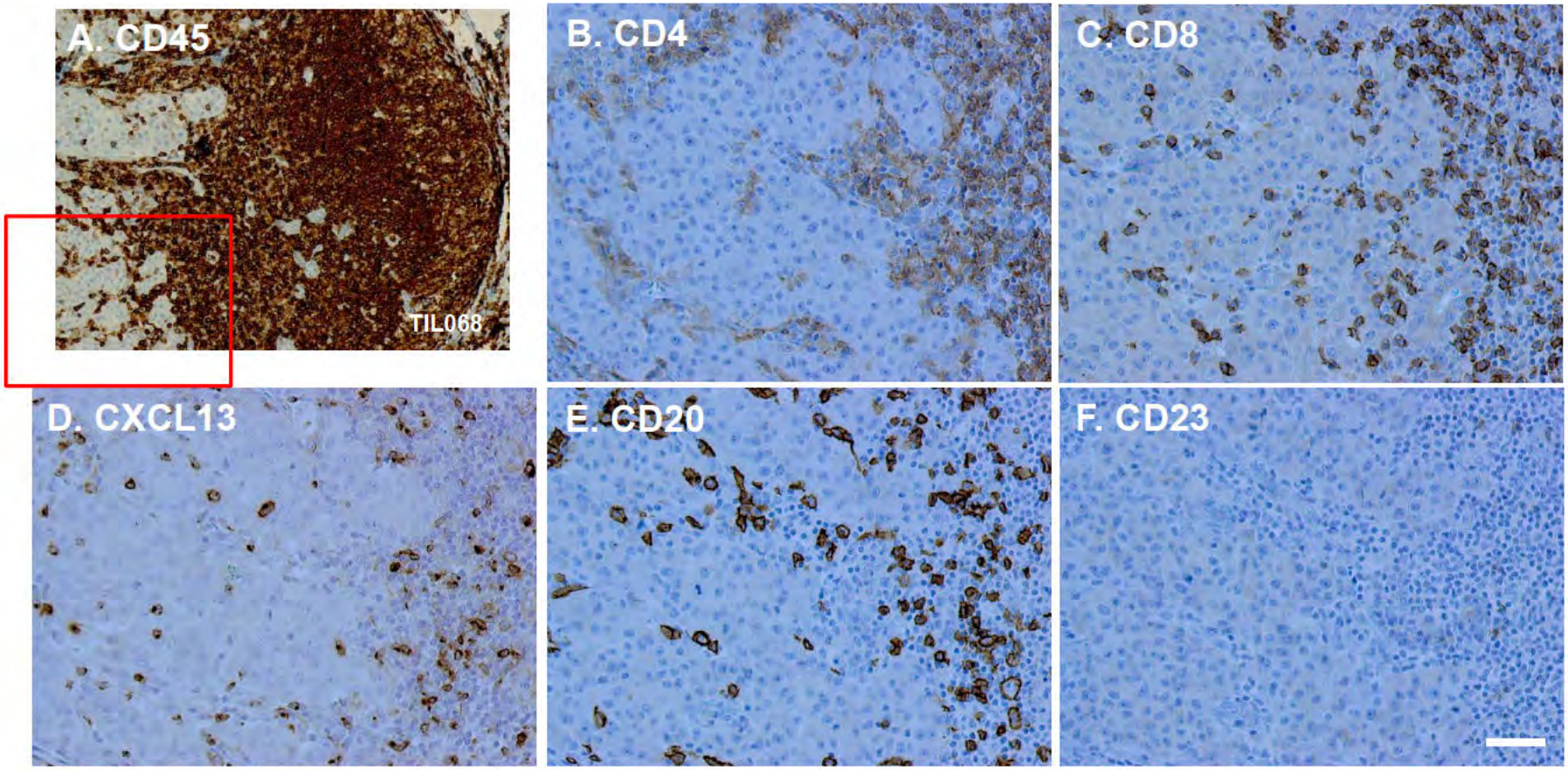


The region highlighted in red (A; 100x) is magnified in B-F (200x); white bar = 50µm

Amon_ the l_mphoc_tes infiltratin_ a tumor cell re_ion, CXCL13 is a_ain expressed b_ cells that co-localize with CD20⁺ B cells. Cytoplasmic CXCL13 staining is essentially limited to small lymphocytes.

*This supplemental figure shows higher magnifications of subregions for some of the markers shown in Figure 10C.

Figure S4F:
Leukocytes Infiltrating Breast Carcinomas
TLS in TIL068; IHC staining

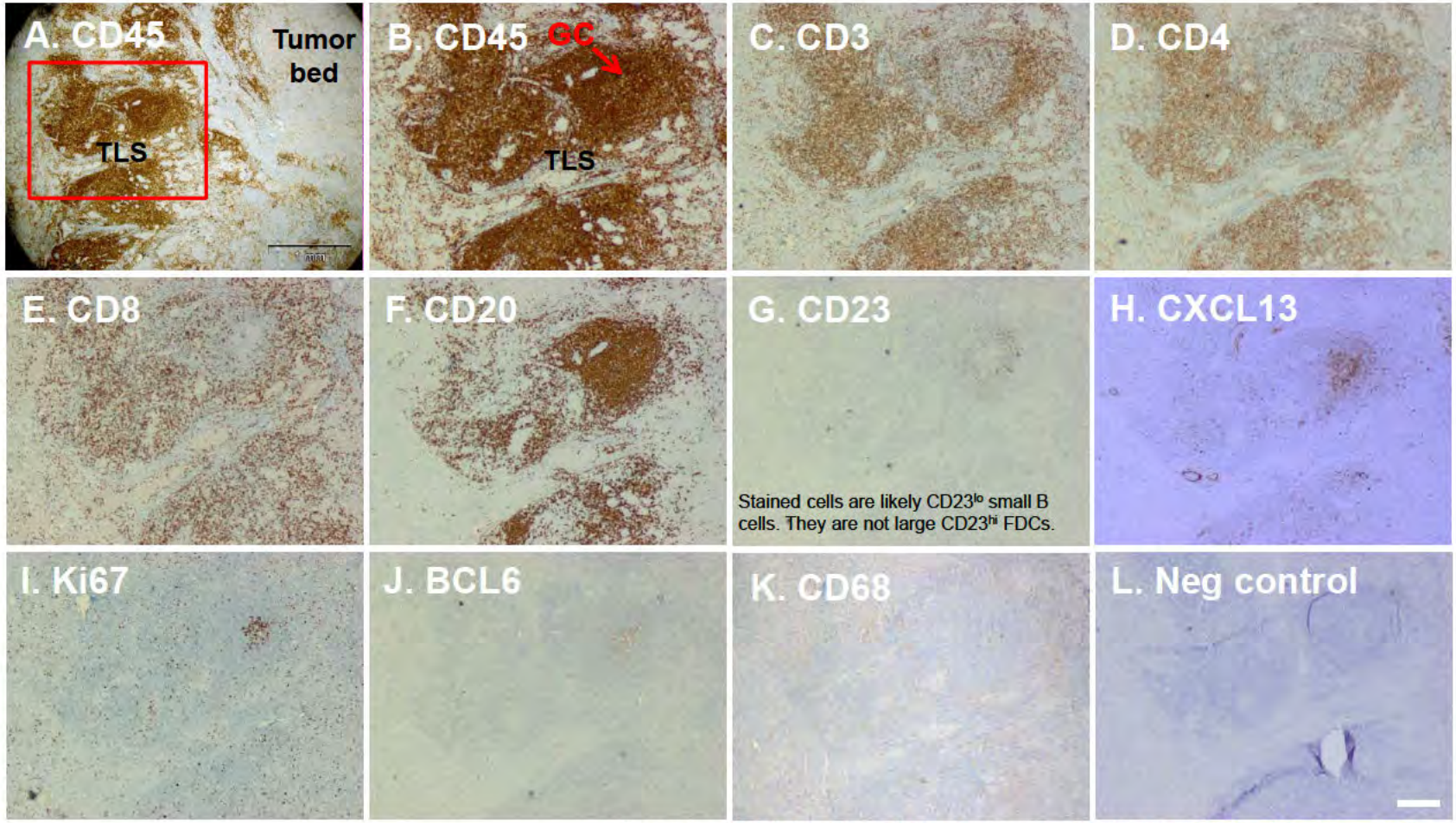


The region highlighted in red (A; 100x) is magnified in B-F (200x); white bar = 50μm

Amon_ the l_ mphoc_tes infiltratin_ a different re_ion of the tumor, CXCL13 is a_ain expressed b_ cells that co-localize with CD20⁺ B cells. Cytoplasmic CXCL13 staining is also essentially limited to small lymphocytes.

*This supplemental figure shows higher magnifications of subregions for some of the markers shown in Figure 10C.

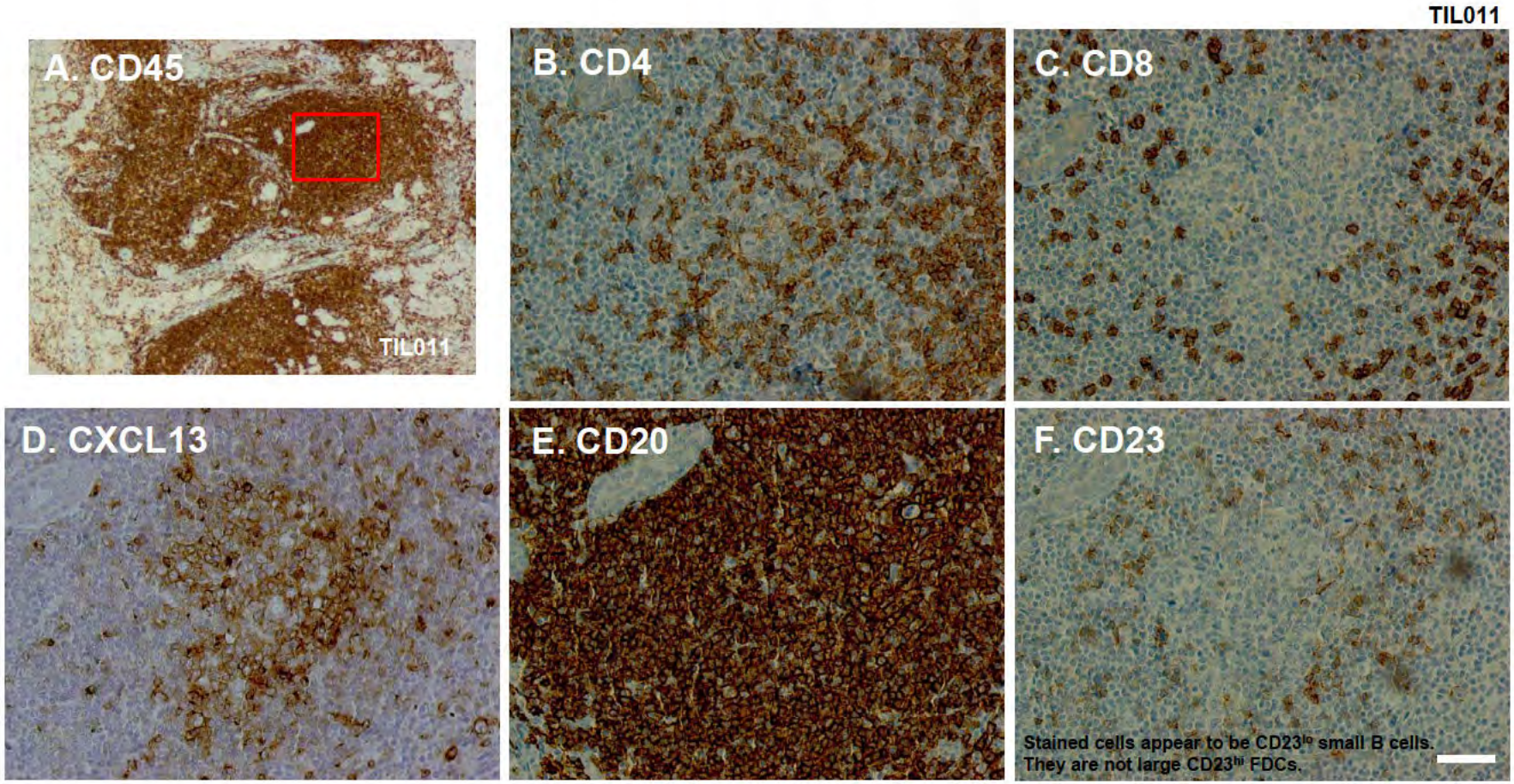
Figure S4G:
Leukocytes Infiltrating Breast Carcinomas
TLS in TIL011; H&E and IHC staining



The region highlighted in red (A; 20x) is magnified in B-F (40x); white bar = 250μm

CD3⁺ T cells form aggregates with CD20⁺ B cells in this TLS; however ,CD23^{hi} FDCs were not detected in the small reactive GC (undetectable on H&E slides). This GC contains a small center of proliferating Ki67⁺ centroblasts and BCL6⁺ cells. The majority of CD3⁺ cells in the T cell zone are CD4⁺ T cells, with some in the B cell follicle GC. CD8⁺ T cells are at a lower concentration in the T cell zone and absent from the GC.

Figure S4H:
Leukocytes Infiltrating Breast Carcinomas
TLS in TIL011; IHC staining

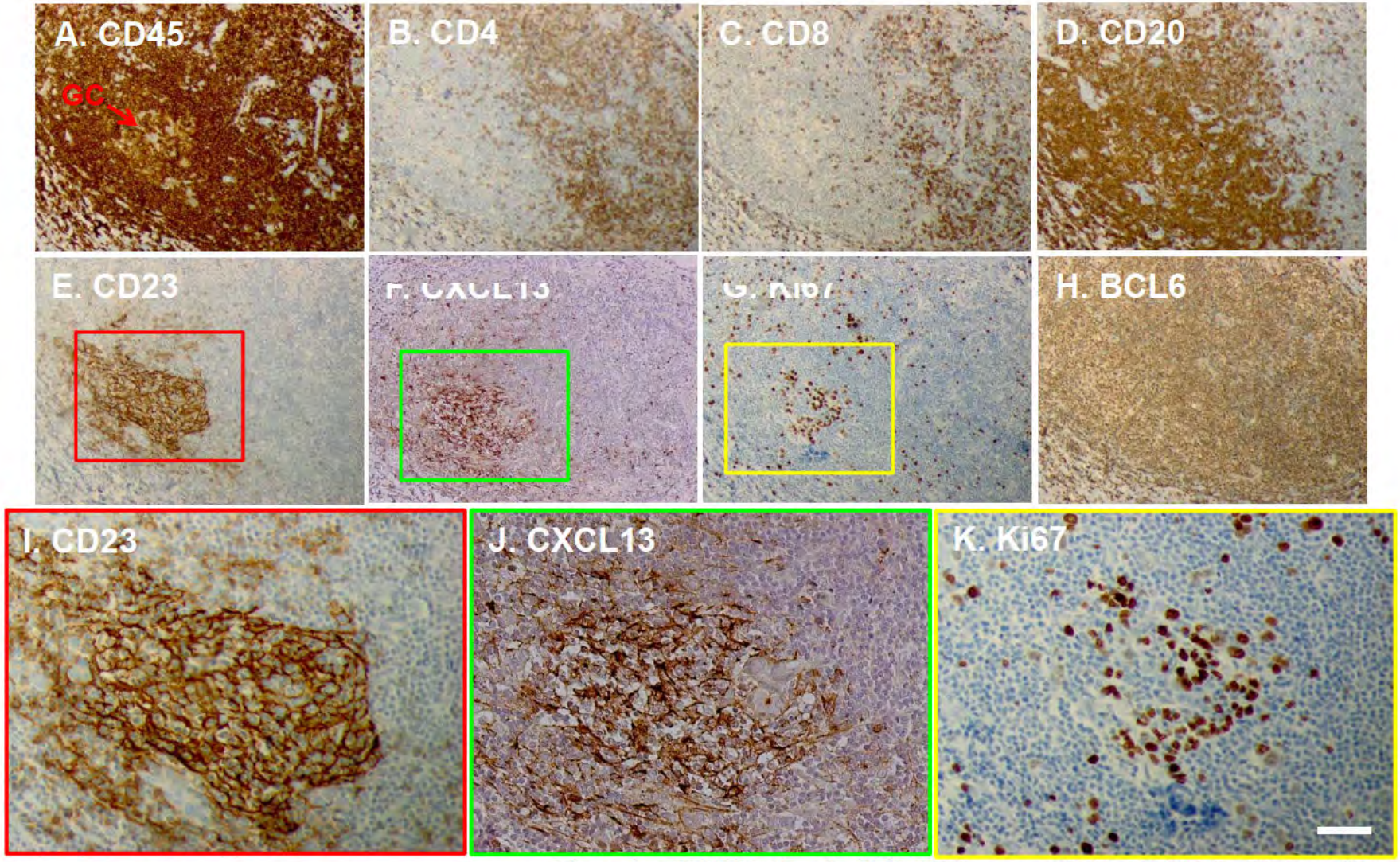


The region highlighted in red (A; 40x) is magnified in B-F (200x); white bar = 50μm

This TLS does not contain large CD23^{hi} FDCs. CXCL13 is ex. ressed . rinci. all. in the GC within the B cell follicle where some CD4^{+} T cells (likely Tfh) are co-localized with CD20^{+} B cells.

*This supplemental figure shows higher magnifications of subregions for some of the markers shown in Figure S4G.

Figure S4I:
Leukocytes Infiltrating Breast Carcinomas
TLS in TIL064; IHC staining



The regions highlighted in color (A-H; 100x) are magnified in I-K (200x); white bar = 50μm

This TLS contains many CD23^{hi} large FDCs and a massive number of BCL6^{+} lymphocytes but few highly proliferating Ki67^{+} centroblasts. CXCL13 is essentially expressed by the FDCs in the GC within the B cell follicle.

Figure S4J:
Leukocytes Infiltrating Breast Carcinomas
Overview of IHC staining analysis

Patient ID[a]	Parameter	CD45		CD4		CD8		CD20		CD23		CD68		CXCL13 Global Intensity[b]	BCL6[b]		Ki67[b]	
		Intra-tumoral	Stromal	Intra-tumoral	Stromal	Intra-tumoral	Stromal	Intra-tumoral	Stromal	Intra-tumoral	Stromal	Intra-tumoral	Stromal		Tumor	TIL	Tumor	TIL
TIL011	Intensity[b]	++	+++	+	++	++	++	0/+	+	0	0	0/+	+	+	0/+	0/+	++	+
	Cell organization[e]	D	D/A	D	A/D	D	D/A	D	D/A			D	D					
TIL019	Intensity	++++	+++	++	++	+++	++	+	+	0/+	+	+++	+	+++	0/+	0/+	+++	+
	Cell organization	D	D/A	D	D/A	D	D/A	D	A/D	D	A	D	D					
TIL047	Intensity	0/+	+	0	+	0	0/+	0	+	0	n.d.[f]	0	0/+	0/+	n.d.	n.d.	+++	0/+
	Cell organization	D	A/D		A/D		A/D		A/D			D	D					
TIL064	Intensity	+++	++++	+	+++	++	+++	0/+	+++	0	+	0/+	0/+	++	++	++++	+++	++
	Cell organization	D	A/D	D	A/D	D	A/D	D	A/D		A	D	D					
TIL068	Intensity	++	+++	+	++	+	+	0/+	++	0	+	+	++	++	+	0/+	+++	+
	Cell organization	D/A	A/D	D	A/D	D	D/A	D	A/D		A	D	D					
TIL027	Intensity	+	++	0/+	+	+	+	0	+	0	0	0	0	+				
	Cell organization	D	A/D	D	A/D	D	D		D/A									
TIL029	Intensity	++	++	+	++	++	+	0	+	0	0/+	++	+	+				
	Cell organization	D	D/A	D	A/D	D	D/A		A/D		A	D	D					
TIL040	Intensity	0/+	+	0	+	0/+	+	0	+	0	+	0/+	+	0/+				
	Cell organization	D	A/D		A	D	D		A		A	D	D					
TIL042	Intensity	+	++	0/+	++	+	+	0/+	+	0	0	0/+	+	+				
	Cell organization	D	A/D	D	A/D	D	D	D	D/A			D	D					
TIL049	Intensity	0/+	+	0/+	0/+	0	+	0	0/+	0	0/+	0/+	+	+				
	Cell organization	D	D/A	D	D		D/A		D/A		A	D	D					
TIL054	Intensity	0/+	++	0	++	0/+	+	0	+	0	0	0/+	+	+				
	Cell organization	D	A/D		A/D	D	D/A		D/A			D	D					
TIL055	Intensity	0/+	+	0/+	+	0/+	0/+	0	+	0	0/+	0/+	0/+	0/+				
	Cell organization	D	A/D	D	A/D	D	D/A		A		A	D	D					
TIL069	Intensity	+	+++	0/+	++	+	+	0	+	0	0/+	+	+	+				
	Cell organization	D	D/A	D	D	D	D		A/D		A	D	D					
TIL074*	Intensity		+++		++		+		++		+		+	+				
	Cell organization		A		A		D		A		A		D					
TIL129	Intensity	++	+++	+	+++	++	+	+	++	0	0/+	++	++	++				
	Cell organization	D	A/D	D	A/D	D	D	D	A/D		A	D	D					

[a]Patients from the confirmation set (Table S1C)

[b]IHC staining intensities were evaluated at 6 levels (relative levels within the assessed tumors); range includes 0, 0/+, +, ++, +++ & ++++

[c]TLS (Tertiary Lymphoid Structure) levels were evaluated at 6 levels (relative levels within the assessed tumors); range includes 0, 0/+, +, ++, +++ & ++++ (determined by CD45 staining)

[d]TLS size was estimated at 3 levels (relative levels within the assessed tumors) including small (S), medium (M) & large (L)

[e]Cellular organization (for CD45, CD4, CD8, CD20, CD23 & CD68 staining) was identified in four patterns: D (= diffuse); A (= aggregate); D/A (= D+A with D>A); A/D (= D+A with A>D)

[f]n.d. = not-determined

*TIL074 was minimally invasive and therefore intra-tumoral lymphocyte levels could not be accurately determined

Figure S4K:
Antibodies used to investigate major leukocyte populations in human breast tumors by IHC staining

Antibody	Gene symbol	Dilution	Source	Reference
Using Endogenous Biotin Blocking Kit (Ventana) and iVIEW DAB Detection Kit (Ventana) on BenchMark XT IHC/ISH slide stainer (Ventana)				
CD3	CD3E	1/100	Dako	A0452
CD4	CD4	1/20	BioSB	BSB5153
CD8	CD8A/B	1/50	Dako	M71031
CD20	MS4A1/2	1/200	Dako	M075529
CD23	FCER2	ready to use	Ventana	790-4408
CD45	PTPRC	1/100	Dako	M0701
CD68	CD68	ready to use	Dako	IS60930
Ki-67	MKI67	1/50	Dako	M724001
BCL6	BCL6	1/20	Leica Microsystems	NCL-L-BCL6-564
Using the method described below				
CXCL13	CXCL13	1/100	R&D Systems	AF801