Full title: Cell-intrinsic genetic regulation of peripheral memory-phenotype T cell frequencies

Short title: Cell-intrinsic genetic regulation of T cell memory

Amanpreet Singh Chawla[1*], Parna Kanodia[1*], Ankur Mukherjee[2], Vaibhav Jain[1], Gurvinder Kaur[3], Poonam Coshic[4], Kabita Chatterjee[4], Nitya Wadhwa[5], Uma Chandra Mouli Natchu[5], Shailaja Sopory[5], Shinjini Bhatnagar[5], Partha P. Majumder[2], Anna George[1†], Vineeta Bal[1†], Satyajit Rath[1†], Savit B. Prabhu[1,5†]

[1]National Institute of Immunology, New Delhi, India.

[2]National Institute of Biomedical Genomics, Kalyani, India.

[3]Laboratory Oncology, Dr. BR Ambedkar Institute Rotary Cancer Hospital, All India Institute of Medical Sciences, New Delhi, India.

[4]Department of Transfusion Medicine, All India Institute of Medical Sciences, New Delhi, India.

[5]Pediatric Biology Center, Translational Health Science and Technology Institute, Faridabad, Haryana, India.

[*]These authors contributed equally to the manuscript.

[†]These authors contributed equally to the manuscript.

Address for correspondence: Savit B. Prabhu, Translational Health Science and Technology Institute, NCR Biotech Science Cluster, 3rd Milestone, Faridabad – Gurgaon Expressway, PO box #04, Faridabad – 121001, Haryana, India (tel: +91129 2876345; email: savit@thsti.res.in).

Abstract

Memory T and B lymphocyte numbers are thought to be regulated by recent and cumulative microbial exposures. We report here that memory-phenotype lymphocyte frequencies in B, CD4 and CD8 T-cells in 3-monthly serial bleeds from healthy young adult humans were relatively stable over a 1-year period, while recently activated -B and -CD4 T cell frequencies were not, suggesting that recent environmental exposures affected steady state levels of recently activated but not of memory lymphocyte subsets. Frequencies of memory B and CD4 T cells were not correlated, suggesting that variation in them was unlikely to be determined by cumulative antigenic exposures. Immunophenotyping of adult siblings showed high concordance in memory, but not of recently activated lymphocyte subsets, suggesting genetic regulation of memory lymphocyte frequencies. To explore this possibility further, we screened effector memory (EM)-phenotype T cell frequencies in common independent inbred mice strains. Using two pairs from these strains that differed predominantly in either CD4EM and/or CD8EM frequencies, we constructed bi-parental bone marrow chimeras in F1 recipient mice, and found that memory T cell frequencies in recipient mice were determined by donor genotypes. Together, these data suggest cell-autonomous determination of memory T niche size, and suggest mechanisms maintaining immune variability.

Introduction

Gene-environment interplay in immune phenotypes has been extensively studied using steady-state cellular immune profiles (1–3), functional immune responses (2,4–8), post-vaccination responses (9,10) and V,D and J usage biases in naïve and memory T and B cell compartments (11). Some of these studies have identified genomic correlates associated with specific steady-state immune phenotypes (2) and vaccine responses (12). However, there are conflicting findings regarding the relative importance of genetic versus environmental factors in regulation of immune phenotype (1,2,13,14), warranting further investigations at the population-level in humans and mechanistic studies in mice on regulation of individual immune phenotypes.

Memory subsets in T and B lymphocytes are immune populations that are generated in response to past immunogen exposure. Immunological memory, providing long-term persistence of antigen-experienced cells contributing to rapid and robust responses following re-exposure (15), is likely to have evolved in an ecosystem where environmental challenges including repeated infections would be the norm (16) and the persistence of long-lasting antigen-specific cells generated during immune responses would confer survival advantage (17). However, it is possible that larger memory lymphocyte pool sizes may carry costs such as restriction of space for the more repertoire-diverse naive T cell compartment (18), attrition of pre-existing memory (19), and other bioenergetic costs (20). Such selection may well result in ‘optimum’ sizes of memory lymphocyte pools (21), and these pool sizes could show population diversity, depending on the diversity of pathogens in the ecosystem and their exposure rates (22). Diversity in pool sizes of memory lymphocytes in a population could thus be determined by a combination of genetic variability and diversity of environmental exposures.

A number of mechanisms can be envisaged regulating the pool size of the memory T cell compartment, including cumulative life-time antigen-exposure and re-exposure, antigenic persistence, degree of expansion, cell survival, attrition and niche-space availability (21,23–26). Immune cells occupy a limited niche space in lymphoid organs (27) or in the periphery (28). This niche size could be a function of size and/or structure of supporting lymphoid tissue architecture (29), along with intrinsic properties of cells occupying the niche. Similar determinants could affect steady state levels of transient cell populations such as immediate-effector T cells (30) and plasmablasts, but their steady state levels would be expected to fluctuate more with short-term environmental changes.

On this background, we have explored the possibility of gene-environment interplay affecting the steady state pool sizes of lymphocytes post-activation. We have used peripheral blood leucocyte immunophenotyping in serial bleeds from healthy young adult human volunteers to assess inter-individual and temporal variability in lymphocyte memory and effector compartments, and report that while both categories showed inter-individual variability, temporal variability was far less in memory than in effector subsets. We have next used peripheral blood leucocyte immunophenotyping of human siblings, and report that memory lymphocyte subset pool sizes were far closer between siblings than between non-siblings. We then further confirmed a role for genetic factors in determining memory lymphocyte pool sizes in independent inbred mouse strains, in which bi-parental mixed bone marrow chimeras demonstrated a parental genotype-driven inheritance of CD4 and CD8 T cell memory pool sizes. Together, our data from both humans and mice provide novel insights in the determination of memory lymphocyte pool size.

Materials and methods

Ethics approval and consent to participate

Human:

Informed written consent was obtained from all human volunteers who participated in the study. The study was approved by ethics committees of National Institute of Immunology (approval number IHEC/AKS/45/2013) and All India Institute of Medical Sciences (approval number IEC/NP-471/2013). The methodologies used in the study were in accordance with approved guidelines. All experimental protocols used in this study were approved by Institutional Ethics Committee (Human Research) of National Institute of Immunology and All India Institute of Medical Sciences.

Animal:

Mice were maintained and used according to the relevant rules and regulations of Government of India and with the approval of Institutional Animal Ethics Committee, National Institute of Immunology (approval number 381/15). All mice experimental protocols were in accordance with approved guidelines and were approved by Institutional Animal Ethics Committee of National Institute of Immunology.

Human subjects and blood collection:

Forty-five healthy adult volunteers were recruited into the study and were bled 3-monthly for 12 months (total of 4 bleeds). Individuals with acute or chronic illness, medication, recent vaccination or pregnancy in women were excluded. Ten ml heparinized peripheral blood was collected from healthy subjects after informed written consent. The study was approved by institutional ethics committee of National Institute of Immunology (NII). All protocols were in accordance with approved guidelines. For sibling study, 37 families were recruited, in which two or more siblings were willing for participation. The family-based study was approved by institutional ethics committees of NII and All India Institute of Medical Sciences (AIIMS).

Exclusion criteria were similar to that mentioned above. For looking at correlations between the various memory and effector cell subsets, we utilized the published raw data from a larger (n=71) cohort of adult healthy volunteers who participated in a previous study (31).

Sample processing and flow cytometry of human samples:

Blood samples were processed without delay (less than three hours). Total leukocyte counts and differential counts were obtained from standard hematologic methods. Peripheral blood mononuclear cells (PBMCs) were separated by density gradient centrifugation. PBMCs were washed, counted and divided into aliquots of about 1 million cells per ml per vial and cryopreserved in 10% DMSO in bovine serum until assays were performed. For flow cytometry, PBMCs were thawed, washed and incubated with the following antibodies organized into two cocktails: The T cell cocktails consisted of CD3 (UCHT1, eBioscience), CD4 (OKT4, eBioscience), CD8 (SK1, eBioscience), CD45RO (UCHL1, eBioscience), CCR7 (150503, BD) and the B cell cocktail consisted of CD19 (SJ25C1, BD), CD20 (2H7, eBioscience), CD38 (HIT2, eBioscience), CD43 (eBio84-3C1, eBioscience), CD27 (M-T271, eBioscience). A single control sample was run along with all the samples to ensure that gating is comparable across experiments. Samples were acquired in BD Verse and analysis was done using flowjo (Treestar). CD4 and CD8 memory cells were defined as the CD45RO+ fractions (32) of CD8 and CD4 cell compartments. Transient T effector-memory, RA+ve (TEMRA) cells were defined as CD45RO-CCR7- subsets (32) of CD8 cells and of CD4 cells. Memory B cells were defined as CD27+CD43- fraction of the CD19+CD20+ B cell subset (32). Plasmablasts were defined as the CD38+CD20- subset of CD19+ B cells (32).

Mice:

The following strains of mice were used for the study: C57Bl/6J, BALB/cJ, SJL/J, CBA/CaJ, B6.SJL-Ptprca Pepcb/BoyJ (B6.SJL), CB10-H2b/LilMcdJ (BALB/b), FVB/NJ, C3H/HeOuJ. Mice strains used in the study were obtained as breeder stock from the Jackson Laboratory (Bar

Harbor, ME) and bred and maintained under specific-pathogen free conditions in the Small Animal Facility of the National Institute of Immunology. All mice were maintained and used according to the relevant rules and regulations of Government of India and with the approval of Institutional Animal Ethics Committee, National Institute of Immunology. All experiments used age and gender matched mice. Littermates were used as controls. Phenotyping of mouse spleen T cells was done using the following antibodies (all from eBioscience): CD4 (RM4-5), CD8 (53-6.7), CD44 (IM7), CD62L (MEL-14), CD25 (PC61.5), CD45.1 (A20) and CD45.2 (104). NK cells were phenotyped using the following antibodies (all from eBioscience): CD90 (53-2.1), B220 (RA3-6B2) and Ly49b (DX-5). Surface staining was done by incubating 1x106 cells in staining buffer (PBS containing 2% BSA and 0.05% NaN3) for 30 mins on ice. The cells were washed thrice with cold staining buffer and samples were acquired on BD Verse and analysis was done using flowjo (Treestar).

Ex vivo cell preparations

Spleen was dissected from mice euthanized by cervical dislocation and teased between a pair of frosted slides to obtain single cell suspension. Red blood cells in the suspension were lysed by osmotic shock using water, washed and then re-suspended in complete medium.

Mix bone marrow chimeras

F1 hybrids were irradiated at 800rads in gamma chamber (BARC, Mumbai) with Co^{60} as a source for gamma rays. Bone marrow cells from each parental strain were mixed in 1:2, 1:1 and 2:1 ratios and a total of 15 million cells transferred into irradiated F1 generation mice (n=24 for B6.SJL - CBA/CaJ chimera and n=19 for SJL/J - BALB/cJ chimera) and reconstitution allowed for two months. Based on the ratio of CD45.1 to CD45.2 in reconstituted chimera, biological outliers showing extreme ratios more than 10- fold from the expected were removed from further analysis (4 outliers in B6.SJL - CBA/CaJ chimera and 2 outliers in SJL/J - BALB/cJ chimera were removed). Phenotyping was done by using above mentioned CD markers. Doublets were

excluded from the population using height and area parameters of forward and side scatter. CD4 effector memory cells were defined as the CD4+ CD25- CD44hi CD62Llo fractions and CD8 effector memory cells were defined as the CD8+ CD44hi CD62Llo fractions.

Statistical analysis

For frequency and cell count comparisons, a non-parametric test (Mann Whitney) was used. Correlations were estimated using Spearman's correlation coefficient. The Bonferroni method was used to correct for multiple comparisons in testing the significance of elements of the pairwise correlation coefficient matrix. Bootstrapping (resampling) methods were used for comparison of intra-individual variances with between-individual variances and for comparison of sibling pairs vs unrelated pairs. Comparison of multiple inbred mouse strains was done using Analysis of Variance (ANOVA) with post-hoc testing for pair-wise comparison.

Gene expression analysis

Gene expression (microarray) data of sorted splenic 'naïve' CD4+ve CD62L-ve CD4 T cells of mice published previously (33) was obtained from GEO database (accession: GSE60337) and analysed using GEO2R tool. Genes with p-value of less than 0.05 after multiple testing (Benjamini) was considered significantly differentially expressed. Gene Ontology enrichment analysis was done using web-based Gene Ontology enrichment analysis and visualization tool (34) (GORILLA) and Database for Annotation, Visualization and Integrated Discovery (DAVID)(35). To reduce false positive results in enrichment, in all enrichment analyses, the enrichment of candidate genes that were differentially expressed were analysed against a background of all the genes that were used in array experiment.

Results

Human peripheral T and B memory cell levels show temporal stability

Firstly, we examined if variation in the memory T and B cell subsets in the population were relatively stable over time, or were significantly affected by fluctuations possibly contributed by short-term environmental exposures. To test this, we characterized the peripheral blood leucocyte subset phenotype of a group of human volunteers [n=45, male: female ratio = 26:19, mean age = 27.16 years (SD = 6.75)] every 3 months for a period of 1 year. Gating strategy for the memory subsets is indicated in supplementary figure 1 (Fig S1).

Representative data of frequencies of each subset over 4 time points are shown in Fig S2 and are shown quantified in Figure 1. Variations of memory B, memory CD4 and memory CD8 frequencies were significantly lower within individuals than between individuals (Fig. 1A, 1B, 1C). On the other hand, intra-individual variances in plasmablasts (Fig. 1D) and CD4 TEMRA cells (Fig. 1E) were no different from inter-individual variances. When absolute cell numbers per µl blood were calculated, broadly similar trends were seen (Fig. 1G, 1H, 1I, 1J, 1K). Curiously, CD8 TEMRA frequencies (Fig. 1F) and counts (Fig. 1L) tended to be similar to CD8 memory (Fig. 1C, 1I) in that they showed lower intra-individual than inter-individual variation. These data suggested that intra-individual short-term fluctuations in memory B and T subsets are unlikely to be a major explanation for the variation that is seen in memory lymphocyte subsets in humans.

Figure 1: **Comparison of intra-individual and inter-individual variance for immune subsets in Humans.**

Box plots indicate median and interquartile ranges of variances of cell frequencies and counts in Human volunteers. Upper whisker extends till the highest value that is within 1.5 times the

interquartile range from 3rd quartile. Lower whisker extends till the lowest value that is within 1.5 times the interquartile range from 1st quartile. Outliers are shown as dots. Intra-individual variances indicate variance of subset frequency or count over 4 time points in each individual (n=45). Inter-individual variances indicate variance of subset frequency or count in randomly chosen set of different individuals (n=45). P-values obtained by bootstrapping are as indicated in the panels.

Correlations between CD4, CD8 and B cell memory

Since memory cells can be long-lived and could have been generated by antigenic exposure in the relatively remote past, it remained possible that cumulative antigen exposures contribute to determining steady-state memory T and B cell levels. Microbial exposure commonly leads to generation of CD4 T-dependent B cell responses, since CD4 T cells are necessary for B cell germinal centre responses and hence vital to memory B cell generation. Thus, it would be expected that there would be coordinated accumulation of memory cells in the CD4 T and B cell compartments leading to a positive correlation between CD4 T memory and B memory cell frequencies, if cumulative antigenic exposures were to contribute substantially to determining memory T and B cell frequencies. However, there was no correlation between CD4 T cell memory frequencies and B cell memory frequencies in a given individual (Fig. 2A). Thus, cumulative antigen exposure may not be a major determinant of memory CD4 and B cell levels. Interestingly, CD4 and CD8 memory T cell subsets showed a strong correlation (correlation coefficient = 0.61, $p < 0.001$) (Fig. 2B), suggesting shared determinants of memory cell subsets within the T cell lineage. On the other hand, frequencies of CD4 TEMRA or CD8 TEMRA cells

236 or plasmablasts did not show any correlations with memory T or B cell frequencies or with each
237 other (Table 1).
238

Pairs of variables	Correlation coefficient (Spearman)	Adjusted p-value (Bonferroni)
Plasmablast – B memory	0.17	1
CD4 memory - B memory	0.0009	1
CD4 TEMRA - B memory	-0.006	1
CD8 memory - B memory	-0.09	1
CD8 TEMRA - B memory	0.038	1
CD4 memory - Plasmablast	-0.04	1
CD4 TEMRA - Plasmablast	0.09	1
CD8 memory – Plasmablast	-0.0017	1
CD8 TEMRA - Plasmablast	0.16	1
CD4 TEMRA - CD4 memory	-0.26	0.335
CD8 memory - CD4 memory	0.6	8.90E-08
CD8 TEMRA - CD4 memory	-0.064	1
CD8 memory - CD4 TEMRA	-0.16	1
CD8 TEMRA - CD4 TEMRA	0.32	0.065
CD8 TEMRA - CD8 memory	0.2	1

239
240 **Table 1: Correlation coefficient (Spearman) of correlation between cell subsets in adult**
241 **volunteers**
242

243

244

Figure 2: **Correlation between memory B and memory CD4 T cell frequencies (A) and between memory CD4 and memory CD8 T cell frequencies (B)**. Memory CD4, memory CD8 and memory B cell frequencies are expressed as percentage of total CD4, CD8 and B cells respectively. Each dot represents data from one individual donor (n=71). Correlation coefficient (Spearman) and p-value are as indicated.

Siblings show concordance in T and B cell memory levels

Since these descriptive and correlative data from humans showed associations that suggested that memory phenotypes might be genetically regulated, we explored this further by a family based study where peripheral blood leucocytes from 80 full siblings from 37 families were immunophenotyped [male: female ratio = 38:42, mean age = 31.8 years (SD=12.7)]. We compared the differences in cell frequencies between siblings with the differences between random pairs of non-siblings for memory CD4, memory CD8 and memory B cells, and for plasmablasts, CD4 TEMRA or CD8 TEMRA cell frequencies. The analysis showed that differences in cell frequencies between siblings were significantly less than the differences between random pairs of non-siblings for memory CD4, memory CD8 and memory B cells, but not for plasmablasts, CD4 TEMRA or CD8 TEMRA cell frequencies (Table 2), suggesting that there was indeed likely to be a genetic contribution to the variations in memory cell frequencies, although possible contributions from early life co-habitation could not of course be ruled out.

Table 2: P-values of difference between within-sibling pairs and between-sibling pairs

Compartment	Sibling (Mean ± SD)	Unrelated Individuals (Mean ± SD)	p-value
B cell memory	6.88 ± 4.45	8.84 ± 1.08	0.03
Plasmablasts	2.05 ± 3.08	2.32 ± 0.33	0.207
CD4 memory	11.17 ± 8.24	15.41 ± 1.91	0.01
CD4 TEMRA	1.56 ± 1.39	0.92 ± 0.14	0.99
CD8 memory	12.66 ± 10.67	17.64 ± 2.19	0.008
CD8 TEMRA	10.04 ± 8.76	10.2 ± 1.27	0.46

Mouse strains differ in their memory T cell frequencies and show cell-autonomous regulation of memory T cell phenotype

These data from humans indicated the possibility of a genetic contribution to the determination of memory T and B cell pool sizes. To examine this possibility more rigorously, we examined independent inbred strains of mice. Since specific markers for B cell memory are uncertain (36) and since this compartment in mice may be smaller than in humans (37), we restricted these studies to T cell memory subsets alone.

The gating strategy used for defining effector memory (EM) CD4 (Fig 3A to 3D) and CD8 T cells (Fig 3E to 3H) in a group of representative mice available in our laboratory are shown in Figure 3. EM populations in CD4 and CD8 subsets were distinguished by unambiguous contours, whereas in some strains (eg. BALB/cJ) the central memory (CM) populations were difficult to gate out from the CD44 negative naïve pool using the conventional markers of memory (CD44 and CD62L). CM phenotype CD4 T cells in the BALB/c strain have been previously shown to contain recent thymic emigrants with a CD44-high phenotype (38). CM phenotype CD8 T cells have also been reported to contain antigen-inexperienced, non-memory ("virtual memory") cell types (39) as well as regulatory CD8 T cells (40), further confounding our analysis of memory subsets. Hence, we have avoided CM phenotype T cells and focused on the unambiguously defined EM populations for interpreting differences in memory between the genetically different strains.

Figure 3: **Gating strategy for memory subsets in mouse splenic CD4 T cells (A to D) and CD8 T cells (E to H).** Each plot shows representative gates from a single mouse strain as indicated. CD4 memory subsets are gated on conventional CD4+ CD25- gate. CD8 memory subsets are gated on CD8+ gate.

Our preliminary data showed that CBA/CaJ and C57BL/6J showed differences in both EMCD4 and EMCD8 compartments with C57BL/6J showing higher EMCD4 and EMCD8 frequencies and counts than CBA/CaJ (Fig 4A - 4D). On the other hand, BALB/cJ and SJL/J differed in EMCD4 but not EMCD8 compartment (Fig 4E - 4H), with SJL/J showing significantly higher EMCD4 frequencies and counts than BALB/cJ (Fig 4E, 4F).

Figure 4: **EMCD4 and EMCD8 frequencies and counts (in millions) in spleen in C57BL/6J versus CBA/CaJ (A to D) and BALB/cJ versus SJL/J (E to H).** Each dot represents data from one individual mouse (n > 25 per group). EMCD4 and EMCD8 frequencies indicate frequencies of EM compartment out of total conventional CD4 (CD4+CD25-) and CD8 T cells respectively. Cell counts are expressed in millions. Box plots indicate median and interquartile ranges. Upper whisker extends till the highest value that is within 1.5 times the interquartile range from 3rd quartile. Lower whisker extends till the lowest value that is within 1.5 times the interquartile range from 1st quartile. The p-values are as indicated in the panels.

To test the source of the differences observed in EMCD4 frequencies between the strains, we prepared bi-parental mixed bone marrow chimeras in which bone marrow cells from both parents, in various ratios, were transferred into irradiated F1 mice. We used heavy irradiation that is reported to successfully engraft parental donor cells in F1 recipients (41–44) and ensured endogenous NK cell depletion post-irradiation (Fig S3 A and B). We also confirmed that chimerism is successful by comparing injected donor cell ratios with ratios of CD45.1 to CD45.2 lymphocytes in reconstituted F1 mice (Fig S3 C and D) and found no statistically significant differences, ruling out the possibility of unequal reconstitution by donor strains. The mean frequency of endogenous cells was 10.1 % in B6.SJL – CBA/CaJ chimera (SD-6.5) and 25.9 % in BALB/cJ - SJL/J chimera (SD-12.3). After waiting two months for generation of memory cells from the donor genotypes, these mice were phenotyped for CD4 and CD8 EM compartments. Memory CD4 and CD8 frequencies of the donor-derived cell population (Fig 5A to 5D) in these chimeras showed similar trends as that in parental strains (Fig 4). These results suggest that genetic factors contribute in cell intrinsic fashion to determining EM T cell frequencies.

Figure 5: Frequencies of EMCD4 and EMCD8 in the spleen from donor partner in mixed bone marrow chimera of B6.SJL-CBA/CaJ pair (A, B) and BALB/cJ-SJL/J pair (C, D). Each dot represents data from one individual mouse (n > 9 per group) and individual mice data are connected by lines. The p-values are as indicated in the panels.

CD4 gene expression in strains that differ in CD4 effector memory phenotype

Since bone marrow chimera data suggested cell-autonomous regulation of memory phenotype, we hypothesized that intrinsic gene expression differences in naive CD4 T cells between the mouse strains could contribute to differences in EMCD4 frequencies that are observed across the strains. We tested this using gene expression data (microarray) of splenic CD4 T cells from multiple mouse strains available from the immunological genome project (33). Immgen gene expression data comes from sorted CD4+ CD62L+ spleen cells ('non-EM' CD4 T cells), which could potentially include CD44-high CMCD4 population in addition to genuine naïve CD4 T cells. Hence, we selected 2 mouse strains (CBA/CaJ and SJL/J) that showed maximum differences in EMCD4 frequencies and counts, without showing differences in CMCD4 levels (Fig S4 A to E, p-values shown in supplementary table S5). CMCD4 frequencies were very low in these 2 strains (Figure 3B and 3D), making comparison of CD4+ CD62L+ gene expression between these strains interpretable. A comparison of gene expression profile of CD4+ CD62L+ve splenocytes (obtained from GEO accession: GSE60337) of CBA/CaJ and SJL/J identified 290 genes that were significantly different (p values < 0.05 after multiple correction) (supplementary table S6). Gene ontology enrichment analysis identified significantly enriched GO processes associated with the 290-gene list (supplementary table S6). As expected, MHC molecule related

361 transcripts were differentially expressed and came up as significantly enriched, reassuring that the
362 analysis we performed could actually pick up biologically relevant differences (table S6). In
363 addition, there was significant enrichment for lipid biosynthesis (11 genes, p value: 1.78E-05,
364 adjusted p value: 0.004) and lipid metabolism (16 genes, p value: 1.13E-04, adjusted p value:
365 0.02) (supplementary file S6), implying that these pathways could be of biological relevance to
366 regulation of CD4 T cell memory phenotype. When data from multiple strains were pooled
367 together, there was a positive correlation between EMCD4 and EMCD8 frequencies (Fig S4F),
368 similar to what is observed for humans (Fig 2B), suggesting similar mechanisms in regulation of
369 CD4 and CD8 effector memory or dependence of one subset on the other.
370
371
372

373 **Discussion**

374

375 There are a number of possible explanations for the observation of a substantial extent of
376 variation in the relative frequencies of memory-phenotype B cells, CD4 T cells and CD8 T cells.
377 Such variations in biological parameters are commonly interpreted as errors in technical or
378 experimental factors, although in real populations, biological variance may well have meaningful
379 interpretative value (31). At the population level, we have used descriptive characterization of the
380 human peripheral blood leucocyte phenotype to examine the possible significance of population
381 level variation in these leucocyte subsets.

382

383 By comparison of intra-individual variance of memory and effector subsets of T and B cells, we
384 find that immune subsets that respond to day-to-day fluctuations, such as the plasmablasts and
385 CD4 TEMRA cells, show similar intra-individual and inter-individual variance in the population,
386 while B cell memory, CD4 memory and CD8 memory cell frequencies are very much stable over
387 a one-year period within individuals. This suggests that memory subsets are not determined by
388 short-term environmental fluctuations, unlike CD4 TEMRA and plasmablasts. Within-individual
389 variation can thus possibly be an explanatory factor in the population-level variation for effector
390 T cells and plasmablast subsets, but not for memory T and B cell subsets. A recent study (45) that
391 quantified technical and biological variation in human immune phenotype also found high intra-
392 individual variation in CD4 TEMRA cells in comparison to other immune lineages. Our data
393 show that CD8 TEMRA cells did not behave similar to CD4 TEMRA cells in that they do not
394 show considerable intra-individual fluctuation. This could be related to indications that CD8
395 TEMRA cells may be functionally different from CD4 TEMRA cells, despite being
396 phenotypically similar (46).

397

Our findings of a lack of correlation between B cell memory and CD4 memory also argue against the possibility of cumulative exposures determining both the memory subsets, although it remains possible that cumulative exposures regulate one but not the other subset. The positive correlation we find between CD4 memory and CD8 memory suggest that T cell memory levels could be regulated by similar mechanisms, distinct from how memory B cell levels are regulated.

We attempted to quantify the degree of similarity in immune phenotype between siblings to see if there was any indication that these memory phenotypes are heritable. Although our sibling study was not as powerful as previously published twin studies (1,2,4,14) for understanding heritability, our sibling data strongly support and extend the interpretations from our serial bleed data. Thus, siblings did not show concordance in effector T and B subsets, suggesting that variation in those subsets is environmentally driven. On the other hand, siblings were far similar to each other in memory subset frequencies than non-siblings were, consistent with a genetic component regulating memory lymphocyte frequencies. However, it must be acknowledged that any similarity between siblings could be attributed to early life influences, since it is well recognized that both intra-uterine (47,48) and neonatal (49) stress impacts immune system development, and a recent study (13) has shown a prominent role for co-habitation in determining immune phenotypes. Our findings are consistent with two previous studies (1,14) which report high degree of heritability for majority of immune subsets. Another study (2), too, reported high degree of heritability for CD4 memory subsets, although it found non-heritable factors to be important for most other immune phenotypes. Since our interpretations are predominantly based on associations, we complemented the study with experimental data from mice.

Although laboratory animals show restricted genetic diversity compared to human populations, we exploited the fact that independent inbred strains are genetically distinct and are grown in a

423 homogeneous environment in a controlled animal facility. Hence, it is worth examining if
424 differences in leucocyte subset levels between strains are genetically driven. We
425 immunophenotyped some common strains of mice and chose two pairs showing substantial
426 differences, SJL/J and BALB/cJ for further experiments on CD4 memory and C57BL/6J and
427 CBA/CaJ for CD8 and CD4 T cell memory phenotype. It is notable that not only did splenic CD4
428 and CD8 memory frequencies show differences, but that total numbers of these cells per organ
429 also showed the same differences, indicating that subset frequencies are reasonable surrogates for
430 the pool size of these subsets. We also have noted that animals maintained in our facility had
431 relatively high frequencies of EM phenotype T cells even at steady state. Even though mice were
432 harbored in a specific-pathogen-free facility, it remains possible that there might be variations in
433 microbial antigenic burden or gut microbiome composition between laboratories, although these
434 factors are difficult to quantify. Long years of reproductive isolation because of inbreeding could
435 also have contributed to these differences.
436
437 Our mixed bone marrow chimera experiments allowed us to examine whether the CD4 and/or
438 CD8 memory levels were genetically determined and whether these genetic influences were T
439 cell-intrinsic in nature. Our data suggest that donor genotype-specific cell-intrinsic factors
440 strongly influence both CD4 and CD8 memory T cell pool size.
441
442 Using gene expression data of mouse splenic CD4+ve CD62L+ve T cells available in the public
443 domain (33) we attempted to characterize genetic differences that could explain differences in
444 CD4 memory phenotype. There are a number of limitations in our approach. Firstly, we do not
445 evaluate gene expression differences in the same mice (or even mice from the same small animal
446 facility) that we experimentally find phenotypic differences for, and differences in phenotype that
447 we observe could be influenced by additional facility-specific environmental factors as well.

448 Secondly, the array based gene expression data available in the public domain that we used (33)
449 contain only 2 replicates for each strain (except for C57BL/6J), limiting statistical power and
450 increasing the chances of picking up false positive differences. Thirdly, the "naive" population as
451 defined by Immgen is based on CD4+ve CD62L+ve gate, and does not exclude CD44+ cells,
452 potentially including central memory cells into the population. Our analysis attempts to overcome
453 this limitation by comparing those strains which differ only in EMCD4, and not CMCD4 subsets.
454 In spite of these caveats, our analysis is still useful as a preliminary exploration that can be
455 extended in further experiments using multiple replicates and strains. Remarkably, lipid/fatty acid
456 metabolism pathways that came up as significant in as limited an analysis as this, was recently
457 suggested to play crucial roles in T cell activation (50). This exploratory analysis is thus likely to
458 provide further clues to direct future mechanistic studies.
459
460 Thus, our data suggest genetic and cell-intrinsic factors as major determinants of the memory
461 pool size of T cells as well as, probably, B cells. However, in a natural ecosystem where animals
462 are exposed to substantial and continual antigenic exposures, the balance between cell-intrinsic
463 genetic effects on memory and the effects of environmental factors is likely to be nuanced and
464 quantitative. Further studies using varying antigenic burdens and experimental immunizations
465 will be necessary to address these issues. Mechanistically, it will be interesting to explore
466 whether genetic cell-intrinsic factors affect memory generation during immune response or
467 memory attrition after the peak of immune response. Our data provide interesting insights and
468 directions for future work in understanding the genesis and consequences of the regulation of
469 niche size of lymphocyte memory.

Acknowledgements:

We acknowledge Department of Biotechnology, Government of India and Department of Science and Technology, Government of India for funding the study.

Funding:

The study was supported in part by grants from the Department of Biotechnology (to A.G. # BT/PR12849/MED/15/35/2009; to V.B. # BT/PR14420/Med/29/213/2010; to S.R. # BT/PR-14592/BRB/10/858/2010; to S.B. BT/MB/01/THSTI-ChBC/2009; and to U.C.M.N. # BT/PR14723/MED/15/44/2010), and from the Department of Science and Technology, Government of India (to V.B. # SR/SO/HS-0005/2011 and #EMR/2015/001074; to S.R. # SB/SO/HS/210/2013). The National Institute of Immunology and the Translational Health Science and Technology Institute are supported by the Department of Biotechnology, Government of India.

Conflict of interest: The authors declare no conflict of interest.

Key words: memory, T-cell, B-cell

List of abbreviations:

EM: Effector memory

TEMRA: T, effector memory, RA positive

CD: Cluster of differentiation

References

1. Orrù V, Steri M, Sole G, Sidore C, Virdis F, Dei M, et al. Genetic variants regulating immune cell levels in health and disease. Cell. 2013 Sep 26;155(1):242–56.

2. Brodin P, Jojic V, Gao T, Bhattacharya S, Angel CJL, Furman D, et al. Variation in the human immune system is largely driven by non-heritable influences. Cell. 2015 Jan 15;160(1–2):37–47.

3. Lu Y, Biancotto A, Cheung F, Remmers E, Shah N, McCoy JP, et al. Systematic Analysis of Cell-to-Cell Expression Variation of T Lymphocytes in a Human Cohort Identifies Aging and Genetic Associations. Immunity. 2016 Nov 15;45(5):1162–75.

4. de Craen AJM, Posthuma D, Remarque EJ, van den Biggelaar AHJ, Westendorp RGJ, Boomsma DI. Heritability estimates of innate immunity: an extended twin study. Genes Immun. 2005 Mar;6(2):167–70.

5. Li Y, Oosting M, Deelen P, Ricaño-Ponce I, Smeekens S, Jaeger M, et al. Inter-individual variability and genetic influences on cytokine responses to bacteria and fungi. Nat Med. 2016 Aug;22(8):952–60.

6. Ye CJ, Feng T, Kwon H-K, Raj T, Wilson MT, Asinovski N, et al. Intersection of population variation and autoimmunity genetics in human T cell activation. Science. 2014 Sep 12;345(6202):1254665.

7. Li Y, Oosting M, Smeekens SP, Jaeger M, Aguirre-Gamboa R, Le KTT, et al. A Functional Genomics Approach to Understand Variation in Cytokine Production in Humans. Cell. 2016 Nov 3;167(4):1099-1110.e14.

8. Horst R ter, Jaeger M, Smeekens SP, Oosting M, Swertz MA, Li Y, et al. Host and Environmental Factors Influencing Individual Human Cytokine Responses. Cell. 2016 Nov 3;167(4):1111-1124.e13.

9. Lee YC, Newport MJ, Goetghebuer T, Siegrist CA, Weiss HA, Pollard AJ, et al. Influence of genetic and environmental factors on the immunogenicity of Hib vaccine in Gambian twins. Vaccine. 2006 Jun 19;24(25):5335–40.

10. Yan K, Cai W, Cao F, Sun H, Chen S, Xu R, et al. Genetic effects have a dominant role on poor responses to infant vaccination to hepatitis B virus. J Hum Genet. 2013 May;58(5):293–7.

11. Rubelt F, Bolen CR, McGuire HM, Vander Heiden JA, Gadala-Maria D, Levin M, et al. Individual heritable differences result in unique cell lymphocyte receptor repertoires of naïve and antigen-experienced cells. Nat Commun. 2016;7:11112.

12. Majumder PP, Sarkar-Roy N, Staats H, Ramamurthy T, Maiti S, Chowdhury G, et al. Genomic correlates of variability in immune response to an oral cholera vaccine. Eur J Hum Genet. 2013 Sep;21(9):1000–6.

13. Carr EJ, Dooley J, Garcia-Perez JE, Lagou V, Lee JC, Wouters C, et al. The cellular composition of the human immune system is shaped by age and cohabitation. Nat Immunol. 2016 Apr;17(4):461–8.

14. Kaczorowski KJ, Shekhar K, Nkulikiyimfura D, Dekker CL, Maecker H, Davis MM, et al. Continuous immunotypes describe human immune variation and predict diverse responses. PNAS. 2017 Jul 10;201705065.

15. Farber DL, Netea MG, Radbruch A, Rajewsky K, Zinkernagel RM. Immunological memory: lessons from the past and a look to the future. Nat Rev Immunol. 2016 Feb;16(2):124–8.

16. Karlsson EK, Kwiatkowski DP, Sabeti PC. Natural selection and infectious disease in human populations. Nat Rev Genet. 2014 Jun;15(6):379–93.

17. Boehm T. Evolution of vertebrate immunity. Curr Biol. 2012 Sep 11;22(17):R722-732.

18. Cicin-Sain L, Smyk-Pearson S, Smyk-Paerson S, Currier N, Byrd L, Koudelka C, et al. Loss of naive T cells and repertoire constriction predict poor response to vaccination in old primates. J Immunol. 2010 Jun 15;184(12):6739–45.

19. Huster KM, Stemberger C, Gasteiger G, Kastenmüller W, Drexler I, Busch DH. Cutting edge: memory CD8 T cell compartment grows in size with immunological experience but nevertheless can lose function. J Immunol. 2009 Dec 1;183(11):6898–902.

20. Boots M, Bowers RG. The evolution of resistance through costly acquired immunity. Proc Biol Sci. 2004 Apr 7;271(1540):715–23.

21. Graw F, Magnus C, Regoes RR. Theoretical analysis of the evolution of immune memory. BMC Evol Biol. 2010;10:380.

22. Wodarz D. Evolution of immunological memory and the regulation of competition between pathogens. Curr Biol. 2003 Sep 16;13(18):1648–52.

23. Antia R, Pilyugin SS, Ahmed R. Models of immune memory: on the role of cross-reactive stimulation, competition, and homeostasis in maintaining immune memory. Proc Natl Acad Sci USA. 1998 Dec 8;95(25):14926–31.

24. Jiang J, Lau LL, Shen H. Selective depletion of nonspecific T cells during the early stage of immune responses to infection. J Immunol. 2003 Oct 15;171(8):4352–8.

25. Gray D. A role for antigen in the maintenance of immunological memory. Nat Rev Immunol. 2002 Jan;2(1):60–5.

26. Vezys V, Yates A, Casey KA, Lanier G, Ahmed R, Antia R, et al. Memory CD8 T-cell compartment grows in size with immunological experience. Nature. 2009 Jan 8;457(7226):196–9.

27. Capece T, Kim M. The Role of Lymphatic Niches in T Cell Differentiation. Mol Cells. 2016 Jul;39(7):515–23.

28. Mueller SN, Mackay LK. Tissue-resident memory T cells: local specialists in immune defence. Nat Rev Immunol. 2016 Feb;16(2):79–89.

29. Junt T, Scandella E, Ludewig B. Form follows function: lymphoid tissue microarchitecture in antimicrobial immune defence. Nat Rev Immunol. 2008 Oct;8(10):764–75.

30. Mahnke YD, Brodie TM, Sallusto F, Roederer M, Lugli E. The who's who of T-cell differentiation: Human memory T-cell subsets. Eur J Immunol. 2013 Nov 1;43(11):2797–809.

31. Prabhu SB, Rathore DK, Nair D, Chaudhary A, Raza S, Kanodia P, et al. Comparison of Human Neonatal and Adult Blood Leukocyte Subset Composition Phenotypes. PLoS ONE. 2016;11(9):e0162242.

32. Maecker HT, McCoy JP, Nussenblatt R. Standardizing immunophenotyping for the Human Immunology Project. Nat Rev Immunol. 2012 Feb 17;12(3):191–200.

33. Mostafavi S, Ortiz-Lopez A, Bogue MA, Hattori K, Pop C, Koller D, et al. Variation and genetic control of gene expression in primary immunocytes across inbred mouse strains. J Immunol. 2014 Nov 1;193(9):4485–96.

34. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009;10:48.

35. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44–57.

36. Bergmann B, Grimsholm O, Thorarinsdottir K, Ren W, Jirholt P, Gjertsson I, et al. Memory B cells in mouse models. Scand J Immunol. 2013 Aug;78(2):149–56.

37. Zuccarino-Catania GV, Sadanand S, Weisel FJ, Tomayko MM, Meng H, Kleinstein SH, et al. CD80 and PD-L2 define functionally distinct memory B cell subsets that are independent of antibody isotype. Nat Immunol. 2014 Jul;15(7):631–7.

38. Lai D, Zhu J, Wang T, Hu-Li J, Terabe M, Berzofsky JA, et al. KLF13 sustains thymic memory-like CD8+ T cells in BALB/c mice by regulating IL-4–generating invariant natural killer T cells. J Exp Med. 2011 May 9;208(5):1093–103.

39. Chiu B-C, Martin BE, Stolberg VR, Chensue SW. Cutting edge: Central memory CD8 T cells in aged mice are virtual memory cells. J Immunol. 2013 Dec 15;191(12):5793–6.

40. Li S, Xie Q, Zeng Y, Zou C, Liu X, Wu S, et al. A naturally occurring $CD8^{+}CD122^{+}$ T-cell subset as a memory-like Treg family. Cellular & Molecular Immunology. 2014 Jul;11(4):326.

41. Sprent J, Boehmer HV, Nabholz M. Association of immunity and tolerance to host H-2 determinants in irradiated F1 hybrid mice reconstituted with bone marrow cells from one parental strain. J Exp Med. 1975 Aug 1;142(2):321–31.

599 42. Gao EK, Lo D, Sprent J. Strong T cell tolerance in parent----F1 bone marrow chimeras
600 prepared with supralethal irradiation. Evidence for clonal deletion and anergy. J Exp Med.
601 1990 Apr 1;171(4):1101–21.

602 43. Sim B-C, Aftahi N, Reilly C, Bogen B, Schwartz RH, Gascoigne NRJ, et al. Thymic
603 skewing of the CD4/CD8 ratio maps with the T-cell receptor α-chain locus. Current
604 Biology. 1998 Jun 4;8(12):701-S3.

605 44. Zhao Y, Ohdan H, Manilay JO, Sykes M. NK cell tolerance in mixed allogeneic chimeras. J
606 Immunol. 2003 Jun 1;170(11):5398–405.

607 45. Burel JG, Qian Y, Arlehamn CL, Weiskopf D, Zapardiel-Gonzalo J, Taplitz R, et al. An
608 Integrated Workflow To Assess Technical and Biological Variability of Cell Population
609 Frequencies in Human Peripheral Blood by Flow Cytometry. The Journal of Immunology.
610 2017 Jan 9;1601750.

611 46. Sathaliyawala T, Kubota M, Yudanin N, Turner D, Camp P, Thome JJC, et al. Distribution
612 and compartmentalization of human circulating and tissue-resident memory T cell subsets.
613 Immunity. 2013 Jan 24;38(1):187–97.

614 47. Rathore DK, Nair D, Raza S, Saini S, Singh R, Kumar A, et al. Underweight full-term
615 Indian neonates show differences in umbilical cord blood leukocyte phenotype: a cross-
616 sectional study. PLoS ONE. 2015;10(4):e0123589.

617 48. Schaub B, Liu J, Höppler S, Schleich I, Huehn J, Olek S, et al. Maternal farm exposure
618 modulates neonatal immune mechanisms through regulatory T cells. J Allergy Clin
619 Immunol. 2009 Apr;123(4):774-782.e5.

620 49. Paparo L, di Costanzo M, di Scala C, Cosenza L, Leone L, Nocerino R, et al. The influence
621 of early life nutrition on epigenetic regulatory mechanisms of the immune system.
622 Nutrients. 2014 Nov;6(11):4706–19.

623 50. Angela M, Endo Y, Asou HK, Yamamoto T, Tumes DJ, Tokuyama H, et al. Fatty acid
624 metabolic reprogramming via mTOR-mediated inductions of PPARγ directs early activation
625 of T cells. Nature Communications. 2016 Nov 30;7:ncomms13683.

626

627

628 **Supporting information:**

629 **Figure S1: Gating strategy of human B and T memory subsets**

630 **Figure S2: Intra-individual variance in immune subsets across one year (4 time points).**

631 Each dot indicates frequency of immune subset at a time point in an individual donor. Individual

632 donors are connected by lines. (n=45; only 10 donors shown for the sake of clarity).

633 Quantification and p-values shown in figure 1.

634 **Figure S3**: **A and B:** NK cell depletion in chimera recipients post-irradiation. Dots represent data

635 from individual mice. **C and D:** Reconstitution efficiency in B6.SJL-CBA/CaJ chimera (C) and

636 BALB/cJ-SJL/J chimera (D). Y-axis indicates ratio of injected CD45.1/CD452 cells to

637 reconstituted CD45.1/CD452 cells for the cell subset indicated. Dots indicate mean and error bars

638 indicate 95% confidence interval. P-values are as indicated.

639 **Figure S4**: **Quantification of EMCD4 and CMCD4 frequencies and counts in multiple**

640 **strains of mice examined** (A to D) (Statistical quantification for these comparisons are tabulated

641 in supplementary table S5). Boxplots show median and interquartile range. Upper whisker

642 extends till the highest value that is within 1.5 times the interquartile range from 3rd quartile.

643 Lower whisker extends till the lowest value that is within 1.5 times the interquartile range from

644 1st quartile. Outliers are shown as dots. Each group consisted of > 10 mice. Cell counts are

645 shown as number in millions.

646 E: mean EMCD4 and CMCD4 of each strain plotted together. CBA/CaJ and SJL/J show

647 differences in EMCD4 frequency (x-axis), but not in CMCD4 frequency (y-axis).

648 F- Correlations between EMCD4 and EMCD8 frequencies with mice from all strains pooled

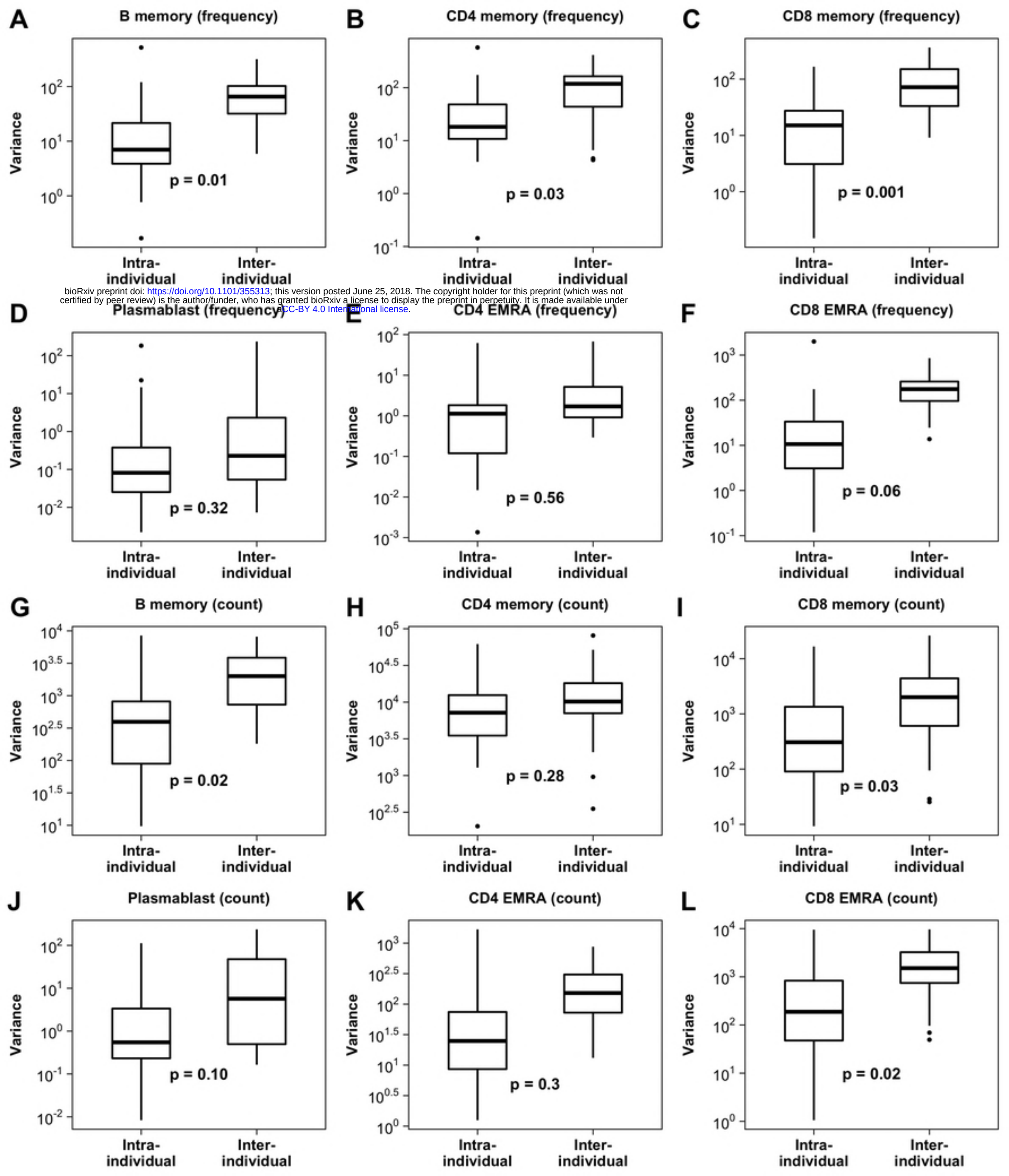
649 together. Each dot represents a mouse (n=92). Correlation coefficient (spearman) and p-values

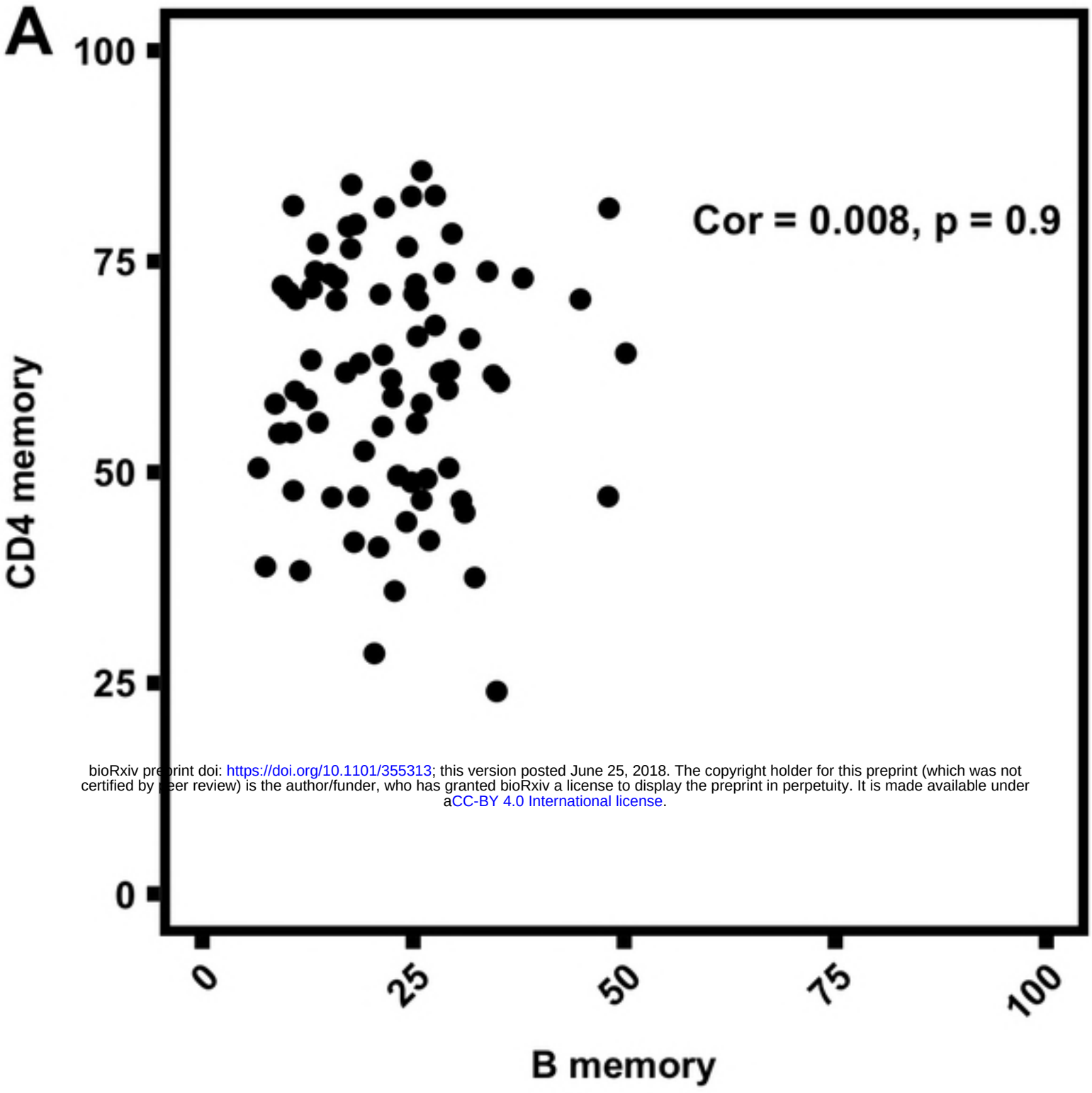
650 are indicated.

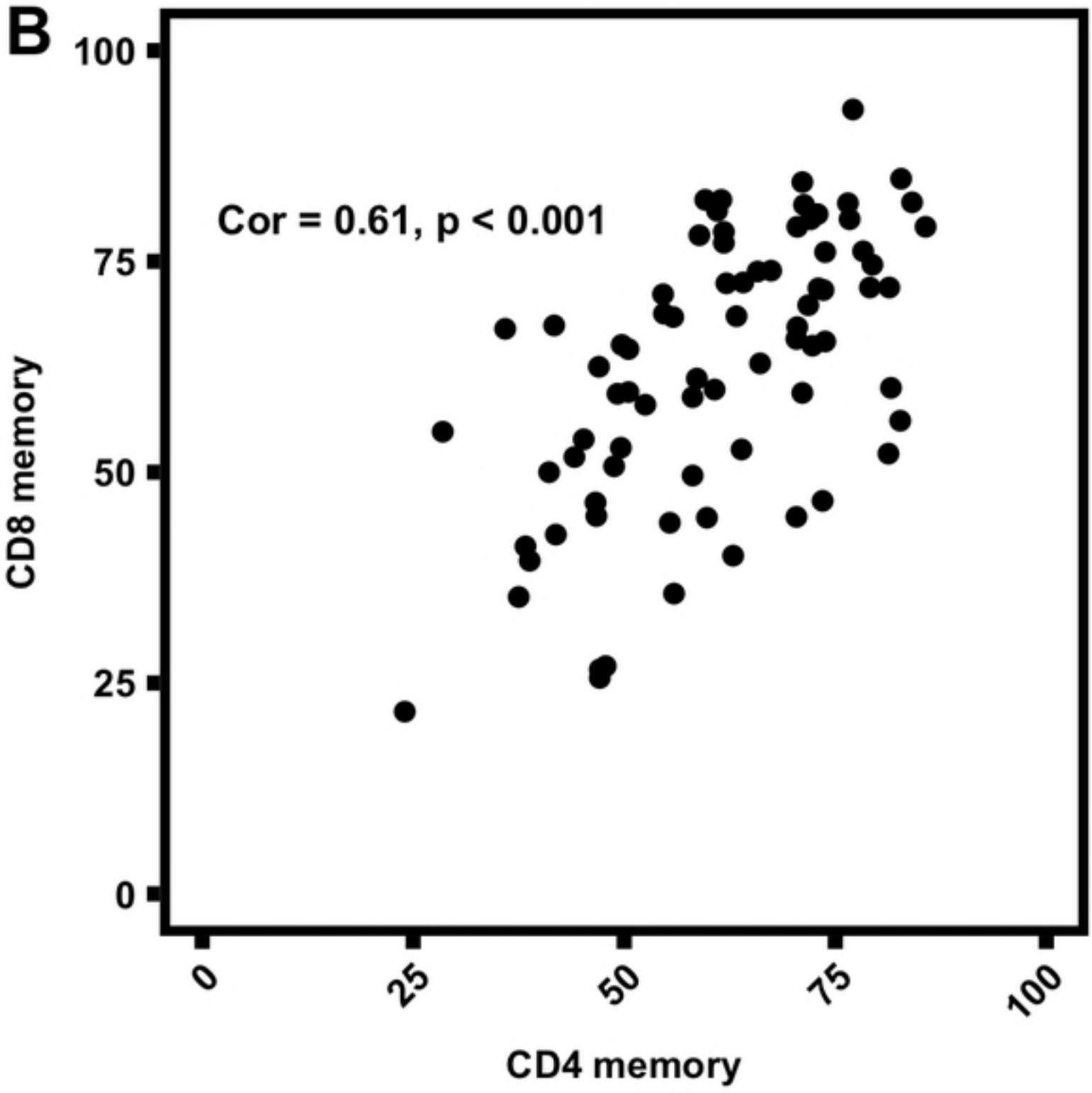
651 **Table S5**: **Statistical analysis of data shown in Figure S4**

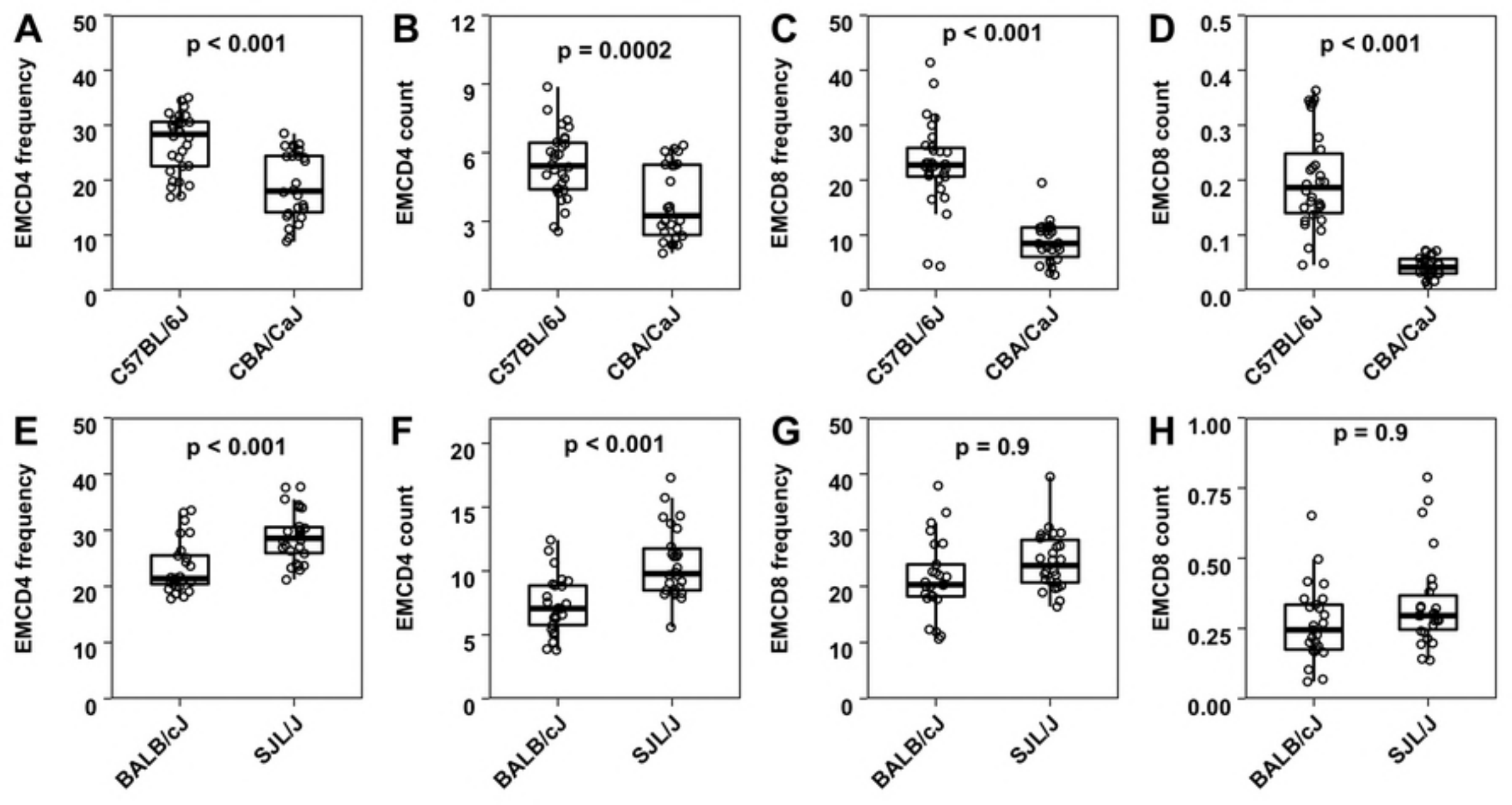
652 **Table S6: List of differentially expressed genes in naïve CD4 T cells between SJL/J and**

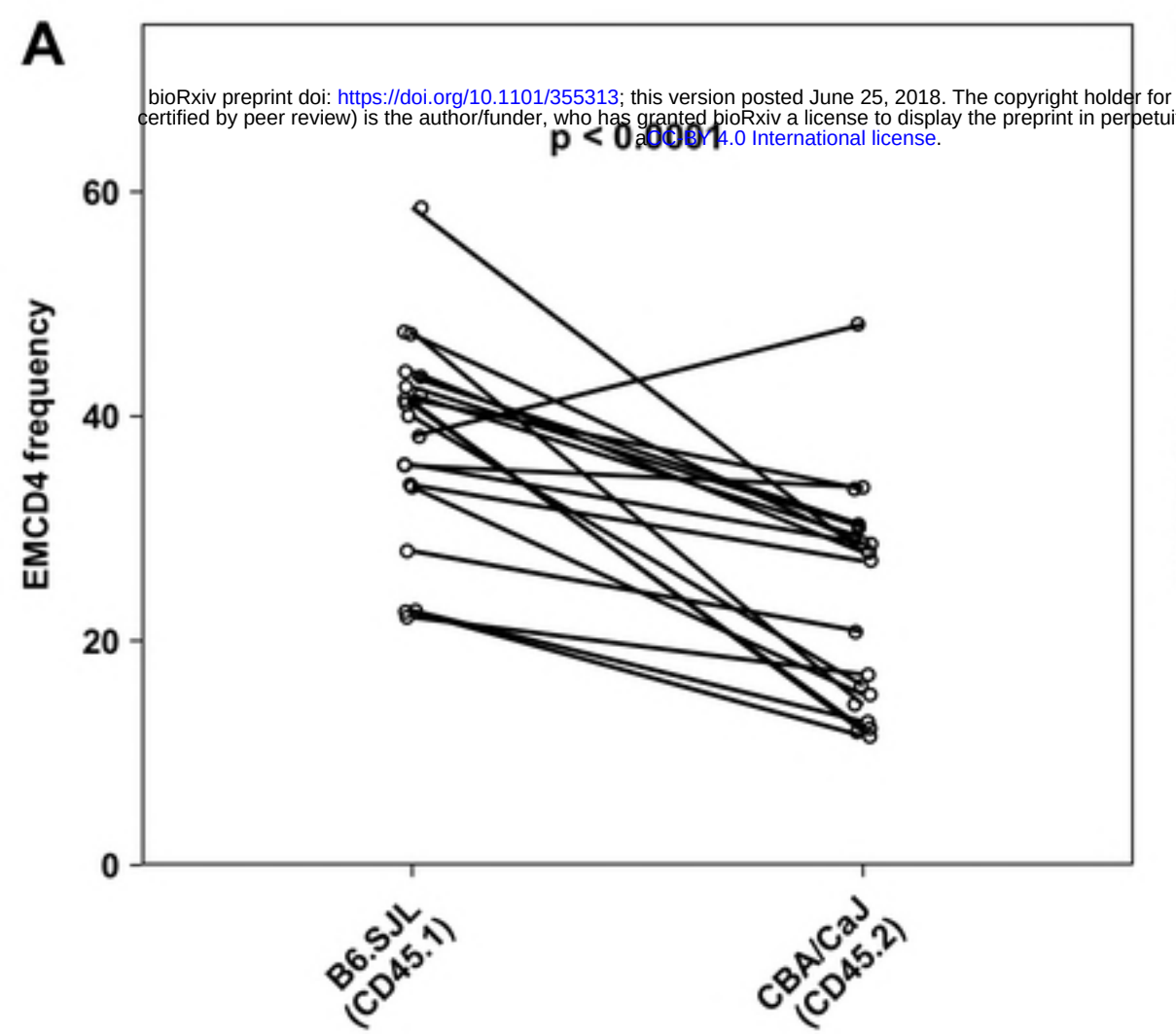
653 **CBA/CaJ and gene enrichment analysis results**

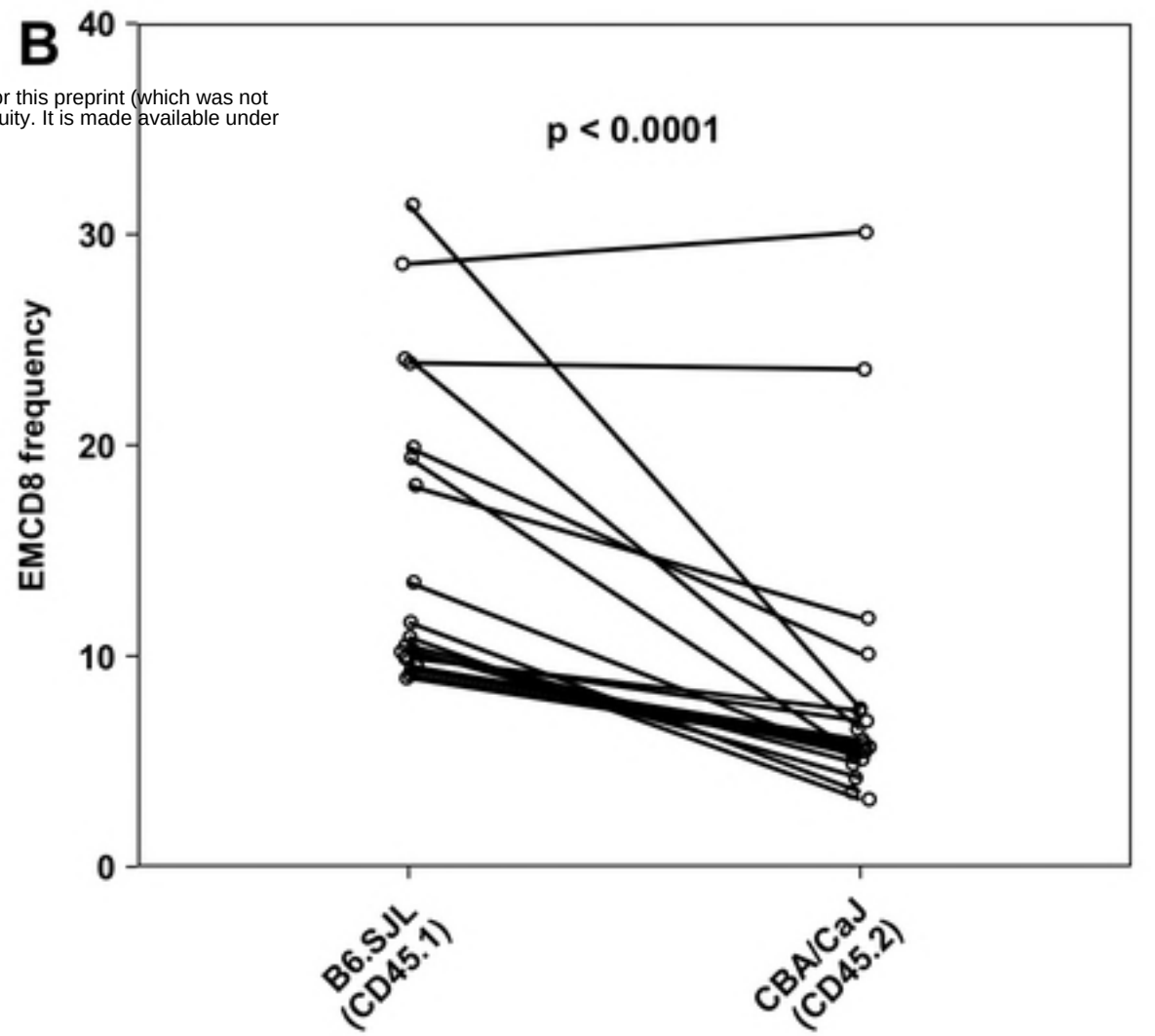


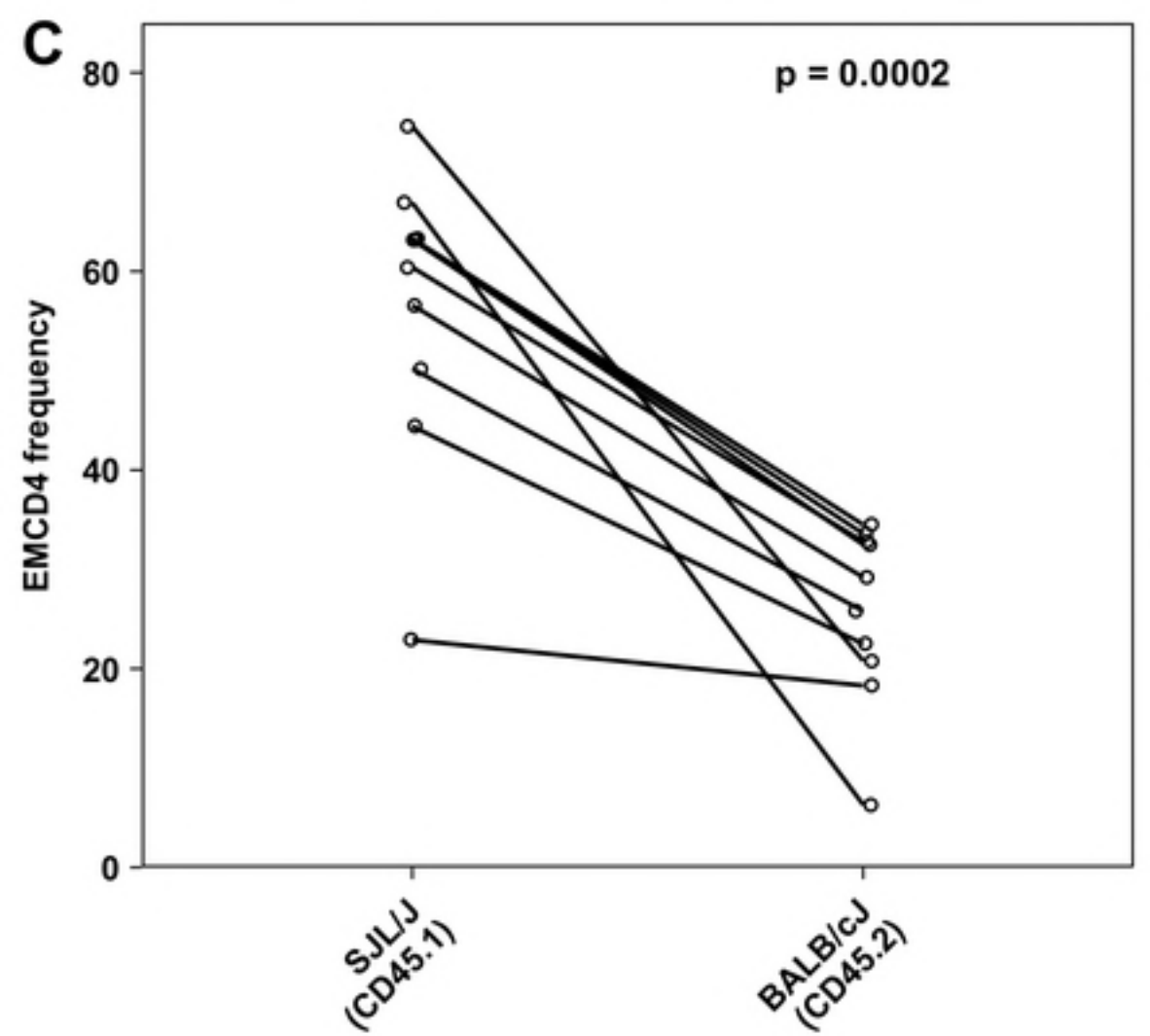


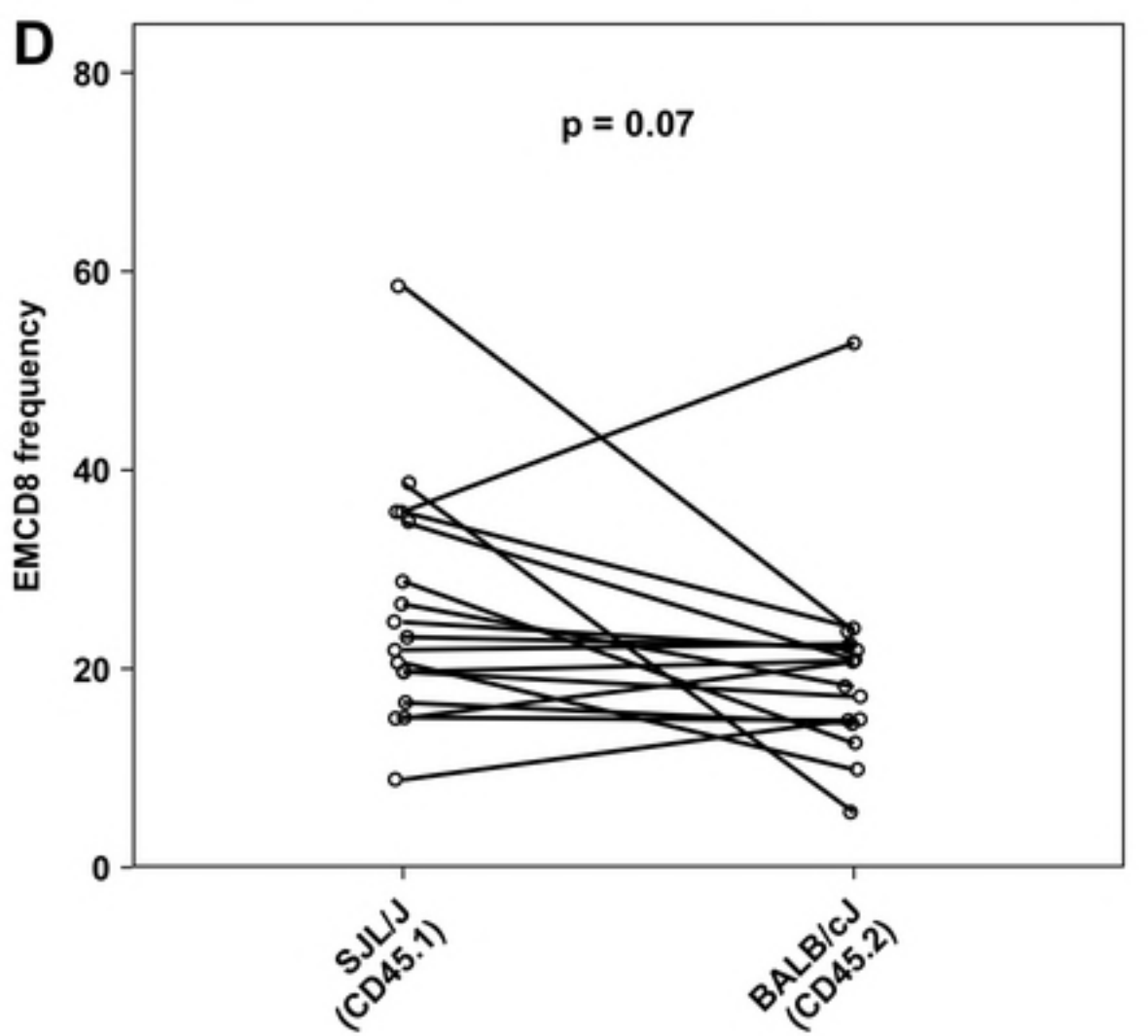



A
EMCD4 frequency
p < 0.001
C57BL/6J
CBA/CaJ
B
EMCD4 count
p = 0.0002
C57BL/6J
CBA/CaJ
C
EMCD8 frequency
p < 0.001
C57BL/6J
CBA/CaJ
D
EMCD8 count
p < 0.001
C57BL/6J
CBA/CaJ
E
EMCD4 frequency
p < 0.001
BALB/cJ
SJL/J
F
EMCD4 count
p < 0.001
BALB/cJ
SJL/J
G
EMCD8 frequency
p = 0.9
BALB/cJ
SJL/J
H
EMCD8 count
p = 0.9
BALB/cJ
SJL/J





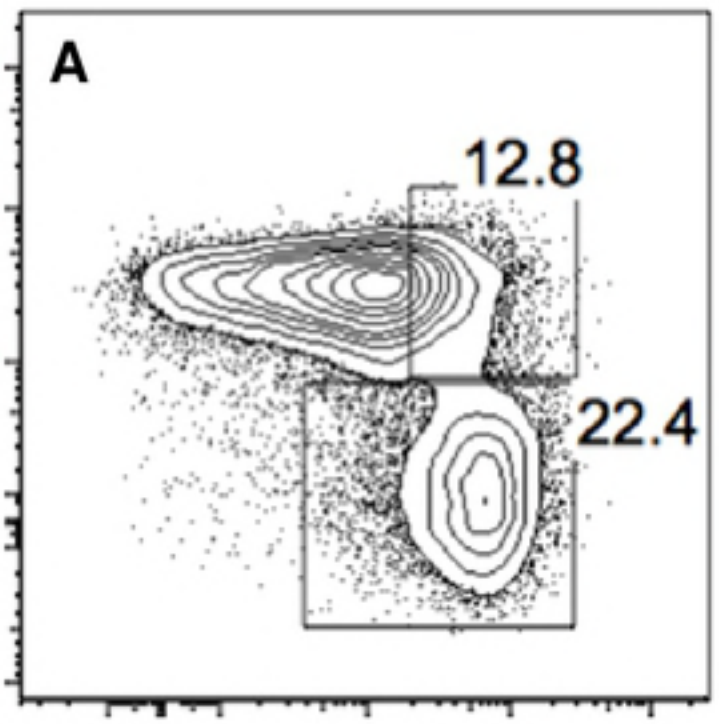


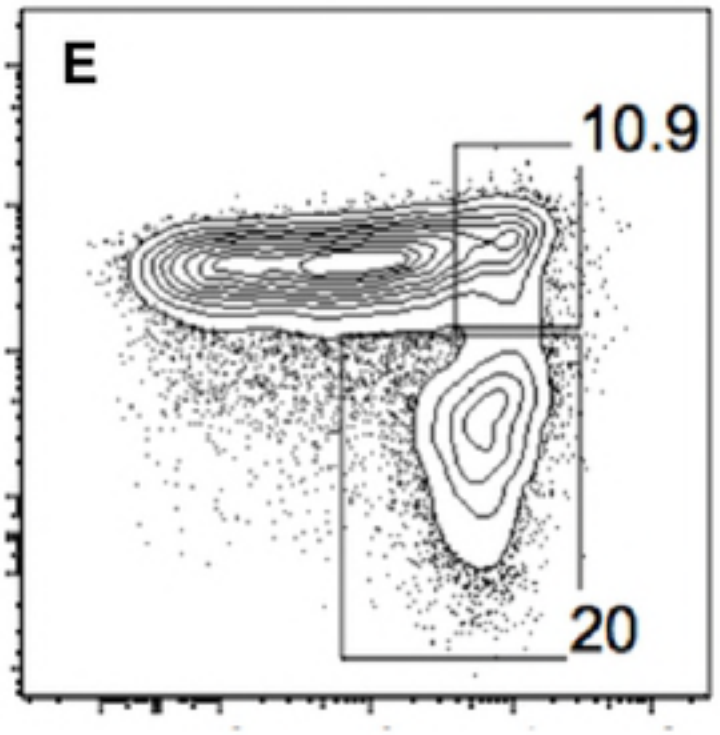


CD4

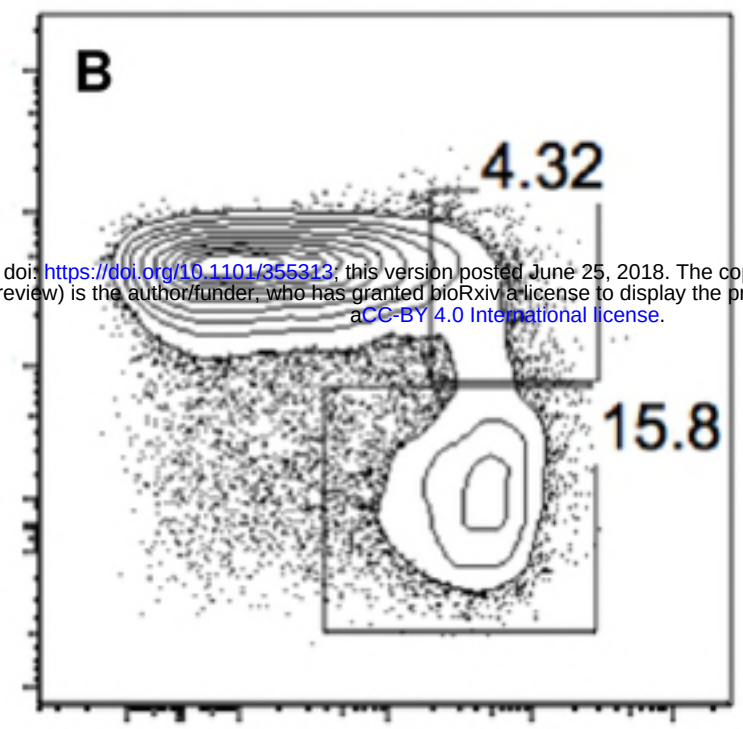
CD8

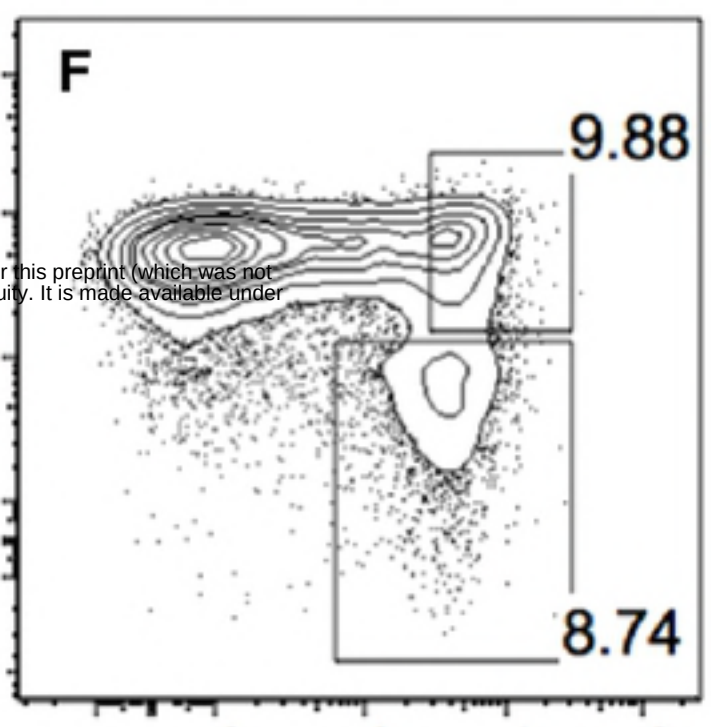
C57BL/6J



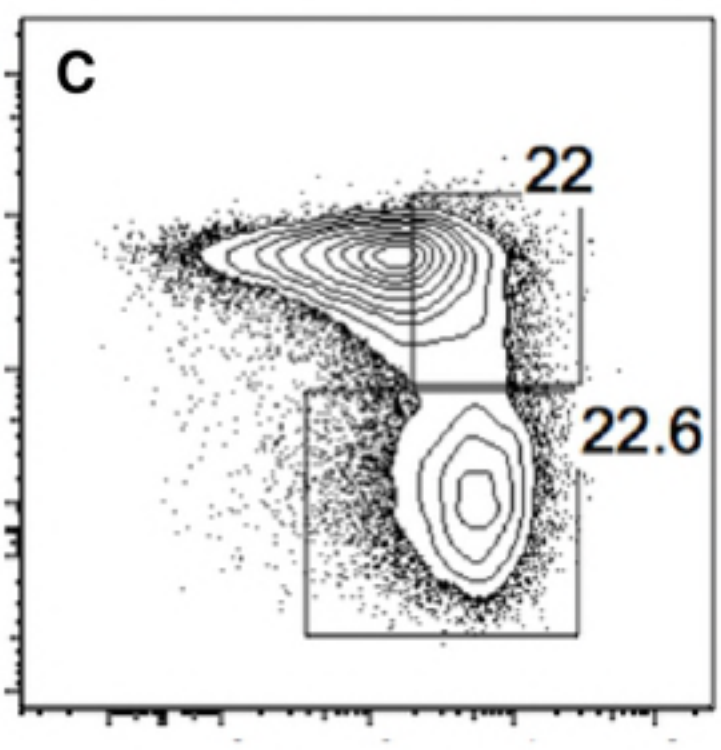


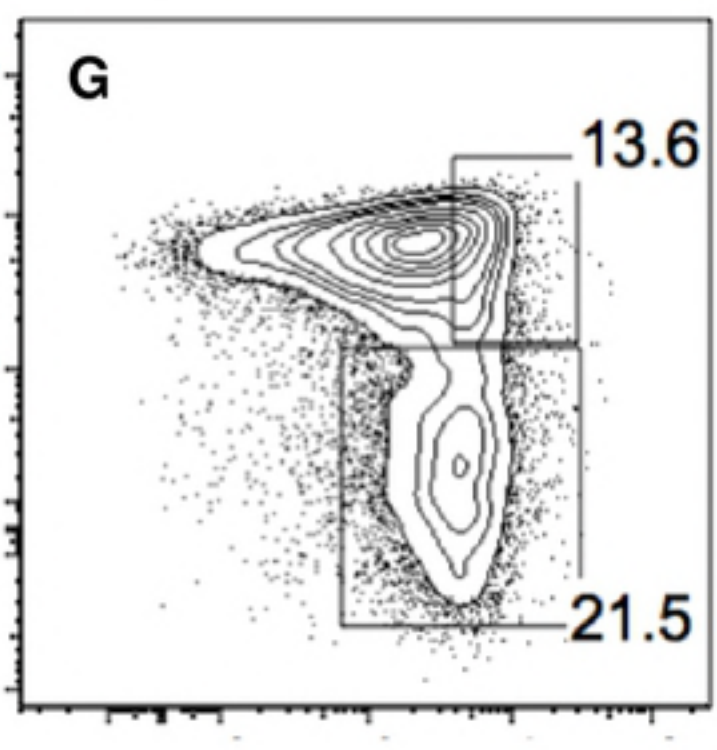
CBA/CaJ



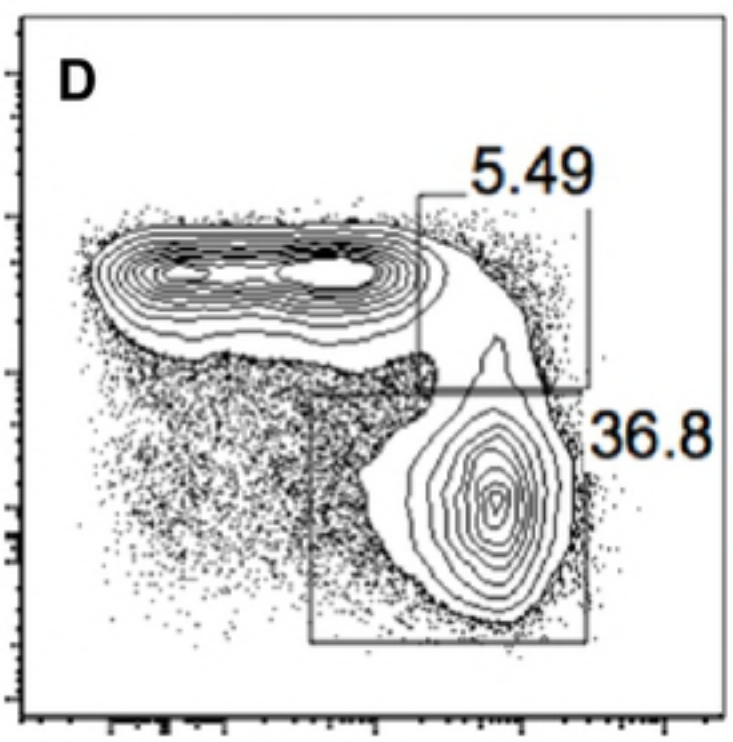


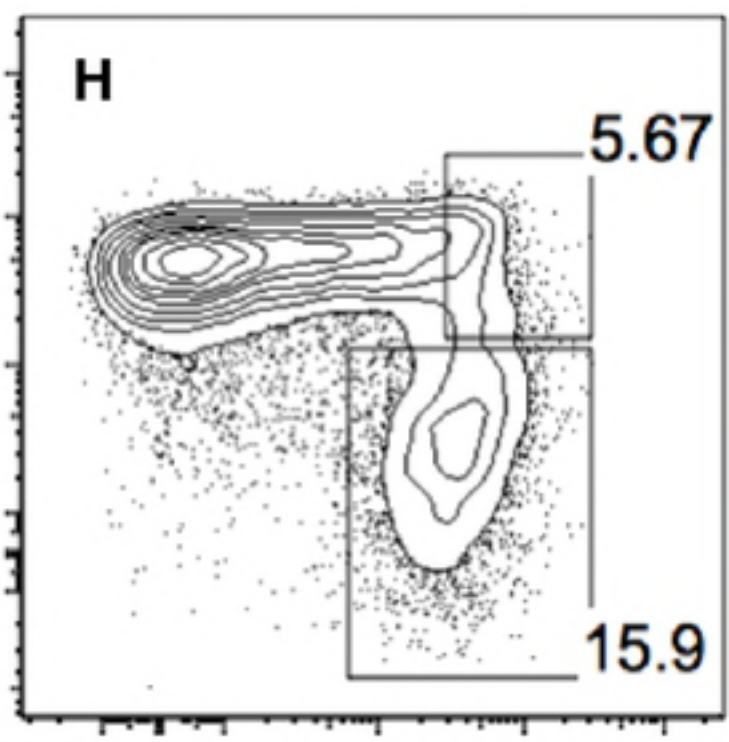
BALB/cJ





SJL/J





CD62L

CD44