THE UNIVERSITY OF CHICAGO

THE T CELL RECEPTOR GUIDES DEVELOPMENT, POLARIZATION, AND PERIPHERAL FUNCTION OF γδ T CELLS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

INTERDISCIPLINARY SCIENTIST TRAINING PROGRAM: IMMUNOLOGY

BY

MAYURI VISWANATHAN

CHICAGO, ILLINOIS

AUGUST 2024

என் அம்மாவுக்கும் அப்பாவுக்கும்

Contents

List of Figures

Acknowledgments

Beginning my PhD studies during the COVID-19 pandemic was an experience with no playbook—nothing like this had ever happened before, and everybody was improvising and adapting to the circumstances as well as we could. While shift work and capacity restrictions made for a bumpy start and isolating circumstances, the kindness, generosity, and good spirits of my fellow labmates kept me coming in every day and I am so thankful for their consistent warmth and brilliance. I specifically want to thank my mentor, Erin Adams, for being a strong advocate for me from early in my training, and Caitlin Castro and Augusta Broughton, who were essential to all of the work I did during my PhD and taught me nearly everything I know. I also want to thank Amrita Ramesh, who is a fount of wisdom and my dear friend for life, and Sean Ryan, whose consistent friendship and insistence on the importance of having a whole life inside and outside the lab has helped me maintain perspective throughout the ups and downs of my PhD. Finally, my most sincere thanks to all of my mentors and collaborators, both past and present, for their generosity, time, and belief in me. I would not be here without every single one of them.

I have been fortunate to have the most incredible support system that has only grown with each new challenge in my life. To my friends, who keep me balanced and joyful, thank you for making my life so full. To my sister, who understands me like nobody else in the world can, thank you for being there for every tear, every triumph, and every interminable fit of giggles. To my partner Roddy, you are my port in the storm and coming home to you is the best part of every day. And finally, to my parents, who have always unwaveringly believed in me, and inspired curiosity, tenacity, and a fierce sense of purpose in me from the moment I was born, thank you for always showing me that with passion and perseverance, anything is possible.

Abstract

γδ T cells occupy a unique immunological niche, bridging innate and adaptive immune function within a single cellular compartment. The lineage-defining γδ T cell receptor undergoes VDJ recombination in the thymus, but can recognize self-ligands, bind to ligands using germline encoded residues, and recognize antigens outside the context of MHC molecules. This dissertation addresses the still unknown role of the γδ TCR in postnatal thymic Vδ1+ T cell development, and shows that variation in the constant region of the γδ TCR durably influences thymic programming and peripheral function of γδ T cells. These roles for the γδ TCR in T cell function are distinct from known programs of αβ T cell development and function, and provide novel insights into the role of the γδ TCR in programming T cells in health and disease.

Chapter 1: Introduction

γδ T cells occupy a unique niche in the human immune system, bridging innate and adaptive function within a single cellular compartment[1]. These cells cover a diverse array of functions by expressing a variety of receptors, including Toll-like receptors and NK receptors along with their T cell receptor, allowing them to respond to invariant signals of tissue stress and infection[2]. Their ability to integrate the signals of various receptors to respond in aggregate to the environment uniquely poises them to monitor tissue stress and perform functions specific to tissue and disease contexts.

γδ T cells are stereotyped, as their name implies, by their expression of the γδ T cell receptor, composed of γ and δ chains rather than the conventional α and β chains. The γδ TCR has been shown to play a functionally important role in developmental cell fate, infection, and cancer, with several studies showing evidence of clonal expansion and memory differentiation in disease contexts[3-5]. In the thymus, signal strength appears to inform effector cell fate, although these mechanisms are less well understood in humans[6,7]. Human γδ T cells are distinct from mouse γδs in several critical ways, including variable gene usage, antigen recognition, and functional polarization, which has made understanding the unique characteristics of these cells in humans challenging[8].

γδ T cell receptor structure and usage

Like αβ T cells, γδ T cells generate heterodimeric T cell receptors by the process of somatic V(D)J recombination. During VDJ recombination, thymocytes rearrange D (diversity) segments to J (junctional) segments, then pair V (variable) segments to the recombined DJ to generate a full VDJ variable segment. α and γ chains lack D segments to rearrange, so in those cases, the

variable part of the TCR is composed of only V and J segments. Additional diversity is generated at the junctions of the genes by proofreading and random nucleotide addition, meaning 2 T cell receptors encoded by the same genes could have substantial differences in their variable regions. The amino acids encoded by these random junctional nucleotides are often referred to as "non-templated" residues. The VDJ region is then spliced after transcription to a constant region to form a complete T cell receptor. The V(D)J regions of each chain, when paired together into the TCR heterodimer, encode the variable part of the T cell receptor, which often contacts antigens[8]. In both human and mouse, the δ locus is embedded between the V genes and J genes of the alpha locus, meaning V-J recombination of alpha gene segments requires excision of the δ locus. For this reason, δ recombination must precede alpha recombination. Delta, like beta, contains V, D, and J segments, but unlike beta , can incorporate multiple D genes into a single T cell receptor, leading to higher variability in length and diversity in the CDR3 loop[8]. This may be one way in which γδ T cells compensate for their relatively fewer V genes to choose from, which otherwise limits their potential combinatorial diversity in the variable region.

Mouse γδ T cells are classified by their Vγ gene usage, which determines when they develop in the thymus, which tissues they reside in, and in some cases, which antigens they recognize[9]. In contrast, human γδ T cells are classified by their δ chain usage, and while Vδ3+ cells do exist, the human γδ repertoire is largely composed of Vδ2+ and Vδ1+ cells. Most peripheral blood γδ T cells express a Vγ9Vδ2 T cell receptor, but at mucosal sites, Vδ1+ cells can dominate the repertoire[2].

γδ T cell receptor ligands

Identifying and structurally and functionally validating ligands of γδ T cells in humans has proved very challenging, largely because they do not follow most of the rules of αβ T cell ligand recognition. Unlike αβ T cells, γδ T cells do not always require an MHC molecule to present antigens[8]. This opens up the universe of possible ligands to include cell surface bound and secreted proteins as well as nonpeptidic antigens. Additionally, mouse and human γδ T cells differ in γ-δ chain pairing, so many of the ligands identified in each species lack correlates in the other. Finally, unlike most αβ T cells, many of the suggested ligands for γδ T cells are self-ligands, suggesting the γδ compartment may have different mechanisms for maintaining tolerance with autoreactive TCRs which also breaks the "rules'' of how we understand typical αβ TCR ligand recognition.

Butyrophilins and butyrophilin-like molecules

Despite these challenges, several ligands have been suggested for both Vδ1+ and Vδ2+ subsets, including both MHC-like and non-MHC ligands. Perhaps the best characterized class of ligands is the butyrophilin (BTN) or butyrophilin-like (BTNL) group of molecules, which are non-MHC cell surface proteins that belong to the B7 class of immune co-regulators. BTN2A and BTN3A molecules transduce increases in intracellular phosphoantigen concentration, which are often indicative of infection or cancer, to Vγ9Vδ2 T cells. Vγ9Vδ2 T cells bind directly to BTN2A1, but activate through a mechanism that is still not fully understood[8]. This interaction is notable as it is mediated entirely by germline-encoded residues on the TCR[10], not non-templated residues in the CDR3 region, meaning that V gene usage is sufficient to confer BTN2A1 ligand specificity to the Vγ9Vδ2 subset. Similarly, germline residues in the CDR1, CDR2 and HV4 loops of Vγ4 directly bind the BTNL3/8 heterodimer leading to T cell activation[11]. This heterodimer is highly

expressed in the gut epithelium, and disrupting this axis leads to the expansion of non BTNL3/8-recognizing Vδ1+ T cells that may exacerbate gluten sensitivity in celiac disease[11,12].

CD1 molecules

Vδ1 T cells also recognize several MHC or MHC-like ligands, of which the CD1 family of molecules is best studied. The nonclassical class I CD1 family of ligands includes CD1a, b, c, and d, of which only CD1d is also found in mice. CD1d is well known as a ligand of NKT cells, an innate like αβ T cell subset originally characterized as binding CD1d loaded with α-galactosyl ceramide (α-GalCer), but this subset of cells has since been found to bind to more diverse lipid antigens as well. Our lab and others have structurally and functionally characterized CD1d binding by Vδ1 TCRs, allowing us to use CD1d as a model ligand for studying Vδ1 TCR signaling and functional polarization[13-15]. The structural and biochemical characterizations of these TCRs reveal that Vδ1 T cells can bind CD1d in both lipid specific and lipid agnostic manners (Fig. 1), suggesting that some TCRs bind to the CD1d platform at baseline unless disrupted by the presence of a lipid with a bulky head group, while others require the presence of specific lipids in the CD1d groove in order to bind and activate[13-15]. The DP10.7 TCR, which our lab crystallized in complex with CD1d loaded with sulfatide, belongs to the latter category, and binds to CD1d-sulfatide with exquisite specificity and high affinity (Fig.1) making it a very useful model TCR for the experiments testing TCR signal described in chapters 2 and 3.

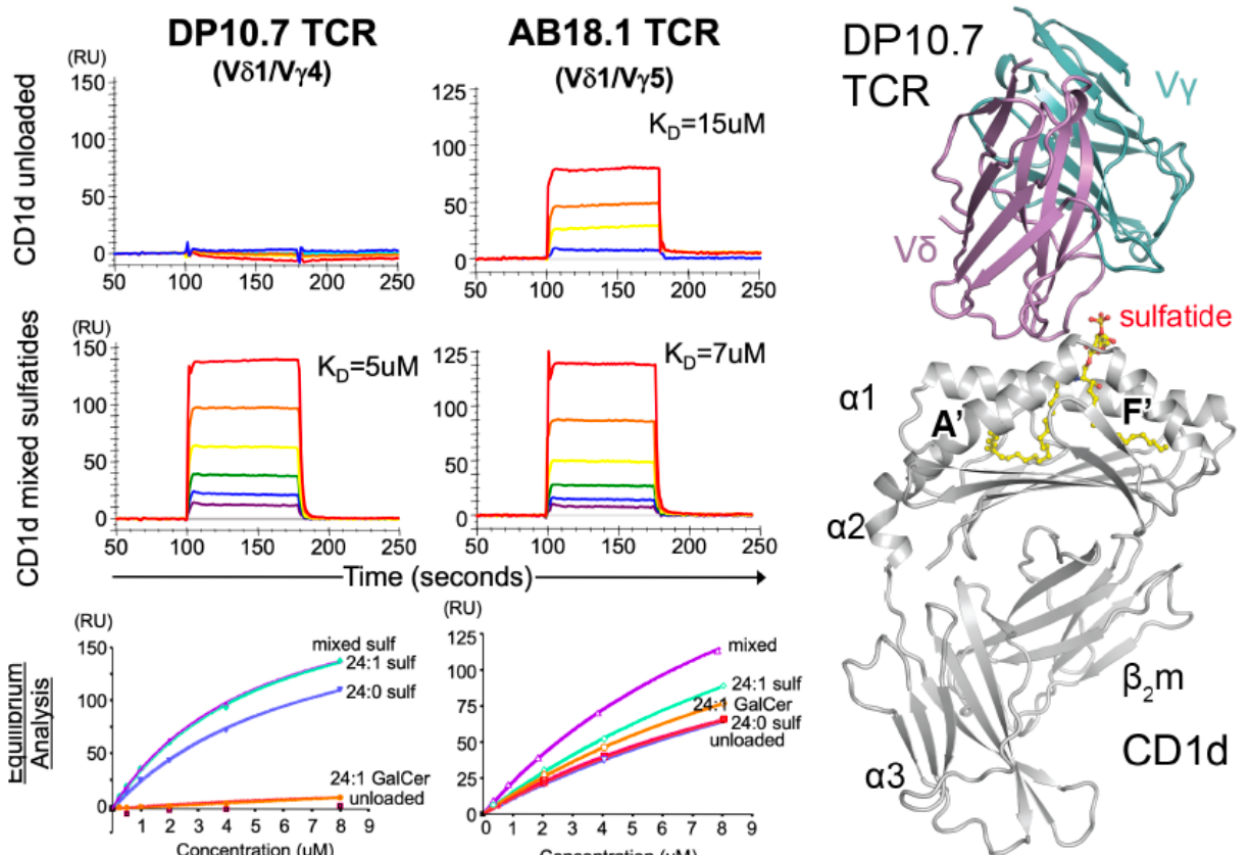

Figure 1: Molecular characterization of CD1d recognition by Vδ1 TCRs.

Left: SPR measurements for DP10.7 and AB18.1 TCRs with CD1d "unloaded" or leaded with sulfatides. Equilibrium binding curves shown at the bottom. Right: Complex structure of DP10.7 TCR with CD1d-sulfatide (PDB: 4MNG, Luoma et al.). The Vδ1 domain mediates all contacts with CD1d. Data shown from Luoma et al., 2013[13].

γδ T cell functional phenotypes

As alluded to earlier, γδ T cells respond to both innate and adaptive signals and can therefore become both innate and adaptive effectors. In several viral infection contexts, a diverse population of γδ T cells clonally expands in the context of viral infection, and shows decreased repertoire publicity in virally infected patients, mirroring a traditional adaptive T cell response[16,17],[1]. In CMV [3], the clonally expanded Vδ1+ T cells also had decreased CD27 expression while retaining proliferative capacity and TCR signaling. Other studies of clonally expanded and TCR stimulated γδ T cells in the periphery suggest TCR stimulation can drive

differentiation of γδ T cells from naïve (TCF7+LEF1+CD27+) to effector (Tbet+Eomes+BLIMP1+GzmB+) states along with clonal expansion. This effector state most closely resembled the T_{EMRA} (CD27-CD45RA+) phenotype in CD8+ αβ T cells. Some γδ subsets, like other innate lymphocytes, such as NK cells and NKT cells, have also been shown to activate and expand in response to cytokine and/or NK receptor signaling, even in the absence of an overt TCR signal or ligand[1]. The role the TCR plays in priming cells to achieve this state, and whether this potential is restricted only to certain semi-invariant or innate-like subsets is poorly understood.

αβ vs γδ T cell development

The development of a functional repertoire of T cells is a challenge for an organism. It must have sufficient diversity in its repertoire to respond to potential threats, and the TCRs it produces must be functional, but this pressure to generate a diverse repertoire can lead to a breakdown in tolerance and autoimmunity. For this reason, T cell development in the αβ compartment is very carefully orchestrated. First, β chains are tested paired with a pTα chain to ensure the beta chain is functional and able to traffic to the cell surface. Once a thymocyte passes this checkpoint, it undergoes cell division to produce a cohort of cells with a functional β chain. Then these cells must make a functional α chain, a process which usually requires multiple recombination events on both chromosomes. After making a receptor that expresses on the cell surface, the αβ thymocyte must undergo positive selection, where it is checked against self-ligands to ensure the TCR can recognize antigen in the context of the self MHC and ensure that CD4 and CD8 are expressed by class II-recognizing and class I-recognizing T cells, respectively. Subsequently, T cells with TCR that bind too strongly to any self peptides derived from healthy self are deleted

after recognition on dendritic cells and medullary thymic epithelial cells (MTEC). During this process, some high affinity TCRs for self-peptide/MHC are diverted to the regulatory T cell lineage to act as dominant immunosuppressors in the periphery or to a specialized CD8alpha/alpha regulatory lineage that resides in the gut epithelium. Each step of this process requires careful spatial and transcriptional regulation, and many details are still being studied[18]. As mentioned in the previous section, many recognized putative γδ T cell ligands are self-ligands. This naturally leads to questions about how these cells undergo selection and how they maintain tolerance in the periphery and implies that their developmental and functional programs diverge substantially from conventional αβ T cells.

Mouse

In mice, γδ T cells develop in the thymus in discrete waves. The first wave of γδ T cells to exit the embryonic thymus are known as DETCs, express the Vγ5 chain, and home to the skin[9,19,20]. The next few waves are also characterized by their γ chain usage and each traffics to a specific tissue, including lung, liver, and gut. All fetal and a large subset of postnatal γδ thymocytes leave the thymus as differentiated effectors, producing either IL-17 or IFN-γ and expressing the associated transcription factors and effector markers[9].

The effector status of γδ T cells leaving the thymus in mice is largely determined by TCR signal strength. The model of γδ T cell lineage selection in mouse thymus, which includes work from several groups, posits that γδ lineage selection and effector fate determination occur either concurrently or sequentially in the thymus. Early DN thymocytes simultaneously undergo rearrangement at the beta, γ, and δ TCR loci. Unlike αβ thymocytes, γδ thymocytes do not

express a pre-TCR chain, and rather recombine and express both γ and δ chains simultaneously. A strong TCR signal from the γδTCR drives DN2 cells into the γδ lineage over the alpha beta lineage, while a weak γδTCR signal will divert the cell from the γδ lineage to the alpha beta lineage at the β-selection checkpoint[21,22]. After γδ lineage commitment, a strong TCR signal likely drives γδ T cell precursors into Type-1 like Tbet+ IFNγ+ effectors, while a weaker TCR signal promotes RORγt+IL-17+ effector commitment. This process is mediated by the precise balance of several networks of transcription factors, including Notch and Egr-ERK-ID3 for lineage selection, and Maf and Tcf7 for effector specification[23-26]. Thus, while the mechanistic details of these regulatory networks are still unclear, it has been widely shown and accepted that for most murine γδ T cells, TCR signal strength is a critical determinant of both γδ lineage and effector selection during thymic development.

Human

Unfortunately, most of these observations do not translate neatly to human γδ development. As mentioned in the previous section, mouse γδ T cells are classified by their γ chain usage while human γδ T cells are classified by their δ chain. Both Vδ1 and Vδ2 T cells can be found, at varying abundances, in different organs and diseases, and do not follow the stereotyped wave pattern of development found in mice. The lack of congruence between human and mouse γδ T cell development may suggest divergent overall functions in the two organisms. However, Vγ9Vδ2s mostly develop in the fetal thymus, while more diverse and adaptive Vδ1 T cells make up the majority of postnatal thymic γδ T cells[27]. There is much less evidence for thymic effector programming in humans, although a recent study observed Type 1-like, Type 3-like, and even a small number of Type 2-like effector cells in the most differentiated cluster of cells in the fetal

thymus, suggesting that functional pre-programming may be a feature of human γδ thymic development as well[28]. Additionally, they saw narrowing of the TCR repertoire and decreasing publicity as cells approached thymic egress, suggesting some form of TCR driven selection in the fetal thymus[28]. The transcriptional networks regulating selection in human γδ T cells have proven challenging to study, and it is still unknown whether the mouse mechanisms tying TCR signal strength to cell fate are conserved in human γδ T cells.

Tissue resident γδ T cells

While some γδ T cells circulate in the blood, most of them, especially of the non-Vγ9Vδ2 subsets, reside in barrier tissues including the skin, intestine, and lungs[2]. The next section describes some specific characteristics of these tissue resident populations. Most of the work on tissue resident γδ T cell populations so far has been done in mice, but work is ongoing in our lab and others to precisely characterize tissue resident γδ T cells in human tissues as well.

Skin

In mouse, the skin is colonized by two distinct subsets of γδ T cells. The dermis largely contains Vγ4+ and Vγ6+ γδ T cells, while the epidermis is dominated by Vγ5+ dendritic epithelial cells (DETCs)[9]. DETCs participate in skin repair, and monitor the epithelial barrier by extending their dendrites towards the stratum corneum layer of the epithelium[29]. In the context of skin injury, they produce insulin like growth factor 1 and keratinocyte growth factors to encourage proliferation of keratinocytes, and recruit macrophages to the sites of injury[30]. DETCs exit the thymus programmed to produce IFN-g, but under certain circumstances may be reprogrammed to produce IL-17 in the skin, which has been shown to have contradictory wound healing and

injury exacerbating functions, depending on context[9]. Both dermal and epidermal γδ T cell subsets promote hair follicle stem cell proliferation during wound healing[9].

Intestine

Intraepithelial lymphocytes (IELs) reside in the intestinal epithelium but are highly motile cells, which allows them to perform dynamic tissue surveillance[9]. There are both αβ and γδ IELs, but for the purposes of this introduction, I will focus on the γδ subset, which have substantial functional overlap with the αβ subset in both mice and humans.

In mice, γδ IELs are largely Vγ7+ cells that bind to and recognize BTNL molecules in the gut epithelium, which aid in their extrathymic maturation. As noted in the previous section, the analogous population in humans is the Vγ4+ population, which binds to BTNL3/8 dimers in the intestinal epithelium using germline encoded residues in the HV4 region of the TCR[11]. IELs have been shown to have a variety of physiological roles, most notably participating in wound healing through the promotion of growth factor expression. IELs are characterized by tissue residency and express high levels of CD69, with effector programs that include granzyme B, aryl hydrocarbon receptor, and NK receptor expression in addition to their TCR, but can also take on a regulatory immunosuppressive phenotype[31]. While their direct roles in specific human diseases are still largely unclear, lack of tonic signaling through recognition of BTNL3/8 by the Vγ4+ TCR may exacerbate celiac's disease[12].

γδ T cells in human disease

Infection

γδ T cells are often considered the first line of defense in an infection because of their ability to rapidly respond to innate and adaptive signals. Both adaptive and innate-like expansions of γδ T cells have been described in different infection contexts.

As described in the previous section, CMV seropositive people have larger clonal expansions of Vδ1+ T cells that correlates with functional effector differentiation, suggesting adaptive clonal expansion and differentiation downstream of foreign antigen-specific TCR signals in this context[3]. While some possible antigen mediators of clonal expansion have been proposed, including EPCR and HLA-DR, the specific antigens driving these clonal expansions are still unknown[9].

Cancer

γδ T cells have shown contradtictory protumor and antitumor effects depending on the cancer type and γδ T cell subset. Work in colorectal tumors suggested Vδ1+ T cells can infiltrate and attack tumor cells, particularly in tumors with mismatch repair defects[4]. Another recent study identified AREG+ γδ T cells as a protumor subset in colorectal cancer, but distinguished this subset from AREG- cytotoxic γδ T cells that infiltrated endometrial cancers and performed antitumor functions[5].

Vγ9Vδ2 cells can also exert antitumor effects by detecting high levels of phosphoantigens in cancer cells as a result of disruption of the mevalonate pathway. This cell subset dominates the circulating γδ T cell population, but is often a minority of γδ T cells in tissues. However, the ability of this subset to detect metabolic disruptions in target cells is being exploited in developing “off the shelf” cellular therapies for cancer.

Summary

Functioning at the intersection of innate and adaptive immunity, γδ T lymphocytes make up a small minority of human peripheral blood T lymphocytes but can be the majority of T cells in barrier tissues[2]. γδ T cells play a critical role in a range of immune and non-immune functions, from antiviral responses[3] and antitumor responses[4,5,32-34] to wound healing[5] and memory[35]. As the lineage-defining receptor, the γδ TCR has been of particular interest to our lab. In my thesis work, I explore how T cell receptor signaling directs thymic development of postnatal Vδ1+ thymocytes, identifying multiple cell fates and evidence of TCR instruction. Additionally, I investigate how structural variations in the constant region of the γδ TCR inform signaling, development, and peripheral function of γδ T cells. In all, dissecting the contribution of the γδTCR to γδ T cell development and function in humans is critical to understanding and manipulating these cells in health and disease.

Chapter 2: T cell signaling instructs Vδ1 T cell differentiation in the thymus

Introduction

Previous studies in human γδ T cell development have established several characteristic patterns of thymic output in humans. Human fetal thymus produces mostly Vδ2+ cells, of which the largest subset is butyrophilin/phosphoantigen sensitive Vγ9Vδ2 T cells[27]. These cells enter circulation and can also home to various tissue sites. Whether Vγ9Vδ2 T cells undergo selection on butyrophilin ligands in the thymus is unclear.

A recent study defined the functional profiles of egressing Vδ2+ cells in the fetal thymus and found that most of these cells have already taken on an effector cell fate. The study also found the TCR repertoire decreased in publicity over developmental time, and that certain public or semi-public TCR sequences were associated with certain effector cell fates, following defined and relatively non-overlapping trajectories to a TCR-dependent phenotypic endpoint[28]. This suggests that γδ T cell precursors in the fetal thymus undergo TCR-driven repertoire honing, and some aspects of TCR signaling, potentially the strength of the TCR signal like in mouse, inform the effector programming of the cell.

Whether postnatal thymocytes, which are dominated by Vδ1+ output, also have TCR-driven differentiation trajectories is not known. Additionally, postnatal γδ thymocytes are thought to exit the thymus as naïve T cells rather than differentiated effectors, further complicating the TCR-instructive thymic cell fate model[27]. Postnatal thymic output is also known to have much more diverse TCR repertoires than fetal thymic output, due to a time dependent switch in TdT activity that leads to more N-additions in postnatally derived γδTCRs[36]. Another challenge to understanding if selection occurs in Vδ1+ thymocytes is the huge diversity of ligands these cells

bind, many of which are self-ligands or thus far uncharacterized, making it difficult to understand what the TCRs may be selected on.

Assuming any TCR selection that may occur is guided by ligands expressed in the thymus, there are two non-classical T cell selection models we can use to hypothesize how postnatal Vδ1+ TCR signaling may inform T cell fate. The first is the elegantly described selection and tissue homing program of mouse dendritic epidermal T cells (DETCs). DETCs emerge from the thymus and populate the skin between days 14 and 18 of gestation. They require expression of butyrophilin family member Skint-1 in the thymus to upregulate skin homing and effector markers and appropriately seed the tissue[20]. Knocking out TCR signaling downstream components, like ZAP-70 or Lck, depletes this cell population substantially. While direct binding of DETC TCRs to Skint-1 has not yet been demonstrated, these experiments suggest DETCs are positively selected on TCR signal from self-ligand, then traffic to the tissue where the ligand is most abundant by upregulating the appropriate tissue homing markers downstream of TCR activation[37]. Once in the periphery, these cells likely perform "normality sensing," relying on tonic low-level TCR activation by Skint-1 in the skin to keep them primed to respond to other signals of tissue stress[38].

The other model for comparison with Vδ1 thymocytes is the selection process of αβ NKT cells. NKT cells and Vδ1 cells share CD1d-lipid as a ligand, and therefore may overlap in peripheral function and thymic TCR selection. At the double positive (DP) stage, thymocytes that express the invariant iNKT TCR bind to CD1d expressed on other DP thymocytes to activate the TCR. In addition to TCR activation, DP iNKT precursors express SLAM molecules, which undergo homophilic binding, providing a required second signal for their positive selection. It is still not

known whether iNKT cells undergo negative selection, and unlike DETCs, iNKT cells do not require continuous CD1d binding and TCR signaling in the periphery to persist or activate[39]. To delve into the contributions of TCR signaling in Vδ1 T cell development, we decided to use CD1d as a model ligand to track antigen-specific binding across development and ideally see if more differentiated clusters are enriched for antigen binding, as a proxy for selection. As an additional tool, I used CD1d platform mutants, originally characterized by previous lab member Adrienne Luoma, to probe specifically for upright CD1d platform binding as a product of thymic selection. To understand the rules of CD1d recognition by Vδ1+ T cells, Adrienne made a series of point mutations on the CD1d platform to investigate which residues are critical for Vδ1 TCR binding.

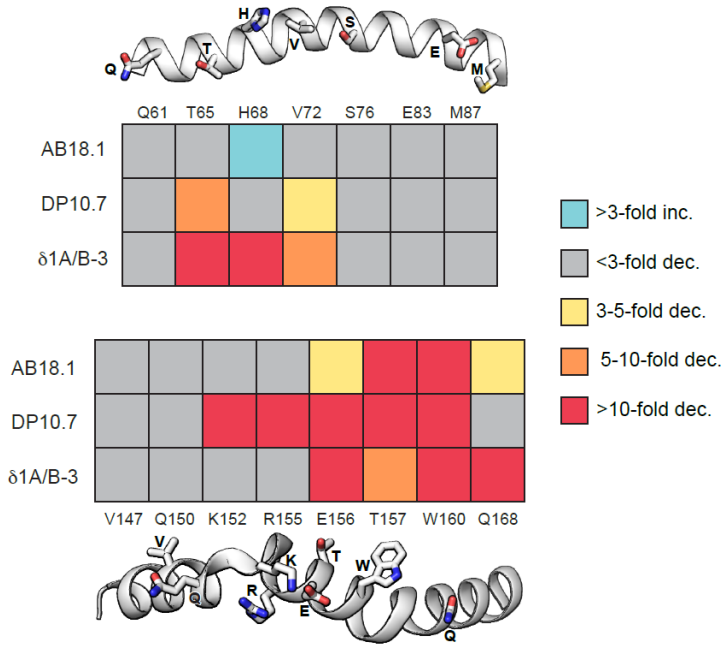

Figure 2.1: CD1d platform mutagenesis highlights key residues for Vδ1 TCR recognition

Summary of CD1d alanine mutagenesis and surface plasmon resonance measurements. Sulfatide-loaded WT or alanine-mutant CD1d (0.625-20 μM) was flowed over a sensor chip with immobilized TCRs. Reference-subtracted traces were used to generate equilibrium binding curves for the determination of affinity constants (K*d*) using GraphPad (Prism). The ratio of alanine-mutant to WT CD1d affinities was determined and represented as fold-decrease or -increase as vs WT as shown. Done by Adrienne Luoma.

As illustrated in Figure 2.1, the W160A mutant ablates binding of all 3 tested CD1d-binding TCRs, while the K152A mutant only disrupts binding of the DP10.7 TCR, which is lipid specific. Hypothesizing that the lipid-specific CD1d binders may undergo a conventional naïve to effector activation mechanism while the lipid-agnostic TCRs may perform "normality sensing" like the DETCs, I leveraged these tetramers to also investigate whether antigen selectivity plays a role in Vδ1 T cell thymic programming (Fig. 2.2A).

In sum, I used a combination of tetramer staining, single cell RNA sequencing, and T cell receptor sequencing to investigate how γδTCR signaling informs postnatal Vδ1 thymocyte development.

Results

To investigate TCR contributions to Vδ1 T cell development, I collected, with the help of Dr. Andrew Koh and Dr. Narutoshi Hibino, 6 thymus samples from pediatric heart surgery patients. The patients ranged in age from 4 days old to 5 months old. Three thymi were from male patients and three from female patients. Initial screening using flow cytometry showed that approximately 1% of CD3+ cells from each thymus were Vδ1+, and between 2-8% of those Vδ1+ cells were tetramer positive, depending on the lipid used (Fig. 2.2B,C). Since I wanted strong amplification of γδ T cell receptors in my single-cell dataset and was interested specifically in the role of the TCR in development, I decided to enrich for my population of interest by sorting on Vδ1+ cells. I then spiked in 10% of tetramer positive αβ T cells to serve as an internal control for tetramer staining and TCR sequencing. After sorting, I generated and sequenced libraries for surface barcoded protein and tetramer, total RNA expression, and paired αβ and γδ TCR expression. To generate these libraries, I used a combination of 10X Genomics'

5’ Immune Profiling kit and my own primers and tetramers, which I helped develop and validate with Augusta Broughton and Caitlin Castro and are detailed further in the materials and methods chapter of this thesis.

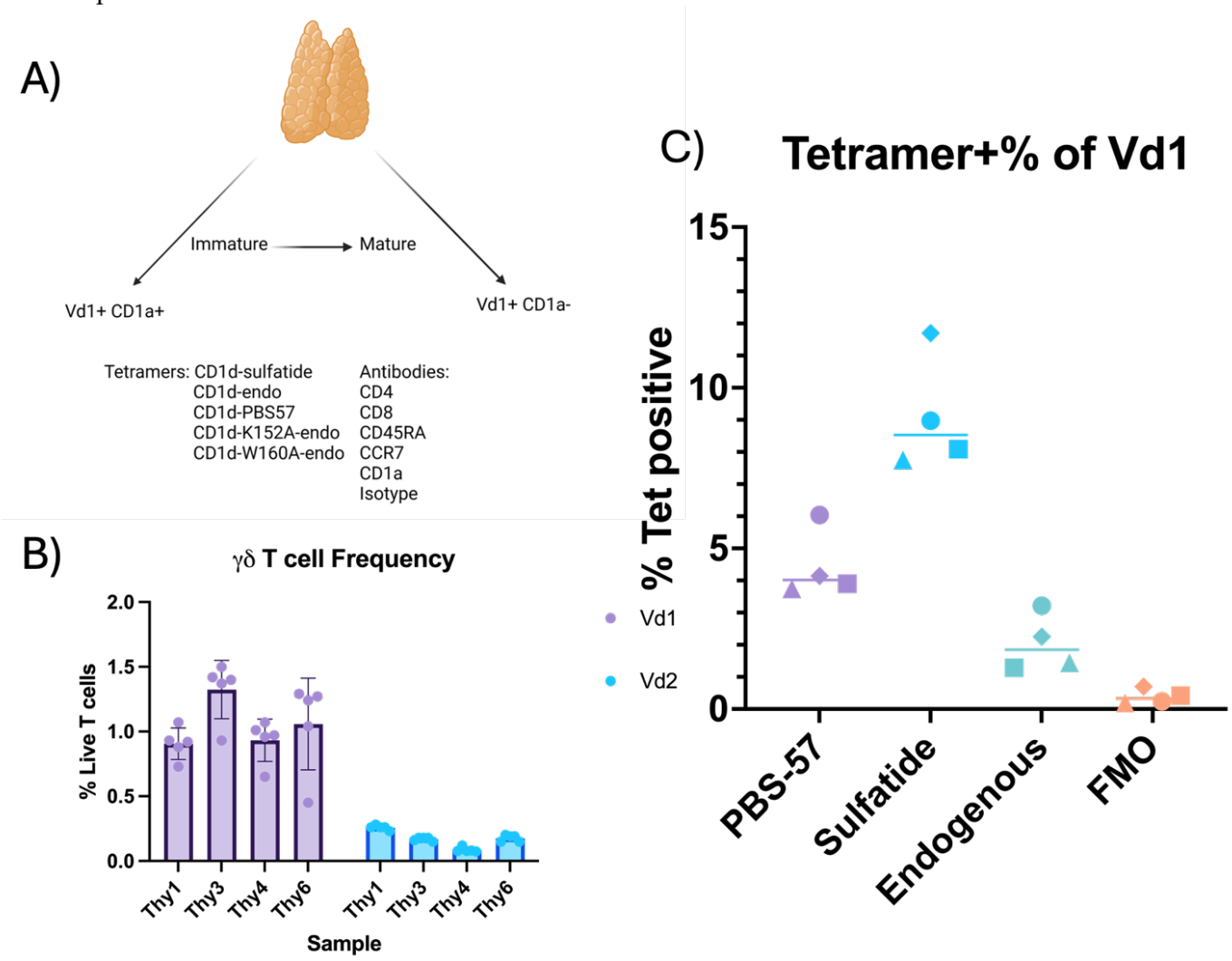

Figure 2.2: Experimental outline and tetramer screen

A) Experimental outline. Vδ1+ cells were sorted from thymi and stained with listed oligonucleotide barcoded reagents. B) Percentage of CD3+ live cells staining with Vδ1 (purple) or Vδ2 (blue) antibody in each sample. C) % Tetramer positive of Vδ1+ cells. Each dot represents a technical replicate.

Postnatal thymocytes have a variety of phenotypes in the thymus identifiable by known immune markers

After filtering and QC, I performed dimensional reduction on the dataset to yield a UMAP with 15 clusters (Fig. 2.3A). Vδ1 T cells from all samples integrated well, except for one cluster of cells derived exclusively from donor 4 (Fig. 2.3B). For the remainder of the analyses shown in figures 2.4-2.7, this cluster was removed and the cells were reclustered to prevent sample-specific biasing of analyses. αβ and γδ thymocytes mostly clustered separately, reflecting their divergent developmental phenotypes (Fig 2.3C). Surface protein counts and gene expression data corresponded well for shared markers and tetramer staining corresponded to expected staining frequencies from flow cytometry, with expected high density staining in the alpha beta clusters, since all the alpha beta thymocytes were sorted on tetramer (Fig. 2.3D).

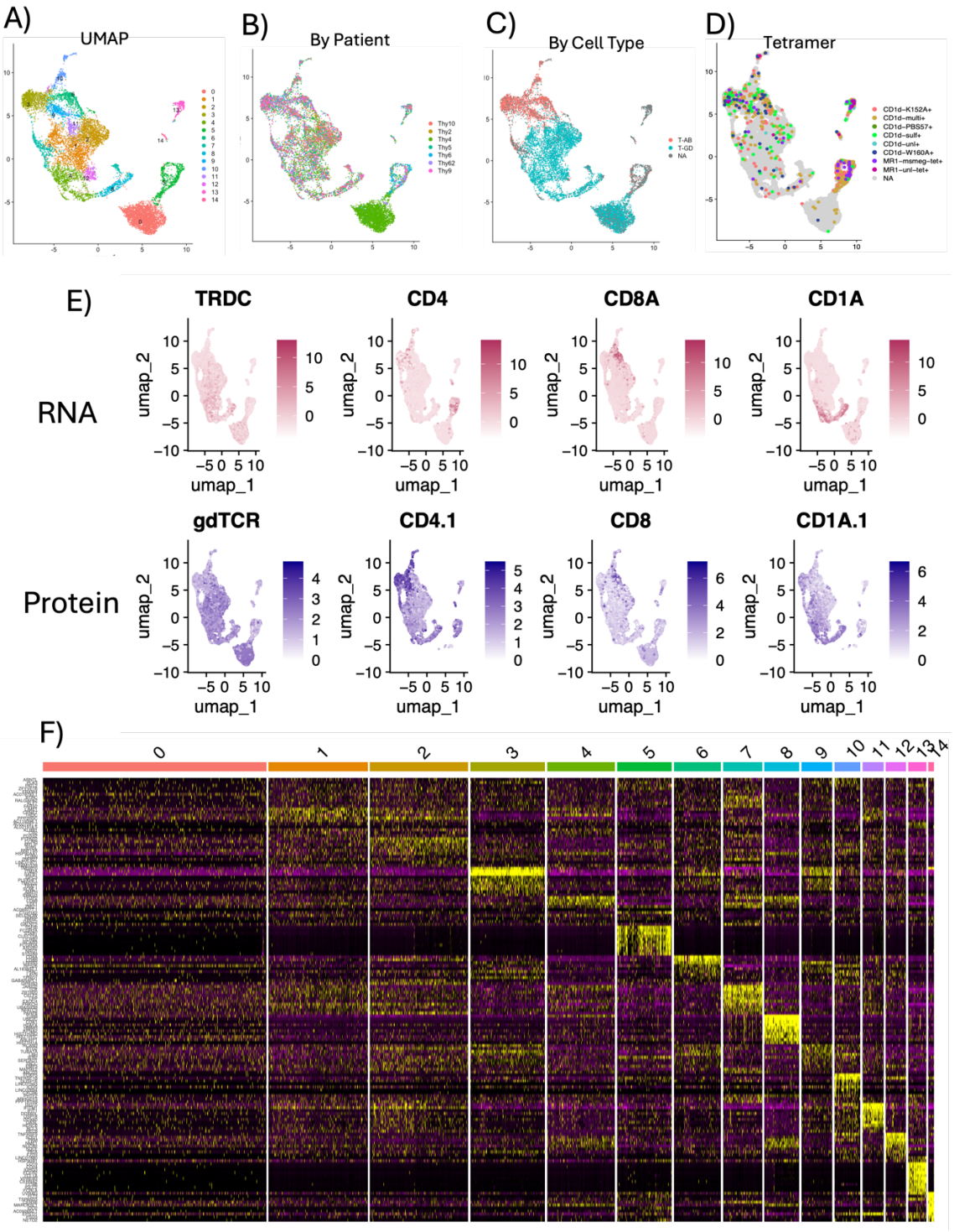

Figure 2.3: Initial clustering and QC

A) UMAP of all samples after Seurat integration. B) UMAP grouped by patient. C) UMAP grouped by cell type. Cell types were classified by expression of a paired TCR. D) Top: RNA expression of key markers over UMAP. Bottom: Protein expression of markers labeled with CITE-seq antibodies after normalization. E) Differentially expressed genes heatmap for full dataset.

Differentially expressed gene analysis revealed recognizable gene sets that allowed labeling of the clusters (Fig. 2.3E). The identified clusters included several innate like effector clusters, αβ Tregs, several clusters of dividing cells, naïve CD4+ and CD8+ cells, as well as several progenitor cell types.

Vδ1 thymocytes develop into Type-1 like, cytotoxic, and regulatory phenotypes

To dive deeper into Vδ1 specific dynamics, I subsetted my combined αβ/γδ T cell dataset on only the cells that expressed a paired γδ T cell receptor to ensure cell identity and therefore appropriate trajectory and clustering predictions. I reclustered the subsetted γδ T cells, and similarly to Fig. 2.3, was able to assign cluster identities to most clusters based on the top differentially expressed genes, highlighted in Fig. 2.4A. There were several clusters of dividing cells and Vδ1 progenitors expressing known developmental markers like *CD1A*. There was also a cluster of Type 1 like effectors, identified by highly differentially expressed genes (*EOMES*, *IFNG*, granzymes) (Fig. 2.4B). This effector identification was corroborated by high gene module score for the Type 1 signature (Fig. 2.4C), taken from Sanchez et al., who used the same gene sets to define fetal γδ thymocyte development. I also identified a naïve/memory like cluster (*SELL, CCR7, CCR9, THEMIS*), an innate-like cluster (*AREG, KLRK1, KLRC4, IL7R*), a CDαα like cluster (*GNG4*, *XCL1*, *CD27*) and an interferon induced cluster (*IFIT1*, *MX1*, *IFNG*, *ISG150*).

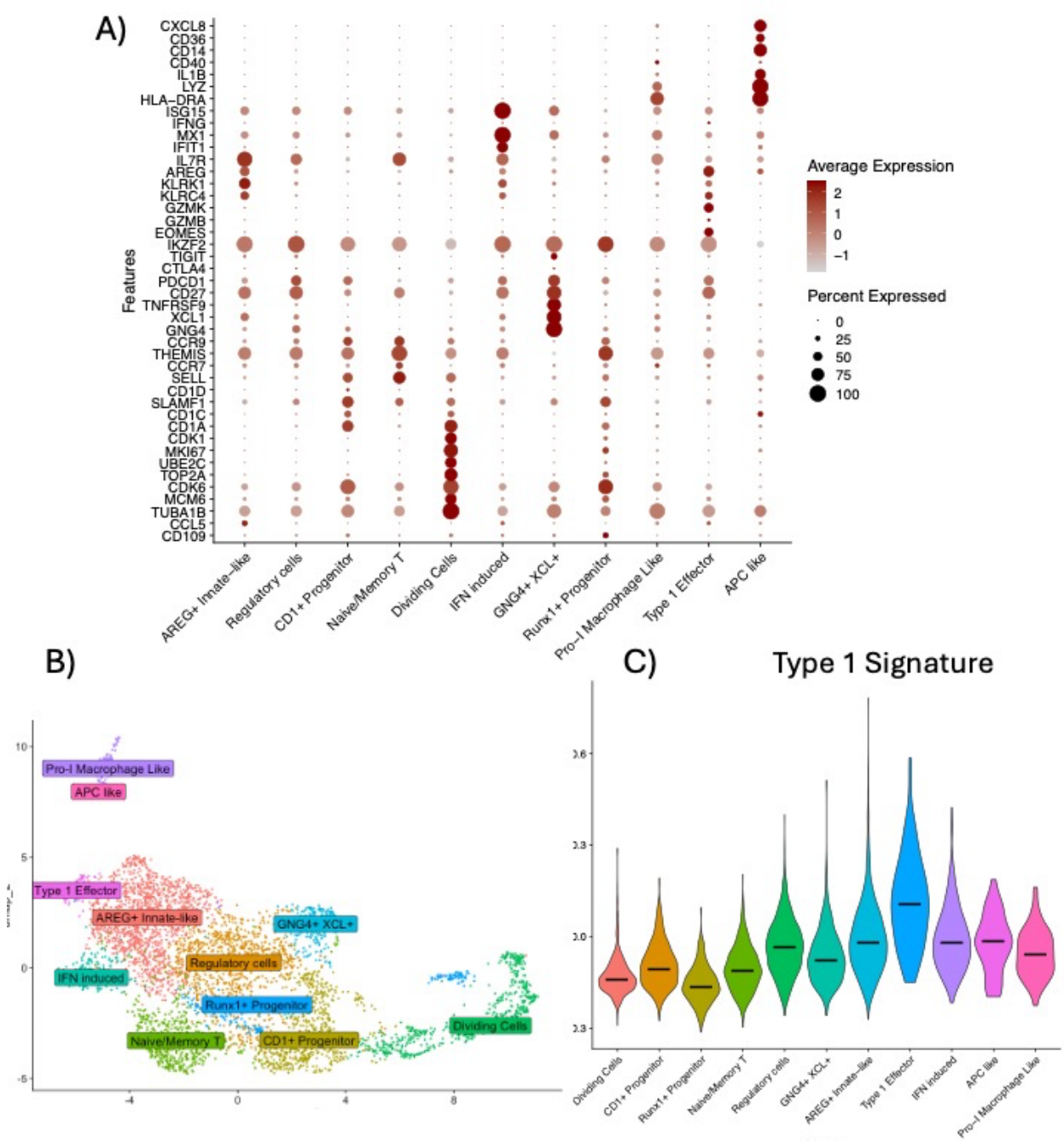

Figure 2.4: Vδ1+ thymocytes cluster into diverse phenotypes

A) Significant differentially expressed genes in each cluster, displayed with renamed cluster IDs on the x-axis. B) UMAP of subsetted γδ T cells only with new cluster IDs. C) Type 1 signature shown by cluster. Comparisons noted in the text are significant ($p<.01$) by Kruskal-Wallis testing and paired t-test between groups.

Notably, there were two strongly pro-inflammatory macrophage-like clusters (Fig 2.4B, upper left) that clustered separately from the rest of the cells and expressed *IL1B* and *CD14*. One of

them specifically highly expressed MHC class II and *CD40*, taking on an APC-like phenotype, while the other highly expressed *CD36* and *CXCL8*. Because these cells all express productive and paired γδ TCRs, it is unlikely that they are a contaminant, and more likely that a small subset of thymic Vδ1s are able to polarize into a strongly pro-inflammatory or APC like phenotype. Finally, there was a large cluster of apparently immunosuppressive regulatory γδs, expressing *PDCD1*, *CTLA4*, *TIGIT*, and *IKZF2* which encodes Helios, a transcription factor known for driving the development of thymic Tregs in the αβ T cell compartment. Thus, clustering of thymic Vδ1s revealed several novel thymic cell populations thus far uncharacterized in developing human γδ T cells.

I next wanted to understand which of these clusters corresponded to a true developmental cell fate. To apply developmental logic to a two-dimensional dataset, I performed pseudotime analysis using the Monocle3 package to superimpose a computationally presumed developmental trajectory onto the clusters. Because the macrophage-like clusters were not continuous with the main UMAP, they were omitted from the pseudotime trajectory, as their temporal relationship to the main clusters could not be inferred. The trajectory analysis was rooted in the clusters most highly expressing CD1A, a known early γδ developmental marker. The trajectory showed differentiation from the proliferating cell clusters through intermediate progenitors into several possible cell fates (Fig 2.5A). The trajectory produced endpoints in 5 different possible cell fate clusters, including the AREG+ cluster, the GNG4+ cluster, the naïve cell cluster, the IFN cluster, and the Type 1 cluster.

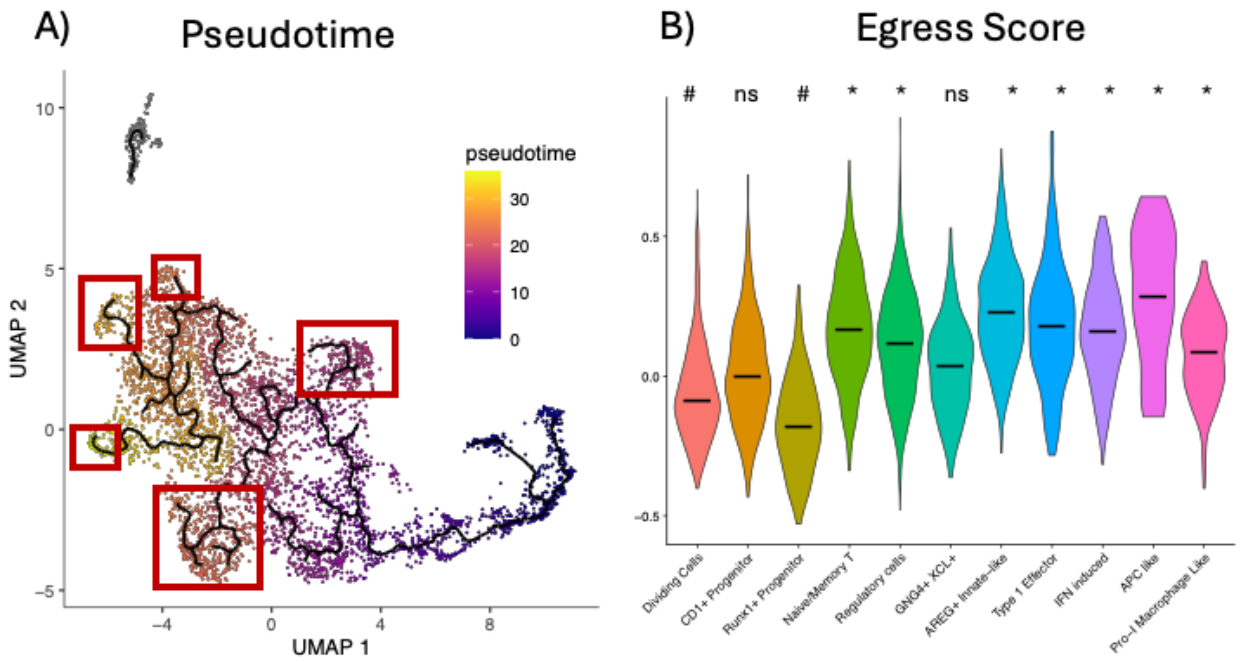

Figure 2.5: Vδ1 thymocytes differentiate into 5 distinct cell fates.

Monocle3 trajectory overlaid onto UMAP of clusters. Yellow represents most differentiated clusters. Hypothesized trajectory endpoints are boxed in red. B) Thymic egress score shown as a violin plot by cluster. *significantly higher than CD1+ progenitor group. #significantly lower than CD1+ progenitor group. p<2.2e-16 by Kruskal-Wallis testing and *# p< .001 by paired ttest between groups.

To understand whether each trajectory endpoint was a genuine cell fate that would exit the thymus and populate the periphery, I calculated the thymic egress score for each cluster using the same gene set used to score fetal thymic γδ egress (Figure 2.5B). This analysis revealed that while all the possible cell fates have elevated thymic egress scores, the AREG+ and Type 1 like clusters, as well as the APC like cluster which was not included in the pseudotime trajectory had the highest egress scores. The IFN cluster and the naïve cluster had the next highest egress scores, with the regulatory cluster close behind. The GNG4+ cluster had a notably lower egress score, suggesting that this cluster is either not terminally differentiated or represents a thymic resident population. All the possible cell fate clusters had statistically significantly higher thymic egress scores than the progenitor or pro-inflammatory macrophage like clusters, corroborating

the pseudotime trajectory prediction. This additionally suggests that the APC-like cells may represent a γδ cell fate that exits the thymus while the macrophage-like cells may be thymus resident or a developmental precursor, but we cannot distinguish between those possibilities with these data. In sum, postnatal Vδ1+ thymocytes progress through proliferative developmental stages and differentiate into Type 1, innate like, regulatory, and naïve phenotypes before leaving the thymus.

TCR signaling is highest in the most differentiated thymocytes

Next, to get to the core of whether TCR signaling itself correlates with cell fate, I calculated the immediate early gene score, which encompasses a range of genes known to be triggered quickly by TCR binding[40,41]. This score therefore serves as a readout of TCR signaling on a per cell basis, which we then aggregate to the cluster as a whole to avoid skewing due to incomplete gene coverage. This analysis reveals that TCR signaling tracks strongly with differentiation by trajectory analysis, with the Type 1, AREG+, and APC-like clusters having statistically significantly higher TCR signaling than any of the other clusters (Fig. 2.6A,B). The IFN-induced gene group closely followed in high immediate early gene expression along with the macrophage-like cluster. These were then followed by the naïve cells, regulatory cells, GNG4+ cells, and Runx1+ progenitor clusters, which have a similar amount of TCR signaling and are similarly "differentiated" in the pseudotime analysis. Finally, the lowest signaling clusters, The CD1A+ progenitors and dividing cells, correspond to the earliest progenitors by pseudotime. The continued and differential levels of TCR signaling as the cells progress into divergent terminal cell fates suggest that TCR signaling informs Vδ1 T cell differentiation fate.

Importantly, the IEG gene signature expression and pseudotime trajectory do not share a 1:1 relationship. There are distinct, highly differentiated cell fates, like the IFN induced fate, which is the most differentiated by pseudotime, that does not have the highest TCR signaling (Fig. 2.6A,B) . This suggests that beyond just acting as a survival signal, strength of TCR signal has some deterministic impact on differentiation fate.

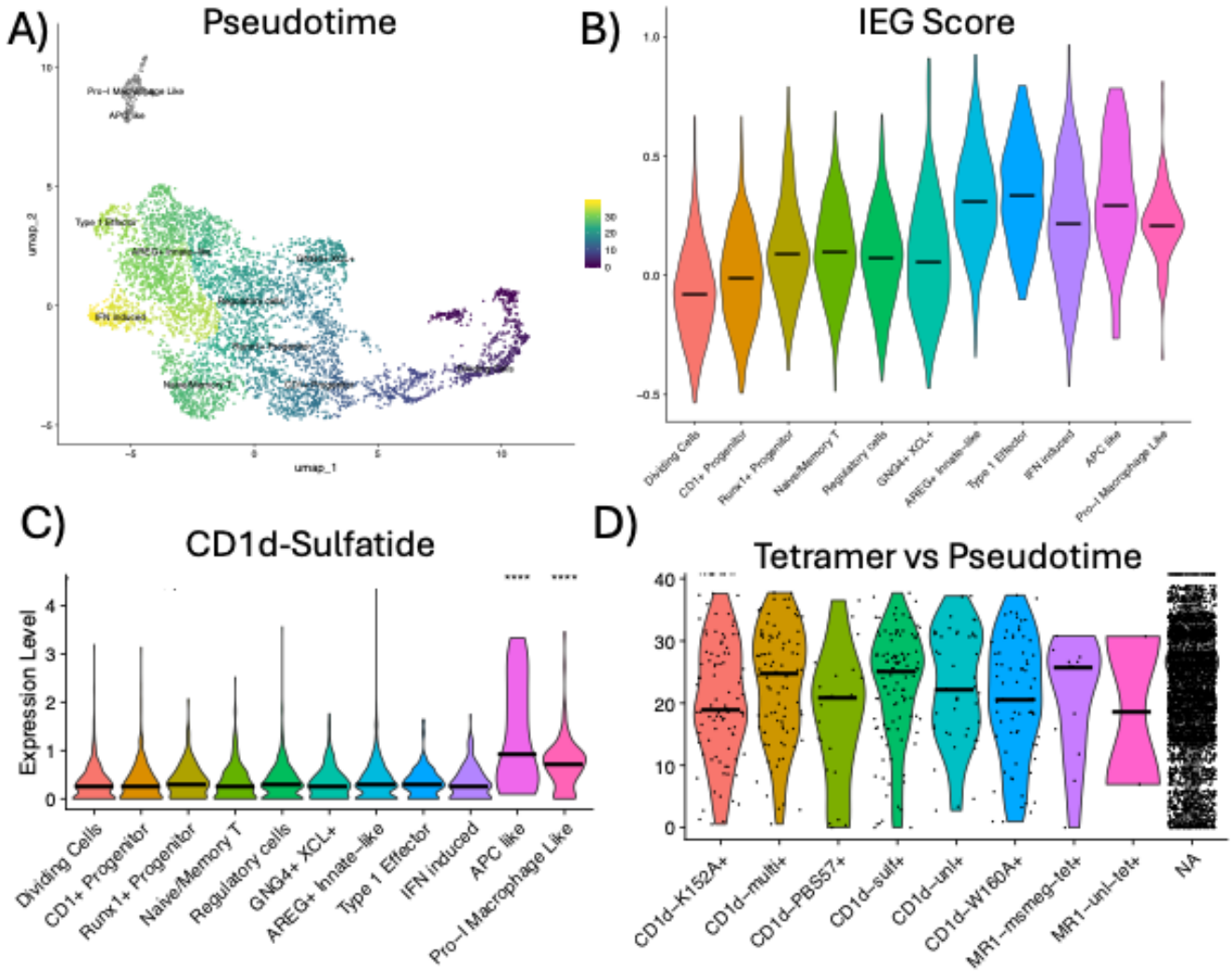

Figure 2.6: TCR signaling and tetramer binding over pseudotime

A) UMAP colored by pseudotime. Higher number/yellow is more differentiated. B) Immediate early gene score calculated using gene list in Appendix II. C) CD1d-sulfatide tetramer by normalized protein counts shown as a violin plot. D) Tetramer staining shown as violin plot over pseudotime on the y-axis. Higher number indicates more differentiated cells. ****p<2.2e-16 by Kruskal-Wallis testing and p<.001 paired ttest between groups.

More differentiated Vδ1 thymocytes are slightly, but not significantly, enriched for tetramer binding

I then hoped to leverage the tetramer reads to identify correlations between TCR signaling and tetramer binding, to determine if high TCR signaling could be explained by antigen-driven positive selection. I was particularly interested in whether the CD1A progenitor cluster, which significantly expressed CD1D, B, and C as well as SLAMF1, represented a progenitor state where Vδ1 T cells might be presenting selecting ligands to each other in a mechanism analogous to iNKT cell selection. Unfortunately, the tetramer coverage was quite low, especially once counts were normalized, and it was difficult to correlate tetramer binding to trajectory. In fact, the clusters most significantly enriched for tetramer binding were the macrophage like clusters, which were not part of the trajectory (Fig. 2.6C). I then gated tetramer positivity on a per cell basis, requiring a minimum of 1.5 normalized counts and 2-fold counts over negative control MR1 tetramer, which led to just under 300 cells classified as positive for at least one tetramer. When plotted against pseudotime, the cells that bound CD1d-sulfatide, CD1d multi-reactive, and MR1 loaded with bacterial ligands were the most differentiated, and the ones that bound the mutant tetramers were the least differentiated (Fig. 2.6D). This is a promising piece of preliminary data as it suggests that within the CD1d-binding compartment, there may be selection for upright binding of the CD1d groove, and this selection may not be lipid specific. The MR-1 binding cells also follow a similar pattern. The MR1-msmeg antigen specific cells are more differentiated in pseudotime while the cells that selectively bind unloaded MR1 are less differentiated, suggesting MR1-metabolite guided development of the T cell receptors in the MR1-binding compartment as well. However, these data were not statistically significant due to

the large spread and few cells, so further studies must be done to conclude on the role of ligand in TCR driven cell fate specification in postnatal Vδ1s.

TCR clusters and motifs correlate with functional specification

To further investigate whether functional specification could be directly correlated with TCR sequence and biochemical features, I performed a graph vs graph analysis between the γδ TCR UMAP and the gene expression UMAP using the package CoNGA[42], which was built to detect correlation between T cell gene expression profile and TCR sequence in single-cell datasets. This analysis produced a few TCR-gene expression correlated clusters, but the most interesting was an effector-like cluster differentially expressing *HLA-C, PARP8, IL7R, TXK,* (Figure 2.7, 2.8A,B) that overlapped in markers perfectly with some of the top differentially expressed genes from the AREG innate like effector cluster.

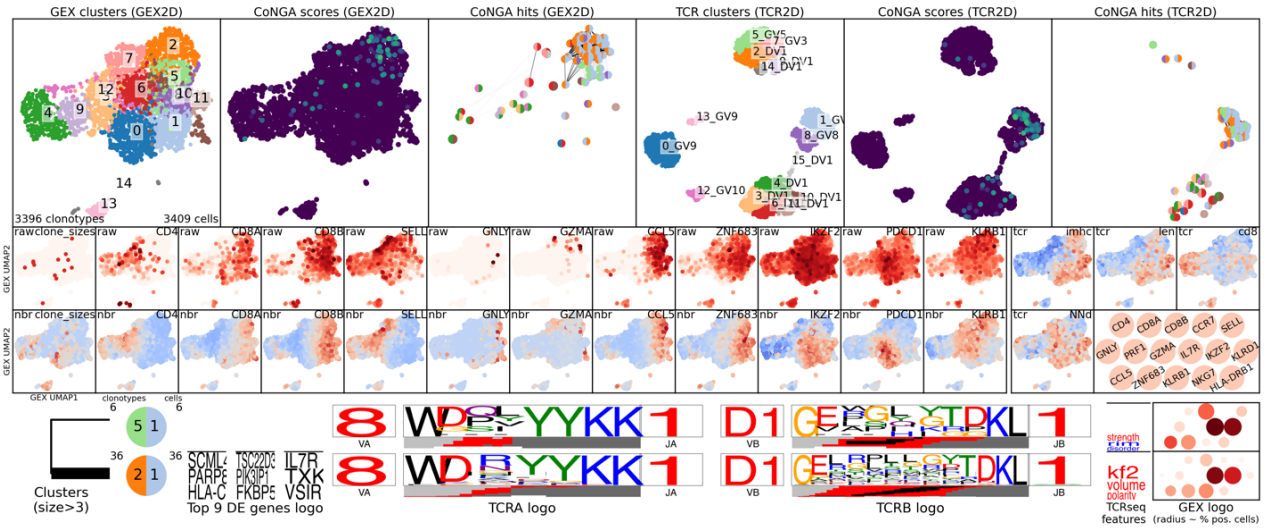

Figure 2.7: CoNGA analysis results

Top row: UMAP projections of RNA (left 2 boxes) and TCR (right 2 boxes) dimensional reductions. The center box and rightmost box highlight the locations of the CoNGA hits, which are the TCR and GEX clusters that have significant overlap. Middle 2 rows: showing expression of various key T cell genes overlaid on GEX UMAP. Bottom: Left: Cluster relationship of CoNGA hits, with number of clonotypes and number of cells listed beside a pie chart. Right side of pie chart shows GEX cluster and right side shows TCR cluster. Next is the list of the top 9 most differentially expressed genes in the CoNGA hit. TCR logos follow, as explained in figure 2.7. Rightmost section shows highly differential TCR features in the CoNGA hit, as well as scaled expression of key genes in bubbles, as listed in the panel directly above.

This TCR-GEX correlation is another independent analysis suggesting that TCR structure and signaling inform cell fate differentiation in Vδ1 thymocytes well beyond the αβ/γδ lineage selection checkpoint.

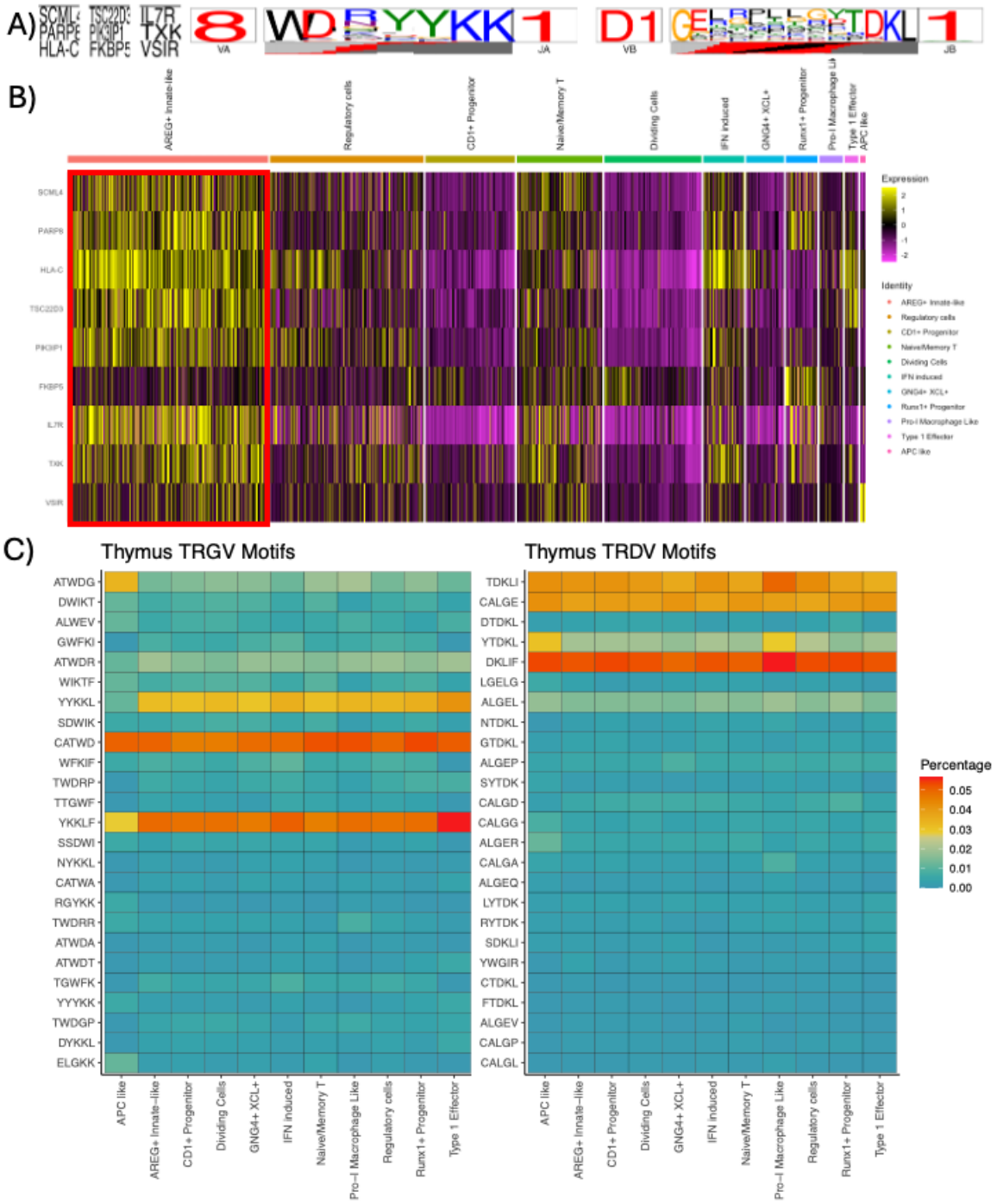

Figure 2.8: Graph vs graph analysis and enriched motifs

γ (left) and δ (right) chain enriched motifs in the TCR cluster highly significantly correlated with AREG+ effector RNA cluster. B) Enriched 5 amino acid γ chain motifs shown by cluster. C) Enriched 5 amino acid δ chain motifs, shown by cluster.

Further corroborating this, the CDR3γ motif enriched in the CoNGA correlated TCR cluster, YYKK, is most highly enriched in the 3 cell fates on the AREG+ effector branch by trajectory analysis—the Type 1 effectors, AREG+ cells and IFN+ cells (Fig 2.8A,B). These are also the three most functionally similar effectors as they are all differentiated, Type 1-like, and cytotoxic in profile, suggesting that enriched γ chain features may participate in guiding these cells down this branch of the pseudotime differentiation trajectory. Finally, it is worth noting one of the most highly represented motifs in the δ chain, YTDKL, in which the tyrosine is a non-templated residue, is most enriched in the two macrophage-like clusters, outside of the trajectory analysis (Fig. 2.8C). These are also the clusters that stained most strongly with the CD1d tetramers, whose recognition by Vδ1 TCRs is often mediated most heavily by contacts with the CDR3 loop of the δ chain[13]. The differential enrichment for shared motifs between the two chains may indicate several TCR guided forks in development that could drive diverse phenotypes while still in the thymus.

Discussion:

The T cell receptor is an energetically demanding and extremely "risky" receptor to recombine—a mistake could lead to a breakdown in tolerance and systemic autoimmunity. For this reason, the process of T cell selection in αβ T cells is carefully orchestrated and tightly regulated, with TCR-cued checkpoints for survival and subsequent clonal deletion of TCRs that signal too strongly. In mouse γδ T cells, TCR signal strength determines effector lineage fate, with most γδ T cells entering the periphery fully differentiated[9]. Many of these cells appear to use their TCR binding to self-ligand as a tonic survival and innate effector status maintenance signal, rather than a direct signal for activation[38]. Fetal human γδ T cells seem to follow a similar paradigm,

with limited repertoire diversity, increasing repertoire publicity, and 3 discrete effector cell fates when leaving the thymus[28].

In this study, I outlined the developmental trajectory of postnatal Vδ1+ T cells using single cell sequencing techniques, combining several lines of evidence to demonstrate that TCR signaling in the thymus informs Vδ1 thymocyte development into as many as 5 different cell fates. Trajectory analysis revealed several forks in development that correlated with differential levels of TCR signaling. The lowest TCR signaling was in the earliest progenitors, with medium levels of TCR signaling driving a naïve or regulatory cell fate. Higher levels of TCR signaling pushed progenitors into a Type 1-like differentiation fork, eventually making them into CTL like effectors, AREG+ innate effectors, or IFN producing and responsive cytotoxic cells. Each of these cell fates had a significant thymic egress score, showing that regardless of TCR signal strength, they will all exit the thymus and are terminal cell fates. Independent graph vs graph analysis also showed a significant correlation between highly similar TCRs and a gene expression profile that directly mirrored the AREG+ innate effector cluster phenotype. This supports the idea that similar TCR features may be "selected" to guide cells towards the Type 1 like cell fates. Compellingly, the overall most highly expressed γ chain motifs across the dataset, several of which overlapped with the CoNGA TCR cluster motif, were most strongly enriched in the three Type 1-like effector fates that emerged from the Type 1-like trajectory fork, suggesting that TCR features guide developmental cell fate. δ chain motifs also correlated with cell phenotypes, further supporting TCR guiding of thymocyte cell fate.

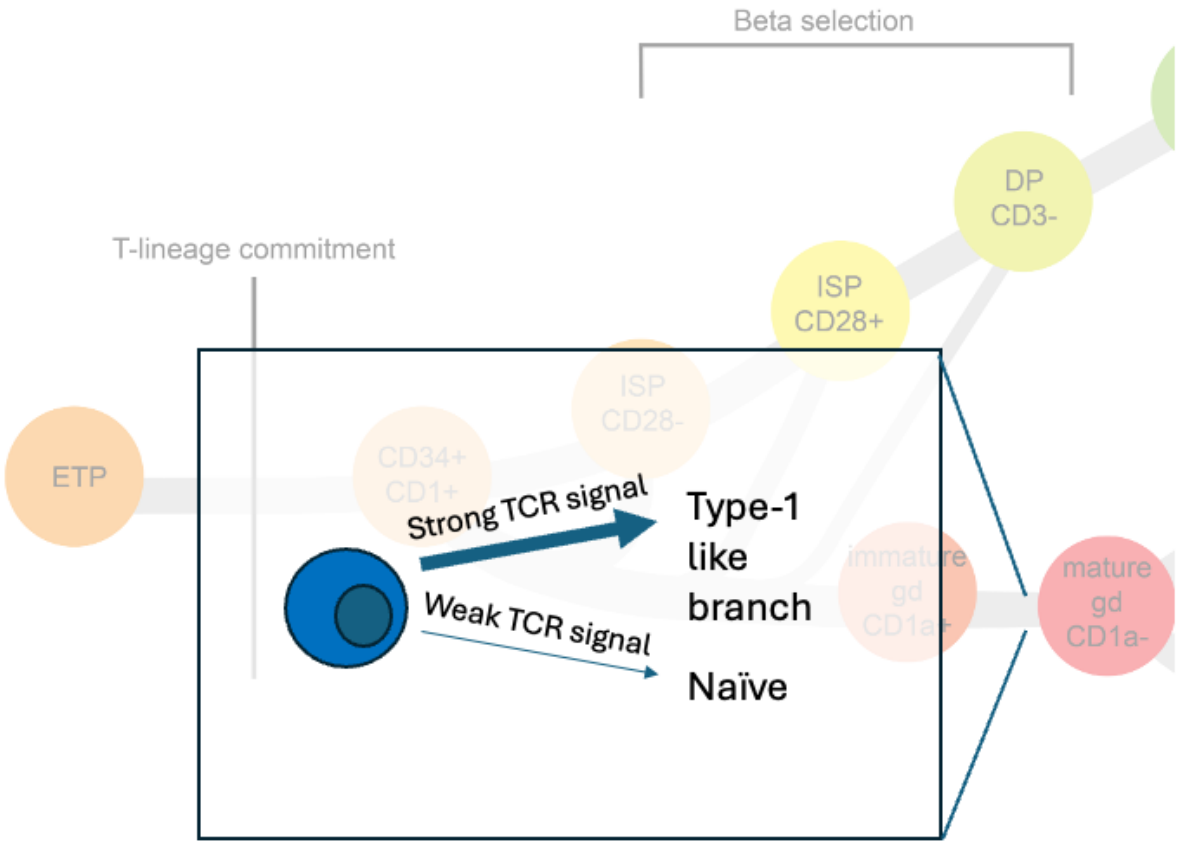

Figure 2.9: Summary model

Summary figure showing TCR signal strength informing thymic effector differentiation in postnatal Vd1+ thymocytes. Development figure adapted from Roels et al., review[27].

This study fills some major gaps in our knowledge about postnatal γδ T cell development. We establish here that, guided by TCR signals, both naïve cells and several classes of effector cells develop and exit the thymus. Despite the enormous diversity of TCRs generated in the postnatal thymus, we show evidence of specific TCR features converging in certain cell fates. Our model of postnatal Vδ1+ development, guided by these data, is that after γδ lineage selection, Vδ1 TCRs signal again, and the strength and persistence of TCR signal guides cell fate decision before thymic egress (Fig 2.9). These findings show yet another critical juncture in a γδ T cell's

life where the T cell receptor signal is critical and will help further our understanding of signal summation and tolerance in these cells.

Chapter 3: γ constant usage tunes γδ T cell receptor sensitivity, thymic programming, and peripheral function

Introduction

Both αβ and γδ T cell receptors are composed of variable domains which contain the complementarity determining loops, and constant domains which encompass the connecting peptide region and transmembrane region. In the αβ TCR, the alpha and beta chains are covalently linked via a disulfide bond between two conserved cysteine residues located proximal to the membrane. To be expressed stably on the cell surface and transduce signal intracellularly, this heterodimer must associate with the CD3 proteins epsilon, γ, δ, and zeta. While the transmembrane regions of both the TCR and CD3 proteins play an important role in assembly of this TCR/CD3 complex, the TCR constant domains are also critical in transducing signals. In the αβTCR-CD3 complex, the α and β constant domain connecting peptides (CPs) bind to and stabilize membrane proximal interactions with CD3 components[43].

Curiously, both the β and γ chain constant domains can be encoded by two separate genes. The β chain constant genes, TRBC1 and TRBC2, differ by only 4 residues in the extracellular portion and appear structurally and functionally interchangeable. Cβ2 makes up a slightly higher proportion of healthy peripheral blood, but both Cβ and Cβ2 are found in every functional subset of CD4+ and CD8+ αβ T cells[44-46], suggesting that TRBC usage is randomly chosen during thymic development and does not affect T cell function downstream. Unlike the β chain, however, the γ chain constant genes encode proteins that differ in their proposed extracellular structure[47]. Cγ2 lacks the interchain disulfide cysteine residue, instead having a tryptophan in

this position. Thus, γδ TCR heterodimers using Cγ2 are not disulfide-linked. Of more interest is the well known length variation between Cγ1 and Cγ2 that lies within the connecting peptide region, proximal to the membrane. Cγ1 is shorter in length than any of the other TCR constant domains while Cγ2 has several allelic variants that differ in length in the connecting peptide (CP) region. These length variants are due to duplication or triplication of exon 2 of the TRGC gene and encode proteins with connecting peptides that are longer by 16 amino acids (multiple alleles) or 32 amino acids (Cγ2*05)[47]. Across human populations, the distribution of these alleles has not been extensively examined, but data from small samples suggest all alleles segregate with detectable frequencies, with the Cγ2*05 triplicate allele at higher frequency in Black Africans than in French, Lebanese, and Tunisian study participants[48].

Early studies examining Cγ use within an individual's repertoire suggested that in peripheral blood, Vγ9Vδ2+ cells utilize Cγ1 in their TCR, while the small minority of Vδ1+ cells present in the blood use exclusively Cγ2[49,50]. However, Vδ1+ cells with Cγ1 constant domains were found in the thymus, suggesting the bias in peripheral blood was driven either by thymic selection or clonal expansion in the periphery[49]. A more recent study examining TRGC transcript expression in the thymus concluded that TRGC1 is expressed early in development and switches to TRGC2 later, but that TRGC1 functions exclusively to promote chromatin accessibility during VDJ recombination and only TRGC2 is incorporated into functional γδ TCRs[51]. These observations contradict each other, illustrating how little is known about the usage and function of these genes. Additionally, none of these studies could differentiate between TRGC2 (2x) and TRGC2 (3x) alleles, so their specific abundances are still unknown.

A recent study visualizing the γδTCR-CD3 complex was unable to resolve the structure of the Cγ and Cδ connecting peptides, and their modeling suggests that differences at a few key residues in the interface with CD3 components abolish direct binding of γδ CPs with CD3, resulting in a high level of flexibility[52]. They also observed that at the same level of CD3 surface expression, TCRs with the same variable region bound more tetramer when expressing Cγ2. Notably, TRGC use also tuned T cell activation, with higher CD69 expression in Cγ1 clones of the same TCR than their Cγ2 counterparts[52].

Thes findings, taken together, suggest that TRGC usage could structurally influence γδ T cell antigen recognition and functional polarization. We set out to test this first in a reductionist model with a recombinant TCR of known antigen specificity, then in human tissues using a PCR-based method, and finally in both published and novel single cell RNA sequencing datasets. We found that TRGC usage tunes activation efficiency, informs developmental phenotypes, and biases clonal expansion in both health and disease.

Results

Cγ1 contributes to stronger signaling relative to Cγ2

We have previously shown that the DP10.7 TCR recognizes CD1d-sulfatide with exquisite specificity, allowing us to use it as a model T cell receptor to probe the effect of γ chain usage on signal strength[13]. To investigate the effect of TRGC usage on γδ T cell receptor function, previous lab member Augusta Broughton, in collaboration with Nicolas Asby in Dr. Jun Huang's lab, paired the full DP10.7 TCR δ chain with its Vγ upstream of either the Cγ1 constant domain, Cγ2*02 constant domain (hereafter known as Cγ2(2x) as it contains the single exon 2 repeat), or Cγ2*05 constant domain (hereafter known as Cγ2(3x)) (**Fig. 3.1A**) and transduced them into

CD1d -/-JRT.3-T3.5 Jurkat cells. The transduced Jurkat T cells activate equally when subjected to anti-CD3 stimulation, implying equivalent machinery downstream of the γδ TCR[53] (**Figure 3.2A**). However, when the Jurkats were co-incubated with C1R cells expressing CD1d and loaded with sulfatide, more Cγ1-expressing Jurkats became activated, as measured by a larger proportion of Jurkat cells expressing CD69 compared to Cγ2-expressing cells (**Fig. 3.1b**). In addition to Cγ1-expressing cells being more activated at a population level (**Fig. 3.1b**), Cγ1+ cells also expressed significantly higher levels of CD69 on a per cell basis when compared with Cγ2(3x) (**Fig. 3.1c**). Attenuation of signaling after antigen-specific activation (controlled for antigen density) was correlated with increasing numbers of exon 2 repeats, with Cγ1 responding most strongly, Cγ2(2x) achieving an intermediate level of signaling, and Cγ2(3x) signaling most weakly[53] (**Fig 3.1c**).

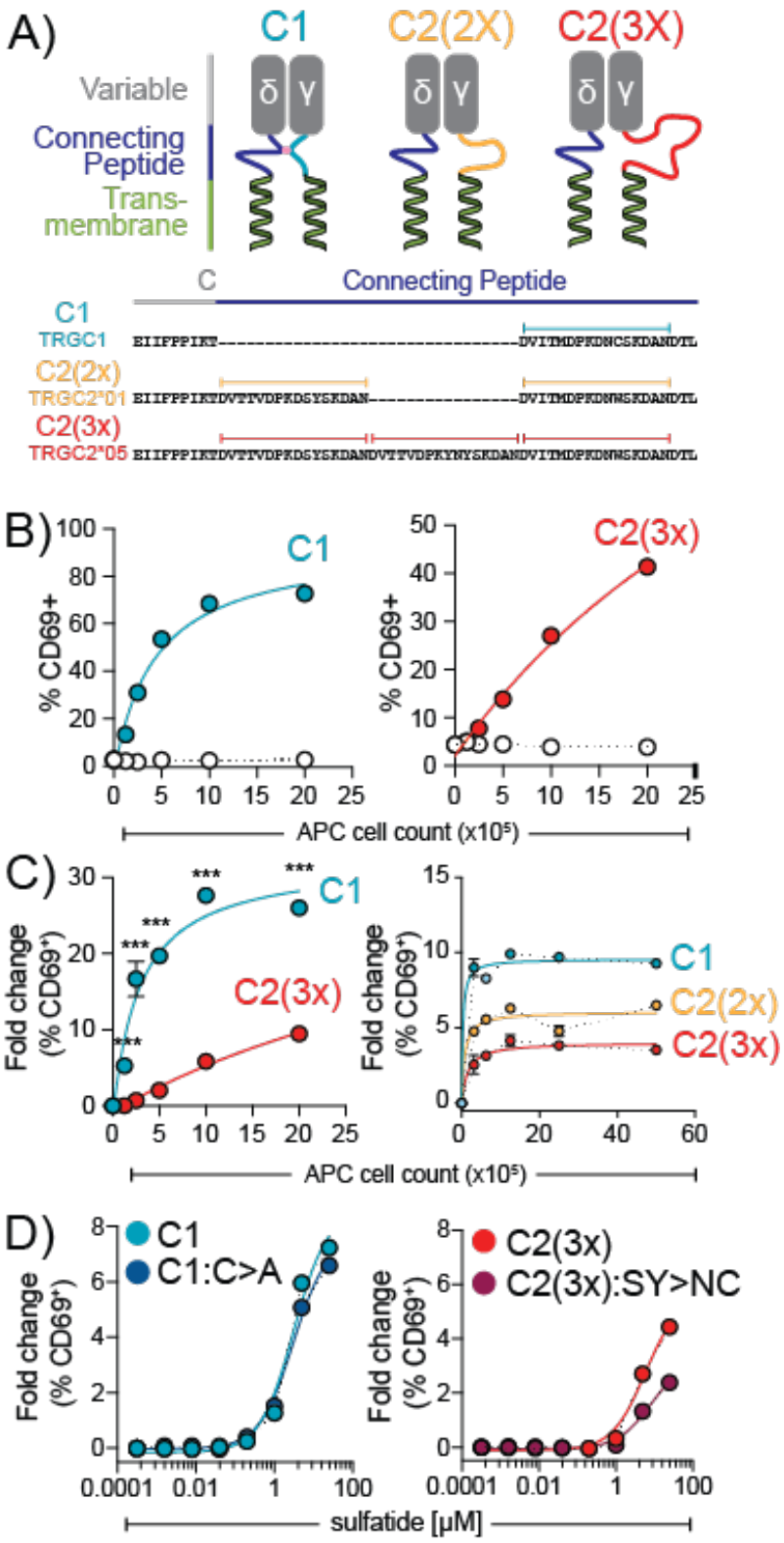

Figure 3.1: Jurkat stimulation with engineered TRGC

A. Alignment of exon 2 of TRGC genes C1, C2 *02 (containing the duplication) and C2*05 (containing the triplication). Graphic illustrates the connecting peptide where the repeats occur in the T cell receptor. Pink dot in C1 graphic represents the interchain disulfide bond unique to TRGC1. **B.** Left: DP10.7 TRGC1 CD69 induction by coculture with increasing # of $CD1d^+$ C1R cells loaded sulfatide (blue filled circles) or unloaded (white circles). Best fit curve (used to estimate EC50) in blue. Right: DP10.7 TRGC2*05 CD69 induction by coculture with increasing # of $CD1d^+$ C1R cells loaded sulfatide (red filled circles) or unloaded (white circles). Best fit curve (used to estimate EC50) in red. **C.** Left: Comparison of DP10.7 TRGC1 (C1-blue) and DP10.7 TRGC2*05 (C2*05- red) CD69 fold change normalized to unloaded $CD1d^+$ C1R for each cell line. Right: Comparison of DP10.7 TRGC1 (C1-blue), DP10.7 TRGC2*01 (C2(2x) - yellow), and DP10.7 TRGC2*05 (C2(3x) - red) CD69 induction normalized to unloaded $CD1d^+$ C1R for each DP10.7 line. Best fit curves (used to estimate EC50) in corresponding colors. **D**. Comparison of CD69 induction with increasing sulfatide concentration among DP10.7 TRGC1 (C1-blue), DP10.7 TRGC2*05 (C2*05- red), and disulfide mutants (DP10.7 TRGC1 C->A, navy; DP10.7 TRGC2*05 SY->NC, maroon). ***p-value<.001

While Vδ1+ TCRs like DP10.7, bind antigens using their CDR loops, cells expressing Vγ9Vδ2 TCRs have a more complex antigen detection mechanism that involves binding BTN molecules with germline encoded residues on the γ chain. To investigate whether the γ constant chain usage also modulates signal transduction of Vγ9Vδ2 TCRs, performed the same experiment as above with the published G115 TCR[54], and observed a similar increase in activation efficiency with Cγ1 compared to Cγ2 (**Fig 3.2b**). Thus, γ constant usage universally impacts γδTCR-ligand sensitivity independently of TCR specificity. Strikingly, the number of exon 2 repeats also

determines the degree of attenuation of activation, suggesting that the CP length indeed contributes to the strength of activation downstream of antigen binding.

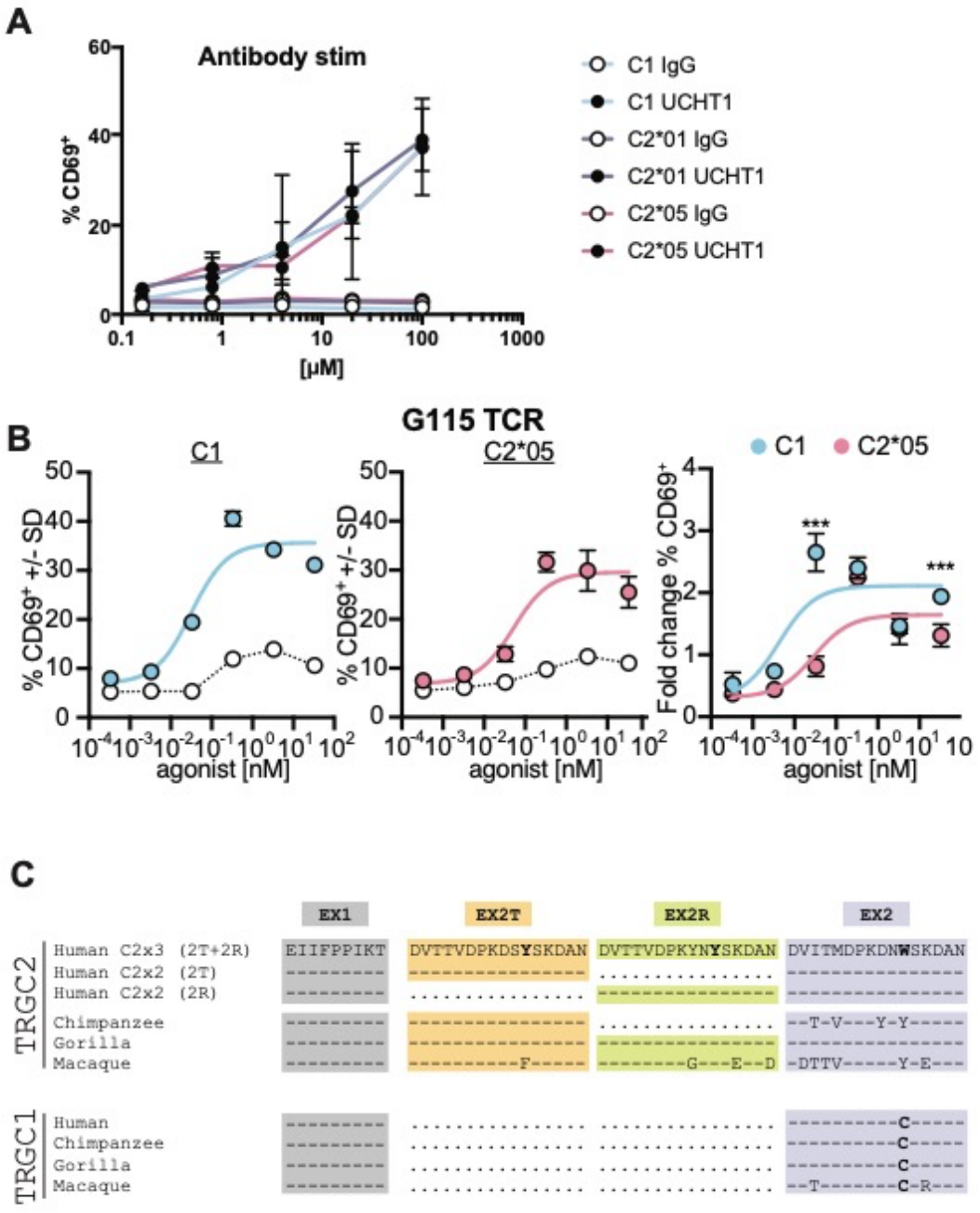

Figure 3.2: TRGC tuning of TCR activation strength is independent of Vγ and conserved in primates

A. DP10.7 TCR with indicated constant chain CD69 induction with plate bound anti-CD3 antibody (clone: UCHT1) (black filled circles) or isotype control (white circles). **B.** G115 TRGC1 CD69 induction by coculture with increasing

concentration of zoledronate (blue filled circles) or unloaded (white circles). Best fit curve (used to estimate EC50) in blue. G115 TRGC2 *05 CD69 induction by coculture with increasing concentration of zoledronate (blue filled circles) or unloaded (white circles). Best fit curve (used to estimate EC50) in pink. **C**. Alignment of TRGC1 and TRGC2 alleles in human, chimpanzee, gorilla, and macaque. Only exon 1 exon 2 with duplications where present are shown Accession numbers are: Chimpanzee.TRGC1 (PNI89314.1), Chimpanzee.TRGC2 (PNI89313.1), Gorilla.TRGC1 (DBA44065.1), Gorilla.TRGC2 (DBA44066.1), Macaque.TRGC1 (AAO75103.1), Macaque.TRGC2 (QJB76138.1).

Variation in Cγ length exists in non-human primates and reflects distinct evolutionary histories

To explore the evolutionary history of these alleles, we aligned the human sequences of these genes and alleles with TRGC genes from macaques, gorillas, and chimpanzees. This alignment shows that, in general, presence of a longer CP in TRGC2 is conserved across primates, suggesting the connecting peptide duplication may have some conserved function. While the extra-long Cγ2(3x) is a minor allele in the human population[48], available TRGC2 CP sequences from gorillas and macaques are the length of the human 3x allele, (**Figure 3.2c**), whereas the chimpanzee sequence is the length of the human 2x alleles. Of note, when compared to the triplicated allele, the duplicated chimp exon 2 aligns to the first repeat (ex2T). Of the 7 duplication alleles in the human population, 1 allele (TRGC2*01) also contains the ex2T, whereas the remainder of alleles align with the second repeat (ex2R) (**Fig 3.2c**). Of course, little is known about polymorphisms in other primate populations, so it is possible that all primate species have both duplicated and triplicated Cγ2s.

Cγ activation efficiency is not dependent on interchain disulfide bond

We hypothesized the difference in activation efficiency between constant domains could be due to presence or absence of an interchain disulfide bond formed between the δ and γ constant domain CPs. Exon 2 of Cγ1 contains the canonical cysteine found in the αβTCR and is capable

of forming an interchain disulfide with Cδ. This cysteine is missing from both Cγ2 (2x) and Cγ2 (3x) and is instead a tryptophan, making this position an attractive candidate for investigation as the mechanism for the observed difference in signaling between the three. However, mutating the cysteine to alanine in Cγ1 to block interchain disulfide bonding failed to diminish the activation efficiency of Cγ1 (**Fig. 3.1d**). Conversely, mutating the serine-tyrosine pair at the corresponding location on Cγ2 to asparagine-cysteine to reconstitute the disulfide bond also failed to boost activation efficiency in Cγ2 (**Fig. 3.1d**). Therefore, the interchain disulfide bond is not responsible for differences in signal strength between Cγ1 and Cγ2[53].

TRGC1, TRGC2(2x), and TRGC2(3x) transcripts are all found in peripheral T cells in vivo

Having established that γ constant chain usage modulates TCR activation efficiency, we wanted to determine whether all 3 TRGC chain variants are expressed in human peripheral tissues. Previous observations suggest that either functional TCRs in the periphery exclusively use Cγ2[51] or the blood is dominated by Cγ1 due to biased pairing of Vγ9Vδ2s with Cγ1[46,50]. Using universal forward and reverse primers flanking the exon 2 repeats of the constant domain (**Fig. 3.3a**) we sorted Vδ1+ and Vδ2+ T cells from healthy donor PBMCs, extracted their RNA, and amplified TRGC transcript from cDNA. PBMCs from multiple donors revealed the presence of both TRGC1 and TRGC2 (2x) usage in both Vδ1 and Vδ2 compartments. One of our two donors, an apparent TRGC2 heterozygote, also showed TRGC2 (3x) expression in addition to TRGC2 (2x). Therefore, both types of γδ T cells, Vδ1 and Vδ2, are able to use either γ constant chain, as well as either TRGC2 allele, and still exit the thymus to become functional T cells in the periphery.

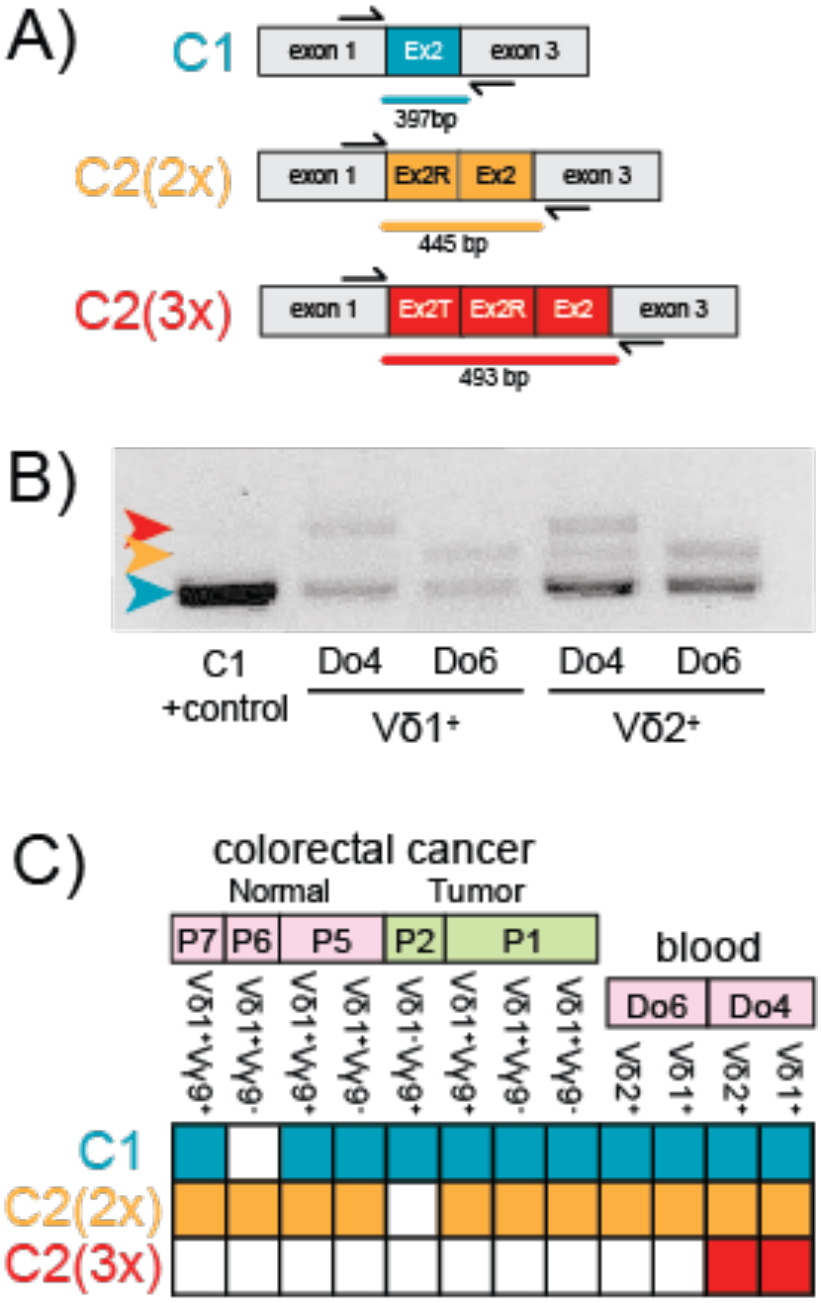

Figure 3.3: TRGC expression in peripheral tissues

A. Schematic indicating amplification strategy for distinguishing TRGC1, TRGC2(2x), and TRGC2(3x) from human samples with expected product size indicated. Black arrows represent approximate primer binding sites. B. Agarose gel visualization of TRGC amplicon from PBMC γδ. Expected amplicon size highlighted with arrows: C1, blue; C2(2x), yellow; C2(3x), red. Lane 1, JR2 a monoclonal cell line uniformly expressing C1 (positive control). Sorted Vδ1+ T cells from Donor 4 (Do4), lane 2, and Donor 6 (Do6), lane 3. Sorted Vδ2+ T cells from Donor 4 (Do4), lane 4, and Donor 6 (Do6), lane 5. C. Agarose gel visualization of TRGC amplicon from CRC IEL cell lines from P1. Expected amplicon size highlighted with arrows: C1, blue; C2(2x), purple; C2(3x), pink. Lane 1, JR2 a monoclonal cell line uniformly expressing C1 (positive control). Lane 2, line sorted on Vδ1+Vγ9-; Lane 3, sorted on Vδ1+Vγ9-; Lane 4 sorted on Vδ1+Vγ9+. D. Schematic of PCR results, presence of corresponding band in each sample is indicated by shading.

To investigate whether both Cγ1 and Cγ2 containing TCRs are found in normal and diseased tissues as well as in peripheral blood, we amplified the TRGC transcripts, as described above, from γδ T cell lines derived from intestinal epithelial lymphocytes (IELs) isolated from either colorectal tumor samples or adjacent normal tissue. Analysis of these samples revealed both TRGC1 and TRGC2 expression in lines derived from either tumor or adjacent normal tissues (**Fig. 3.3d, Fig. 3.4**). Therefore, T cells using both γ constant chains are present in normal and tumor colonic tissue as well as peripheral blood circulation of healthy people.

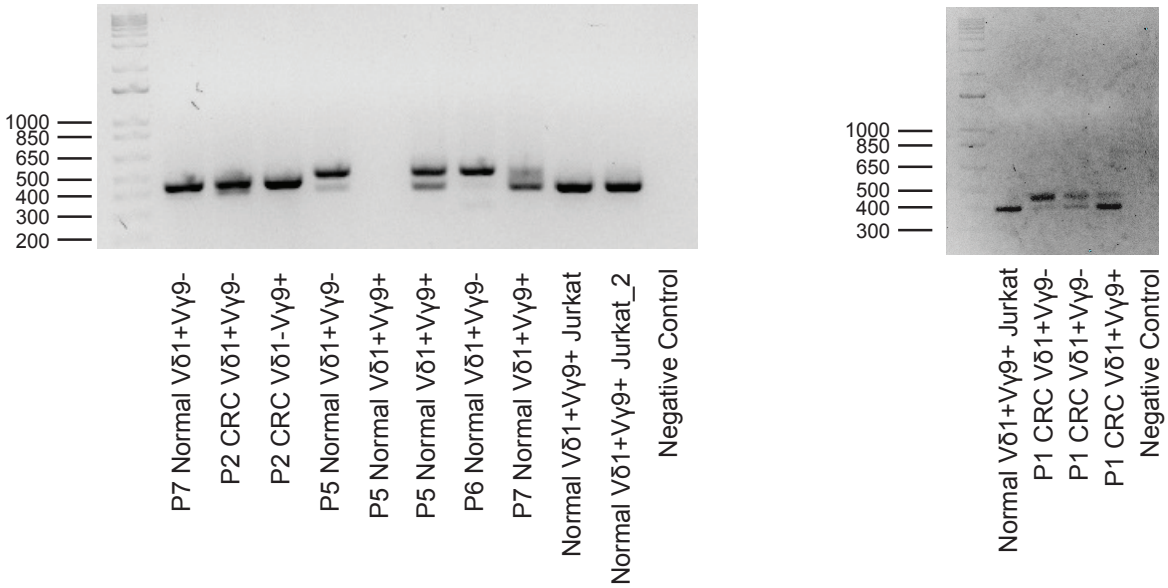

Figure 3.4: SDS-Page Gel shows TRGC1, TRGC2 (2x), and TRGC2 (3x) usage in CRC lines.

Original 3% agarose gels showing amplification of TRGC1, TRGC2 (2x) and TRGC2 (3x) bands illustrated in figure 2.

TRGC1+ and TRGC2+ cells arise in the thymus at different times and have distinct developmental phenotypes

Given that both Cγ1 and Cγ2 TCRs are found in the periphery despite differences in activation efficiency, we wanted to investigate the profiles of Cγ1+ and Cγ2+ γδ T cells through development in the thymus to establish a "time point 0" of constant chain usage proportion and phenotype in γδ T cells. As mentioned previously, there is conflicting evidence in the literature on the timeline and abundances of TRGC usage during development. To address this question, we collected and analyzed single cell RNA sequencing datasets on Vδ1+ γδ T cells from thymi from 6 early postnatal patient samples. Additionally, we reanalyzed published datasets from Sanchez et al., 2022[28], which provided additional data on 6 fetal and 3 pediatric thymi. γδ T cells from all datasets were binned into TRGC1 single-positive or TRGC2 single-positive bins based on normalized gene expression, removing cells expressing transcripts for both chains or below the transcript detection limit for either chain. We determined this to be the most reliable method of binning TRGC use, as VDJ calls were based on very few point mutations in exon 1 of the constant region and corresponded very poorly to gene expression (**Fig. 3.5a**). After binning individual cells based on TRGC transcript expression, we grouped clones based on the TRGC usage of most cells in that clonotype. This ensured that each unique VDJ was assigned its most likely C-based on gene expression, since it is not thought possible for TCR to switch constant chain usage. Additionally, grouping clones together resulted in better correspondence between TRGJP2 and TRGC2, which is the only possible productive J-C pairing after rearrangement to TRGJP2 due to the organization of the γ locus (**Fig. 3.5b,c).**

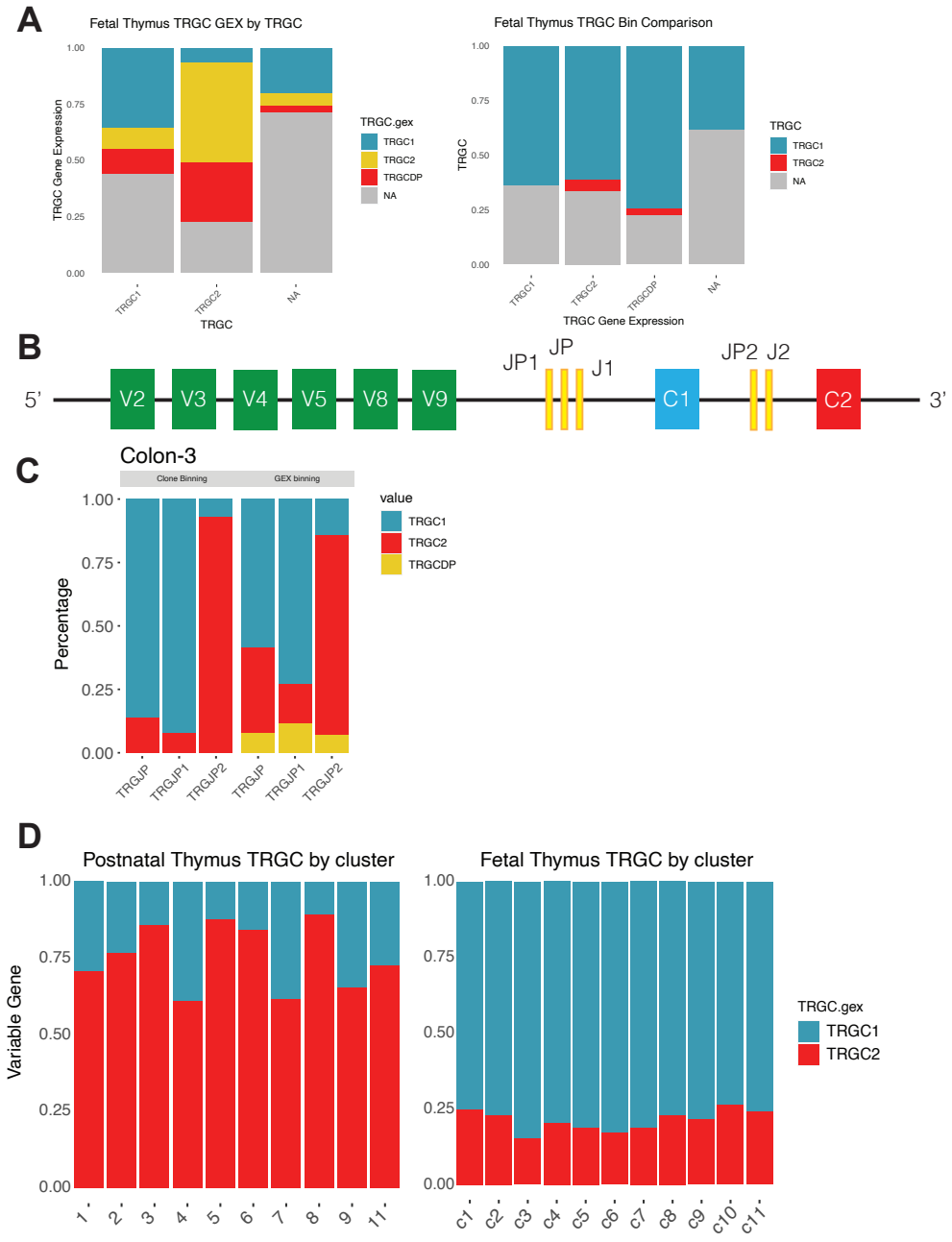

Figure 3.5: Clonal binning provides better concordance of TRGJ-TRGC usage and TRGC1/2 are evenly distributed across thymic clusters

A. Left: Bar graph showing comparison of cells binned by gene expression of TRGC transcripts compared to automatic VDJ assignment by 10x Cellranger v3. X axis indicates VDJ calls and colors indicate gene expression binned calls. Right: X axis indicates gene expression binned calls and colors denote VDJ calls. **B**. Graphical representation of human TRGC locus. **C.** Example bar plot showing TRGJ-TRGC pairing with x axis indicating TRGJ segments and colors indicating the percentage of TCRs in each TRGC bin. Left side of the plot indicates TRGC assignments by binning members of a clone together and right side indicates gene expression based TRGC assignment on a per cell basis regardless of clonal identity. TRGC1 is in blue, TRGC2 in red, and TRGC double positive cells in yellow. Plot

is derived from healthy colon samples from our novel CRC dataset. **D.** Bar graphs showing percentage of cells in each UMAP cluster in each thymus dataset expressing each TRGC gene.

We find that TRGC1-expressing cells dominate the fetal thymic output in both count and proportion, while TRGC2s make up the majority of both postnatal and pediatric thymus (**Fig. 3.6a**). Analysis of the distribution of TRGC1s vs TRGC2s across clusters in each dataset revealed very small enrichments in certain clusters, but overall an equal proportion of TRGC1s/TRGC2s across clusters in both fetal and postnatal thymus (**Fig. 3.5c**), indicating that cells containing both TCR types progress similarly through all stages of development and exit the thymus. However, examining the differentially expressed genes between TRGC1 and TRGC2 cells did reveal small but statistically significant differences in functional and transcriptional programs.

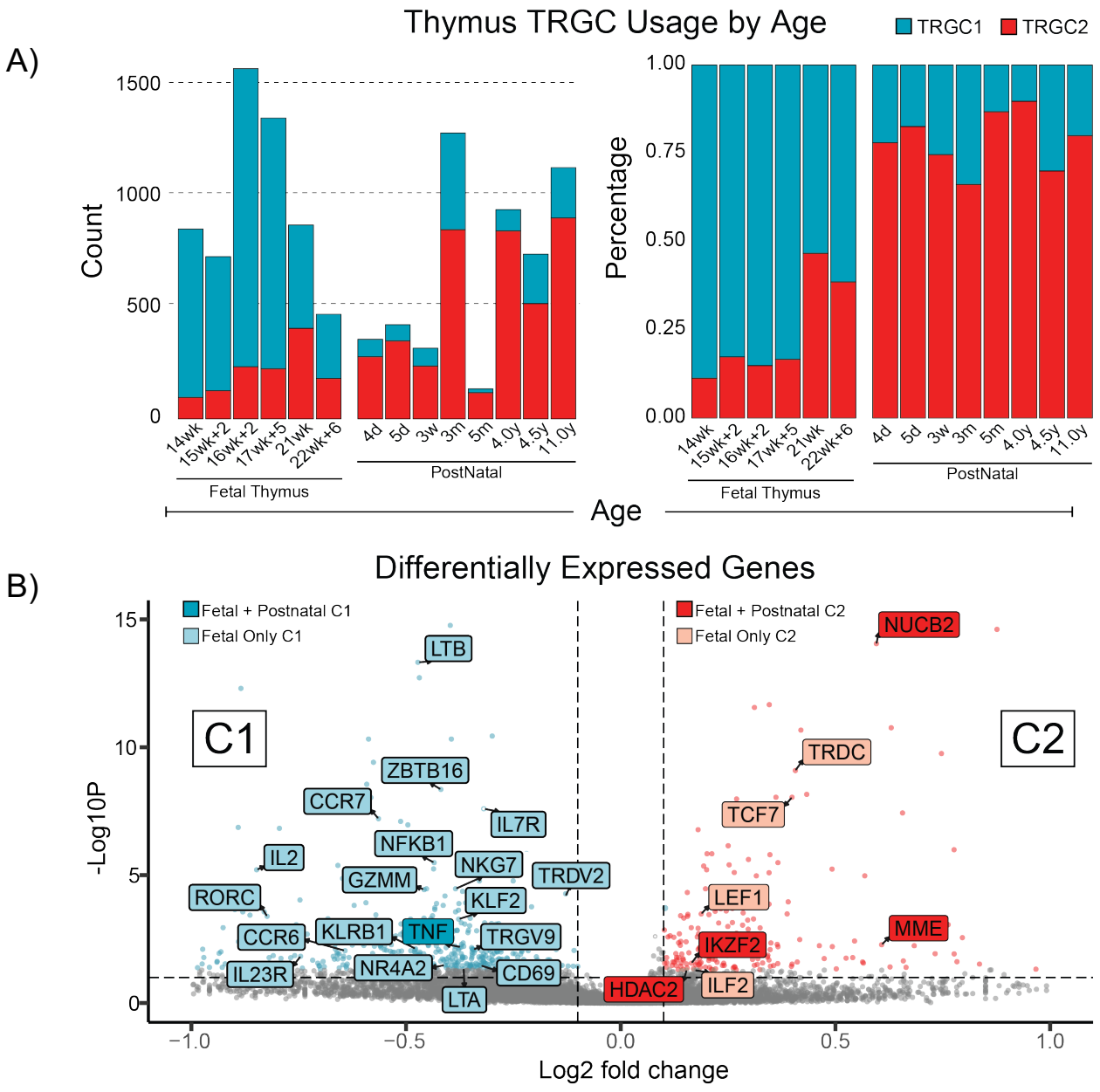

Figure 3.6: Thymus TRGC distribution and DEGs

A. Left: Bar graph illustrating counts of TRGC1 (blue) single positive or TRGC2 (red) single positive cells by age of patient. FT= Fetal Thymus, PNT= postnatal thymus, numbers indicate age in week+day, months, or years, as indicated. Right: Bar graph indicating percentage of total single positive cells that are TRGC1 or TRGC2 positive by age of patient. **B.** Volcano plot highlighting differentially expressed genes between TRGC1+ (blue, right) and TRGC2+ (red, left) cells in fetal thymus. Markers shared between fetal and postnatal thymus are shaded in darker red or darker blue.

TRGC1 cells in both fetal and postnatal thymus show evidence of maturation and cytotoxic differentiation programs, upregulating maturation/thymic egress genes *KLF2*[28,51] and *SELL* [27] respectively, and proinflammatory cytokine *TNF* (**Fig 3.6B**., **Fig. 3.7A**). Fetal TRGC1 cells show high T cell signaling and activation with *CD69*, *NFKB1*, *NR4A2*, and *KLRB1*[55], have a strong cytotoxic signature with high *GZMM*[56] and *NKG7*[56], and upregulate maturation/activation genes *CCR7*[27], *IL2*, and *IL7R*[56] (**Fig. 3.6B**). Postnatal thymic TRGC1s also show evidence of maturation with *LSP1*[57], *CD8B*, and *ID2*[58] (**Fig. 3.7A**). In sum, TRGC1 cells in the thymus collectively show evidence of maturation and differentiation into a cytotoxic program.

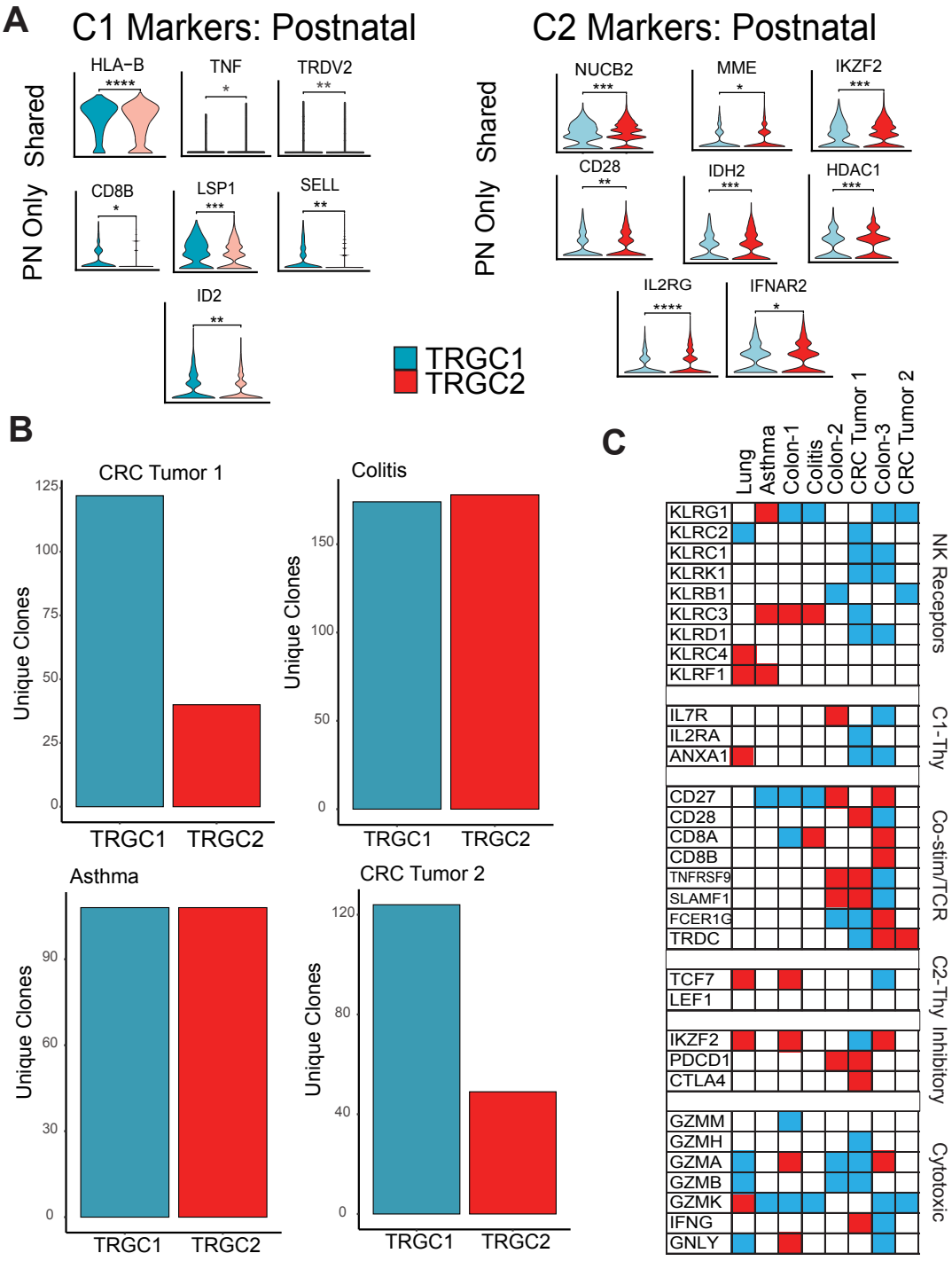

Figure 3.7: TRGC1 and TRGC2 cells have different profiles in postnatal thymus and normal peripheral tissues

A. Violin plots showing a subset of highly differentially expressed genes between TRGC1+ and TRGC2+ cells in postnatal thymus. Statistics show T-test with Bonferroni correction. *p<.05. **p<.01, ***p<.001, ****p<.0001. **B**. Bar graphs showing number of unique TRGC1 (blue) or TRGC2 (red) clones in the disease samples of each dataset. **C**. Summary visualization of a subset of significantly differentially expressed genes between TRGC1 and TRGC2 (p<.05 with Benjamini-Hochberg significance testing for all genes displayed). Differentially expressed genes were determined with Wilcoxon rank sum test using the FindMarkers function in the Seurat pipeline.

In addition to the maturation and cytotoxicity genes above, DEGs from TRGC1 cells from fetal thymus also specifically included *RORC*, *ZBTB16*, *TRDV2, and TRGV9*. When combined with the previously noted *CD69*, *KLRB1*, and *IL7R* (**Fig. 3.6B**) differences, these genes fit the signature for the Type 3 like early effectors described in human[28,59](**Fig. 3.6B**). These data indicate that they are also enriched for the highly antigen sensitive Cγ1 chain, which may contribute to TCR-driven effector programming in the thymus. This enrichment could also be due to chain pairing bias early in development when TRGV9 may preferentially recombine to proximal TRGJ1 or TRGJP1 rather than distal J segments at the locus[60,61].

In contrast, TRGC2 cells from all developmental stages are unified by programs of differentiation into naïve/central memory like phenotypes. Fetal TRGC2s express markers aligning with naive T cell differentiation, including *TCF7*[3,62,63] and *LEF1*[3,63] (**Fig. 3.6B**). In mouse γδ T cells, these transcription factors have also been shown to antagonize the γδ17 cell fate and promote Tγδ1 cell differentiation[6], but Type 1 markers were not broadly enriched in fetal thymic TRGC2s. Postnatal TRGC2+ cells show evidence of a central memory-like phenotype relative to TRGC1s, upregulating *IDH2*[64] and *CD28*[56], and evidence of maturation with cytokine receptors *IFNAR2* and *IL2RG*. Compared to their TRGC1 counterparts, both fetal and postnatal thymic TRGC2+ cells also highly express *IKZF2,* which is most well characterized as a driver of thymic Treg differentiation but has also shown to be highly expressed in IELs[65] (**Fig. 3.6B, Fig. 3.7A**). Taken together, these markers suggest thymic TRGC1s take on a more activated and cytotoxic phenotype while thymic TRGC2s upregulate naïve/memory cell differentiation and maintenance markers.

TRGV-TRGC pairing is biased in fetal thymus, but not postnatal thymus or consistently in peripheral tissues

We next wanted to determine if the phenotypic differences we observed in the thymus were due to differential V gene usage in the TRGC1 and TRGC2 compartments. We looked at the proportion of TRGV usage within TRGC1 and TRGC2 expressing cells to determine whether V gene pairing is biased. V gene - C gene pairing in fetal thymocytes has a significant chi-squared value, meaning TRGV and TRGC use are not independent in these cells. Pearson Residuals show the most significant deviations from expected value are for TRGC1 pairing with TRGV8 and TRGV9, and TRGC2 pairing with TRGV2, TRGV3, and TRGV5 (**Fig. 3.8A**). While the data are insufficient to conclude sequential recombination, the biased pairing of internal V genes with TRGC1 and external V genes with TRGC2 is consistent with this hypothesis and previous data[50,60]. This observation also corresponds with enrichment for the Vγ9Vδ2+ γδ17 phenotype observed in the fetal TRGC1 compartment specifically (**Fig. 3.7B**). However, in postnatal thymus, there are no statistically significant differences in TRGV usage between TRGC1 and TRGC2. This indicates that phenotypic differences seen in postnatal thymocytes are not driven by systematic biases in V gene pairing.

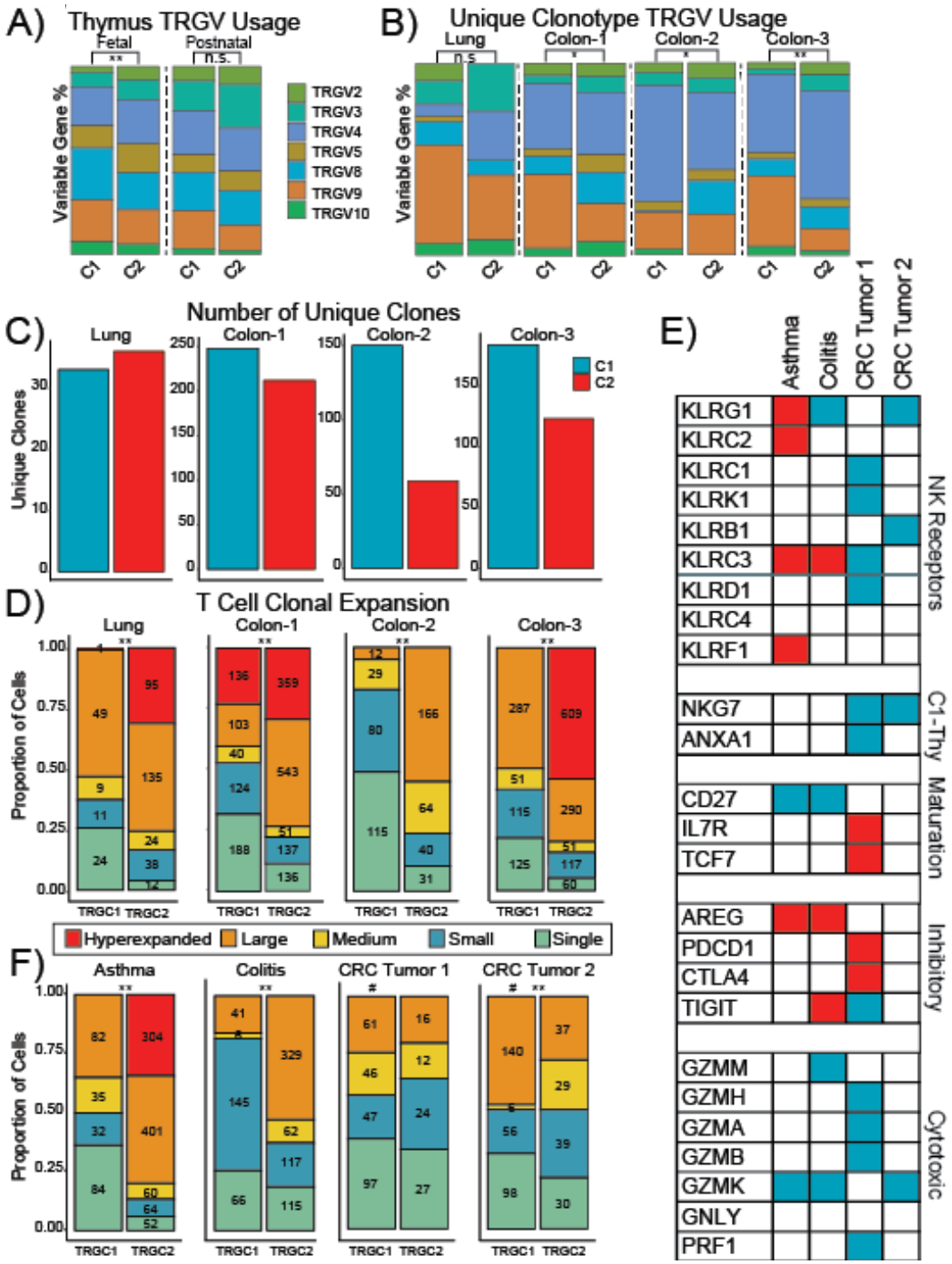

Figure 3.8: Clonal dynamics and gene expression in the periphery

A. Variable gene usage in TRGC1+ and TRGC2+ cells in fetal and postnatal thymus. Only the novel postnatal thymus dataset was used for this figure, for simplicity. *p< .05 by chi-squared testing for significance. **B.** TRGV usage by unique C1 and C2 clones in 3 different healthy colon datasets and 1 healthy lung dataset. Colon-1=Luoma et al., 2020[2], Colon-2=Reis et al., 2022[1], Colon-3=Novel CRC paired healthy tissue dataset. **C.** Number of unique C1 vs C2 clones in 3 healthy colon datasets and 1 healthy lung dataset. **D.** Clonal expansion of C1 and C2 binned by number of copies of each clone from 3 normal colon datasets and 1 non-asthmatic lung dataset. The numbers in the

bars indicate number of cells in that bin, y-axis indicates proportion of total cells.*p< .05 by chi-squared testing for significance. Hyperexpanded = 101-1000 copies, Large = 11-100 copies, Medium = 6-10, Small= 2-5, Single=1. **E.** Clonal expansion of C1 and C2 binned by number of copies of each clone from 2 colorectal cancer tumor datasets, 1 colitis dataset, and 1 asthma lung dataset. The numbers in the bars indicate number of cells in that bin, y-axis indicates proportion of total cells. *p< .05 by chi-squared testing for significance. # p<.05 by chi-squared testing for significance between normal dataset and tumor dataset. Hyperexpanded = 101-1000 copies, Large = 11-100 copies, Medium = 6-10, Small= 2-5, Single=1.**F**. Summary visualization of a subset of significantly differentially expressed genes between TRGC1 and TRGC2 (p<.05 with Benjamini-Hochberg significance testing for all genes displayed). Differentially expressed genes were acquired with Wilcoxon rank sum test using FindMarkers function in the Seurat pipeline.

However, the thymus is a relatively naïve environment, with fewer available antigens and almost no clonal expansion. In order to investigate TCR dynamics in a fully mature context, we used the same strategy described above, generating and analyzing our own datasets from colorectal cancer patient biopsies, as well as reanalyzing published datasets from asthma/non-asthmatic lung, colorectal cancer[4] and colitis[66]. γδ T cells from both asthma and non-asthmatic lung were sorted with CD1d tetramer and are therefore not an unbiased sample of all the γδ T cells present in the tissue, but analyses are internally consistent to the tetramer-positive cells collected and therefore still describe relevant TRGC associations with γδ T cell biology.

First, we examined TRGV usage by comparing unique TCRs across the tissue types. This analysis was conducted on unique clones to avoid biasing TRGV usage by clonal expansion, and therefore is reflective only of selective recruitment to the tissues, not clonal expansion within them. In all 3 adjacent normal colon datasets, representative of 16 patients total, we see enrichment of TRGV4 (**Fig. 3.8b**), a known binding partner to gut-restricted ligand Btnl3/8[12,67]. Across the lung and 2 out of 3 colon datasets, we again observed a trend towards pairing bias between TRGC1 and TRGV9; this pairing bias was not statistically significant in the lung as it was in the fetal thymus, demonstrating that this bias diminishes in significance as cells exit the thymus and enter in the periphery. In colon dataset 1 and 3, TRGV-TRGC pairing are not independent, a trend largely driven by TRGV9-TRGC1 pairing bias. However, in colon 2, the

significance is mostly driven by TRGV8-TRGC2 pairing bias, which contradicts the proximal recombination pattern we observed in the fetal thymus. Thus, while some TRGV segments preferentially pair with certain TRGC segments in the periphery, these biases are not systematic and likely reflect differences in tissue specific recruitment of γδ T cells.

TRGC2+ clones are the majority of unique TCRs in normal colon and are more clonally expanded in all normal tissues

To examine clonal dynamics in normal tissue, we examined total unique clones and clonal expansion in TRGC1 and TRGC2 cells. Analysis of constant chain use in total unique TCRs shows a larger number of unique TRGC2 TCRs in the lung, but more unique TRGC1 TCRs seeding the colon across the datasets (**Fig. 3.8c**). The slight TRGC2 bias in the lung could be attributed to unequal clonal sampling due to tetramer enrichment of γδ T cells in this dataset. In fact, data from a recent study suggest that the increased flexibility of Cγ2's longer CP may allow the TCR to bind more tetramer at the same TCR expression level[52], exacerbating this possible bias. Assessment of clonal expansion within the TRGC1 and TRGC2 compartments also reveals, surprisingly, that TRGC2 clones are more expanded in healthy tissue (**Fig. 3.8d**) despite the heightened activation efficiency granted by TRGC1 use **(Fig. 3.1a**).

Peripheral TRGC1+ clones retain thymic cytotoxic effector signature

To investigate peripheral phenotypes, we analyzed differentially expressed genes between TRGC1+ and TRGC2+ clones in normal peripheral tissues. In all normal colon datasets, TRGC1 cells differentially express both activating and inhibitory NK receptors (**Fig. 3.7C**), showing evidence of activation and effector status. In contrast, TRGC2s from non-asthmatic lungs show

differentially high NK receptor expression, suggesting the NK receptor phenotypes of these cells may be affected by tissue specific cues. In both normal colon and non-asthmatic lung, TRGC1s express high levels of granzymes, particularly granzyme B, with evidence of elevated *IFNG* signaling in one of the normal colon datasets as well, showing that TRGC1s across healthy tissues express cytotoxic and pro-inflammatory effector molecules. Thus, TRGC1s retain the differentiated cytotoxic program in the periphery that they acquired in the thymus.

TRGC1s retain their cytotoxic phenotype in tumor while TRGC1s and TRGC2s functionally converge in pro-inflammatory disease conditions

To understand if constant chain usage influences function in disease, we looked for significantly differentially expressed genes between TRGC1s and TRGC2s in each peripheral tissue disease condition. TRGC1 cells in asthma acquire significant NK receptor expression but lose significant differences in expression of activation genes and granzymes, suggesting phenotypic convergence between TRGC1 and TRGC2 cells in disease. Similarly, the colitis dataset showed very few significantly differentially expressed genes between TRGC1 and TRGC2 cells, suggesting functional convergence between TRGC1s and TRGC2s may be generalizable in pro-inflammatory conditions (**Fig. 3.8E**). The one exception to this functional convergence is the expression of wound healing cytokine *AREG* in the TRGC2s in both asthma and colitis, suggesting TRGC2s may participate in repairing tissue damage in these allergic/autoimmune conditions. In both colorectal tumor datasets, TRGC1s maintain heightened expression of similar genes to healthy colon TRGC1s, including NK receptors and effector granules like granzymes and perforin, suggesting that their differentiated effector function persists in the tumor. TRGC2s in tumors have fewer differentially expressed genes, with the notable exception of TRGC2s

derived from CRC dataset 1, which take on an immunosuppressive phenotype, expressing genes encoding co-inhibitory molecules including *PDCD1* and *CTLA4* (**Fig. 3.8E**).

TRGC1s preferentially expand in tumor tissue

In conjunction with differential gene expression analysis, we examined clonal expansion in colorectal tumor, colitis, and asthma tissues and found a pattern of γδ clonal expansion that was distinct from paired non-diseased tissues. In colorectal tumor tissues, TRGC1s were significantly more expanded (**Fig. 3.8F**) in the tumor compared with proximal non-diseased tissue, which had more clonal expansion in the TRGC2 compartment. In contrast, in both asthma and colitis tissues, there was higher expansion in the TRGC2 compartment. The number of unique clones seeding the disease tissues followed the same pattern seen in the healthy tissues (**Fig 3.7B**), suggesting that the differences in total clonal expansion are due to *in situ* expansion and are not strongly influenced by recruitment of additional cells from circulation.

Taken together, these data suggest that TRGC1s in both normal colon and colorectal tumor tissue maintain aspects of the thymic developmental program, including increased expression of activation and effector markers, and selectively expand in tumor conditions. In normal colon, the TRGC2s, which are less cytotoxic, are more clonally expanded, but the more cytotoxic TRGC1s are more clonally expanded in the tumor itself. In contrast, in colitis and asthma, TRGC1 and TRGC2 phenotypes seemingly converge and we do not see the concordant expansion in the TRGC1 compartment in disease, suggesting that the signaling milieu in these more autoimmune and allergic responses may functionally polarize both TRGC1 and TRGC2 cells and thus not

drive the selective expansion of TRGC1 cells we see in the immunosuppressive tumor microenvironment.

Discussion

In rearrangement of human β and γ TCR chains there is a "choice" between two genes encoding the constant domains. Conventional thought on β constant chain usage has assumed that this choice of constant domain is random, with no functional consequence or selective expansion in disease. In contrast, our work indicates a clear functional outcome to γ constant domain choice, with significant differences in CP length, signaling, and therefore activation as a consequence of γ constant usage. TCRs with the same antigen specificity signal strongest with Cγ1, and activation efficiency decreases with increasing numbers of repeats in the CP. The significant differences in activation between the 3 allelic variants are not explained by the presence or absence of the canonical CP-CP disulfide bond. These *in vitro* signaling differences translate to distinct transcriptional programs in the thymus that endure and influence clonal expansion in the periphery. Additionally, while fetal TRGV pairing shows some biases, TRGV usage does not neatly segregate with TRGC usage, showing that distribution of γ constant domain usage is broad across the γδ TCR repertoire and thereby suggesting a γδ T cell intrinsic function for differences in signaling based on γ constant domain choice.

Constant chain usage also segregates developing γδs into different thymic signatures that endure in the periphery, with Cγ1s showing evidence of increased activation and cytotoxicity across patient samples. Cγ1s also selectively clonally expand in tumor tissues, suggesting their increased activation efficiency and cytotoxic phenotype allows them to respond more strongly in

the tumor context than Cγ2s. In autoimmunity and allergy, however, Cγ2s maintain their higher level of clonal expansion, with functional convergence between Cγ1s and Cγ2s, suggesting that tissue and disease specific signals can override the intrinsic differences in activation efficiency in a pro-inflammatory context.

These data provide clarity to previous findings reporting chain pairing biases[46,49,50] and time-dependent switching of TRGC usage[60,61]. Our data show productive TCR pairing of Cγ1 and Cγ2 chains with both Vδ1 and Vδ2 in thymus and periphery, although there are enrichments in Vδ2 and Vγ9 pairing with Cγ1 in the fetal thymus. In accordance with published findings[51], our data show Cγ2-containing TCRs dominate postnatal thymic output. However, productive Cγ1 bearing TCRs dominate in fetal thymic γδs and are found in both blood and peripheral tissues postnatally. Our data reveal functional roles for both γ constant genes and that more heterogeneity and flexibility in chain pairing and developmental timelines exist than previously thought.

The γ constant chain differences we document here provide yet another dial for TCR to tune signal strength. Fine tuning of signal strength in the γδ T cell compartment may also be particularly important as many putative TCR ligands are self-ligands[8]. Functional outcomes of these T cells rely on a sum of different receptor signals to activate, polarize, and respond specifically to perturbation[35].

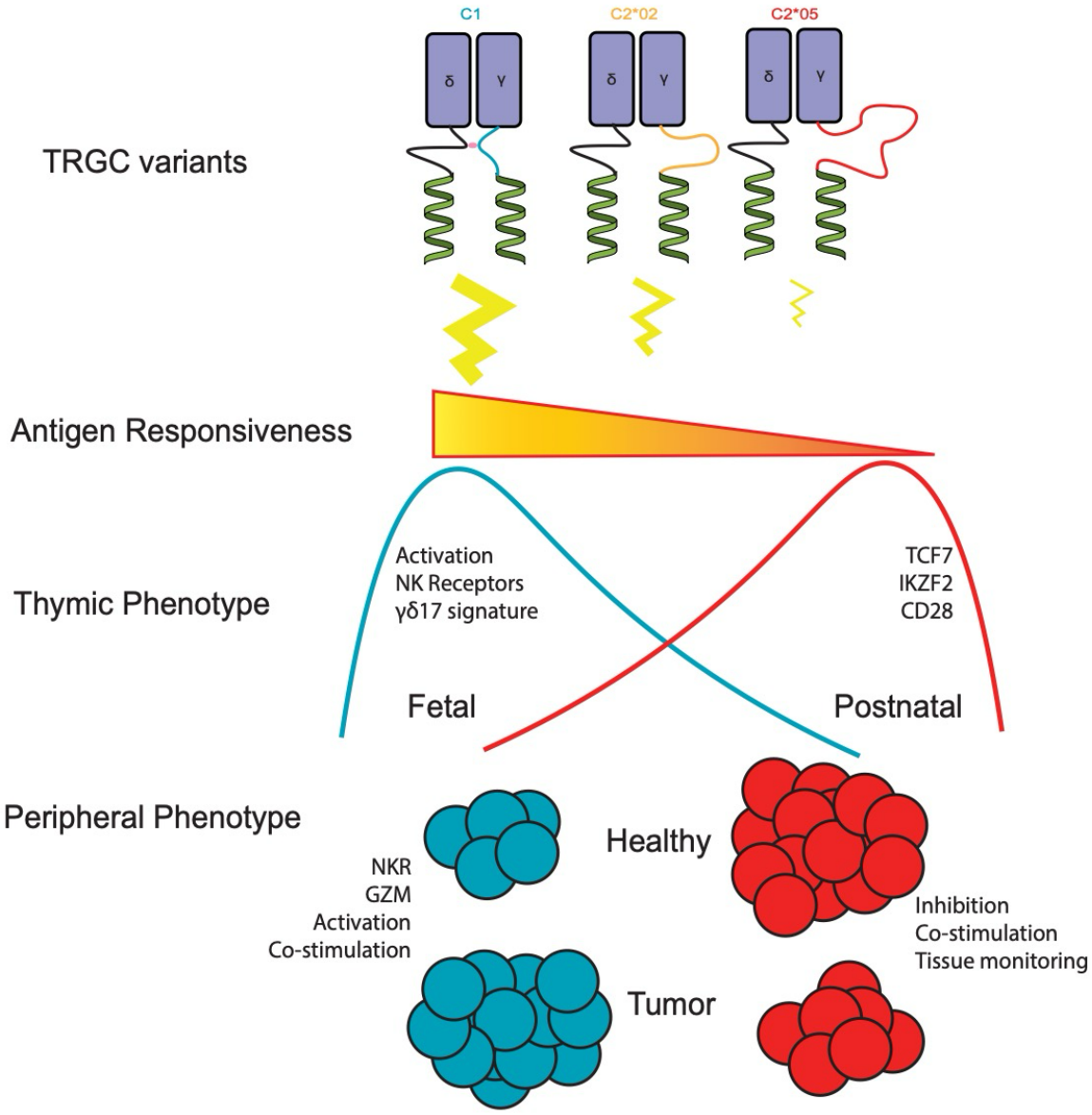

Figure 3.9: Model Summary

Our working model (Fig 3.9), based on the data shown here, is that TRGC1-bearing cells dominate the fetal thymic output, selectively pairing with proximal TRGV segments, most notably TRGV9. This selective pairing promotes the production of Vγ9Vδ2 γδ17-like cells, whose signature is enriched in fetal thymus Cγ1 differentially expressed genes. Thymic Cγ1s as a whole also have a more differentiated and activated profile, likely driven by the inherent

increase in activation efficiency granted by TRGC1 usage. In postnatal thymus, most cells use TRGC2, with diverse and largely unbiased TRGV pairing. These cells exit the thymus and populate peripheral tissues, showing TRGV enrichments reflective of selective recruitment to tissue specific antigens but not consistently biased TRGC-TRGV pairings. In healthy tissues, Cγ2s remain less activated and less cytotoxic, but are more clonally expanded. However, in some immunosuppressive disease conditions, like colorectal cancer tumor, Cγ1s selectively clonally expand. As we observed, Cγ1s respond both more strongly to the same antigen density and achieve a higher maximal activation, which may be favorable in conditions such as cancer to drive clonal expansion. Our data also show that while the phenotypic differences between Cγ1 and Cγ2 cells are established early in thymic development and largely persist in healthy peripheral tissue, disease specific signals can override these intrinsic differences and skew functional polarization in both compartments.

In this way, the functional and clonal dynamics of TRGC1 and TRGC2 γδ T cells provide evidence of another lever available to the human immune system for tuning activation efficiency in immunosurveillance. Further work on these dynamics could help to illuminate how γδ T cells integrate TCR modulation with other receptor signals and how TRGC usage segregates with functional polarization in other disease contexts. Selectively targeting the Cγ1 CP and blocking or boosting the activity of the most antigen-sensitive cells may also prove to be a valuable investigative and therapeutic tool, allowing manipulation of the strength of the immune response while maintaining diversity and antigen coverage of the TCR repertoire.

Chapter 4: Discussion

As the lineage-defining receptor, the γδ T cell receptor is assumed to play critical roles in γδ T cell development, maintenance, and peripheral function. However, the extent to which signals through the TCR determine γδ function and how signals through the TCR are interpreted by the γδ T cell throughout its lifespan is still an active area of research.

Many threads of evidence over decades of work have shown important roles for the γδ TCR in each stage of a T cell's life. In mice, γδ TCR signal strength determines both if a cell becomes a γδ T cell and what kind of γδ T cell it becomes. Several subsets of γδ T cells, most notably DETCs in mice and IELs in humans, show evidence of tonic TCR signaling at rest through engaging with self-antigen in the tissues. On the other hand, several studies have shown TCR-independent survival, trafficking, and activation of γδ T cells in certain contexts. The variety of receptors γδ T cells express, and the diversity of ligands, many of which have yet to be characterized, makes understanding how they sum TCR and non-TCR signals a significant challenge.

The data presented in this dissertation show not only that γδ T cells can exit the thymus as effectors, but they can even take on several different phenotypes, some unique to the γδ compartment with specialized receptors and functions. γδs exiting the thymus can also remain naïve, or take on a regulatory program, or even become antigen presenting cells. This diversity is preprogrammed in the thymus, before the cell is resident in any peripheral tissues or has been exposed to foreign antigens. Importantly, much of this functional pre-programming is orchestrated by γδ TCR signaling. After exiting the thymus, γδ T cells additionally acquire tissue specific adaptations, as described extensively in the IEL literature[4,5,12,65,68].

To participate in such specific functional programming, the TCR must be incredibly flexible, and this work has helped illustrate that the γδ TCR has many mechanisms to modulate signaling. Previous studies have characterized how δ chains can incorporate multiple D segments to dramatically boost repertoire diversity, or how a temporal switch in TdT activity provides an invariant fetal population of γδ TCRs and a much more diverse postnatal population, allowing both coverage of a wide range of antigens and rapid response from innate-like effectors with limited TCR diversity. In my work, I have uncovered yet another mechanism to tune γδTCR signal strength. Diversity in the constant region of the gamma chain provides yet another mechanism to modulate TCR signal strength. The particular benefit of having a constant chain CP that influences signal strength is γδ T cells as a population need not compromise on the diversity of the repertoire—the γδ T cells maintain antigen coverage while tuning signal strength by incorporating both constant domains into functional peripheral TCRs in all compartments.

Thymic selection and tolerance

My studies, taken together, question how "tolerance" is achieved in γδ T cells. It is already clear that γδ T cells have a variety of activation mechanisms—some of which follow the αβ paradigm of direct TCR signaling leading to activation and functional polarization, while others need additional signals of tissue stress to initiate a cytotoxic response. It therefore follows that the establishment of tolerance is unlikely to be one size fits all—if all autoreactive γδ TCR-expressing cells were deleted, there would be no γδ IELs in the periphery, but if every autoreactive TCR-expressing T cell become a cytotoxic killer cell, systemic autoimmunity would ensue. My data suggest TCR signal strength in the thymus plays a crucial role in this pre-programming. Perhaps the most strongly signaling cells become the most differentiated because

they had the strongest responses to self-ligand. Then these cells, the Type 1 like branch, seed the periphery, receive tonic TCR signals, and "bide their time" until tissue stress ligands spur them into action. Weaker TCR signals may be permissive to naïve cell differentiation, and these cells leave the thymus with a working TCR, but not a strongly autoreactive one, to wait for a foreign ligand to mature into effector status. The APC branch, which are also pro-inflammatory, have a high egress score, and high levels of TCR signaling. Where these branch off from the main differentiation trajectory is unclear, but perhaps diverting them into an APC cell fate is a way to salvage a cell that is otherwise too autoreactive through its TCR to be useful.

Once in the periphery, the cells recall their thymic programming in several ways. My work with the γ constant region surprised me in how consistently TRGC usage mapped with γδ T cell cytotoxicity, an association that persisted even in disease. However, the highly inflammatory allergic and autoimmune conditions showed that an extreme signaling milieu can override TCR determinism in the right context.

Flexible TCR complex assembly

The structural mechanism of TRGC tuning of signal strength is another outstanding question. Recent structural work on the γδ TCR-CD3 complex suggested both the accessibility of ligand and flexibility of the receptor were affected by CP length. Their model posited that Cγ2 offers more access to ligand, leading to higher amounts of tetramer binding, but the decreased rigidity ultimately reduced the signal transduction of TCR-ligand binding across the membrane. Our work also suggests that the number of repeats correlates directly with signal strength after controlling for ligand density, suggesting the distance from the membrane or flexibility of the extracellular portion of the TCR may determine signal strength. Additional possibities include

differences in CD3 complex constitution between Cγ1 and Cγ2, where each CP length may preferentially associate with certain CD3 family members. This possibility is less likely however, given that in a Jurkat system where signaling differences between Cγ1 and Cγ2 are present at the same antigen density, anti-CD3 does not recapitulate those signaling differences (Fig. 3.2A.). Thus, the most likely model for CP driven differences in signal strength is that the length or disorder of the connecting peptide decreases the ability of the TCR to transmit the conformational change associated with binding ligand across the cell membrane.

"Adaptate" effectors

The addition of the constant region lever in tuning signal strength adds a layer of regulation to the "adaptate" model of γδ T cell function. First proposed by Adrian Hayday[1], this model describes three paradigms of γδ TCR antigen recognition that all exist in parallel in the human immune system. The first is truly adaptive recognition, where γδ TCRs recognize antigens using their CDR loops, which drives clonal expansion and functional differentiation. An observed example of this type of γδ T cell recognition is found in CMV+ people, who show evidence of repertoire narrowing and effector differentiation concurrent with clonal expansion of a diverse population of γδ T cells. The intermediate recognition model is "quasi-adaptive" recognition of a limited pool of ligands by a restricted pool of TCRs, which is evidenced by expansion of γδ T cells with limited TCR diversity. The final category is "innate" recognition by the TCR, in which fixed chains recognize known self ligands, like in the Vg4+ compartment in human intestine or Vγ9Vδ2 recognition of BTN/phosphoantigen.

This dissertation extends this model to include modulation of signal strength beyond antigen recognition and affinity by the variable region of the TCR. My work shows that Cγ1

preferentially recombines with Vγ9, enriching Vγ9Vδ2+ cells for Cγ1 usage. This suggests that the more potent constant chain frequently pairs with the less diverse TCR repertoire. This preferential boosting of signal in "innate-like" recognition may be an evolutionary mechanism to tune antigen responsiveness in the compartment with stereotyped antigens and lower risk of cross-reactivity and autoimmunity. The preferential recombination of Cγ2 with more diverse "adaptive-like" Vδ1 TCRs may conversely reflect a strategy of dampening signal strength in the TCR compartment that has higher risk of self and cross-reactivity, providing an additional mechanism for maintaining tolerance in adaptate γδ T cell recognition.

Layered immune system

Piecing together the levels of regulation reveals a model of γδ T cell development and differentiation that fits neatly into the hypothesis of the layered immune system originally proposed by Leonore and Leonard Herzenberg, which posits that lymphocytes develop in waves with distinct developmental origins. This hypothesis was originally proposed to explain the development of B-1 cells, which exhibit broad low affinity reactivity and, at least in mice, develop from distinct precursors[69].

It is tempting to situate γδ T cells within this paradigm to organize their distinct trajectories and functions into waves of development. First, in fetal thymus, we see the development of a more public TCR repertoire, with larger proportions of Vγ9Vδ2 cells, and complete effector differentiation before thymic egress. These cells can become Type-1, 2, or 3-like effectors before exiting the thymus. My work shows that these cells also preferentially recombine with the strongly signaling Cγ1 constant domain, providing an early wave of differentiated effectors with relatively restricted diversity and boosted TCR signal strength. These cells likely perform broad

immunosurveillance while the conventional lymphocyte compartments are still developing. Then, the postnatal thymus produces γδ T cells with evidence of TCR-driven fate commitment, but with more naïve cells exiting the thymus. This diverse repertoire of cells preferentially recombines with Cγ2, providing a more tunable, adaptive-like compartment of postnatal γδ T cells to perform complementary functions to the conventional adaptive immune system. These cells may largely perform adaptive antigen recognition, but may require additional signals of tissue stress, like cytokines or NK ligands to fully activate, allowing them to perform specific immunosurveillance in the tissues. Vδ1s also endure in the blood and tissues much later in human life, while the Vγ9Vδ2 population declines, further supporting this hypothesis of a layered γδ T cell compartment[70]. Thus, the "adaptate" function of γδ T cells as a whole, as well as specific patterns of development and TCR recombination at specific time points allows γδ T cells to fill different immunological niches throughout an organism's life.

There are still many outstanding questions in understanding γδ T cell thymic development, peripheral function, and how the TCR contributes to various γδ T cell fates. My thesis work provides several benchmarks in building the map of human γδTCR structure and function, and provides detail and nuance in understanding how these cells detect and respond to signals through their TCR. Finally, my work shows both the centrality and flexibility of the TCR in translating signal strength and programming γδ T cell function, providing novel insights into the strength and tunability of these signals throughout the γδ T cell lifespan.

Chapter 5: Materials and Methods

CD1d protein generation

CD1d protein for all experiments was made recombinantly in Hi5 insect cells using a baculovirus based expression system. The CD1d construct has a C-terminal Avitag for biotinylation, a His-tag for purification, and a 3C cleavage site for removing the His tag before tetramerizing. CD1d was co-infected with untagged b2m and BirA virus to biotinylate protein in vivo, then purified over Ni resin. Protein was 3C treated to cleave the His tag, then cleaned up over Ni resin. Ni flow through was then purified by size exclusion chromatography on a SuperDex 200i column. Pure protein was incubated with a 10-fold molar excess of either mixed brain sulfatides (Avanti Polar Lipids) or PBS-57 (gift from Dr. Paul Savage) for 1 hour at 37C or overnight at 4C. After incubation, excess lipid was diluted with a 10x buffer volume, then the protein was buffer exchanged and concentrated to obtain pure monomer. Successful biotinylation of protein was tested using traptavidin incubated for 10 minutes with protein, then SDS Page gel to ensure multimerization. MR-1 was kindly provided by Nicole Ladd and prepared as previously described in her work.

Tetramer generation

Biotinylated CD1d-b2m and MR1-β2m were incubated with Streptavidin-PE (Agilent) at a ratio of 4.4:1, to achieve 1.1 molar excess of each of the 4 biotin binding sites on each streptavidin molecule. SA-PE was added in 8 equal batches with 5 minutes in between each addition to ensure complete tetramer formation. Oligonucleotide barcoded tetramers for 10x experiments were produced using the same protocol, but Streptavidin-PE was also coupled to an

oligonucleotide barcode before making tetramer. Barcodes were ordered with a 5' amine modification to allow coupling to streptavidin using the Protein-Oligonucleotide Conjugation kit from Vector Labs, which uses a S-Hynic/4FB reaction to couple. After barcode conjugation, free oligonucleotide was purified away from protein using spin columns, then purified again using size exclusion chromatography.

Thymus tissue processing

Thymocytes were collected in collaboration with Dr. Andrew Koh and Dr. Narutoshi Hibino from patients undergoing heart surgery (IRB: 2020-203). Dead cells were separated using Ficoll density gradient and thymocytes were counted and frozen in FBS + 10% DMSO until thawing for staining and sequencing.

Colon tissue collection and processing for single cell analyses (done by Caitlin Castro)

Leftover surgical tissue samples collected at UCMC were collected from patients who consented to research under University of Chicago IRB# 10-209-A and then processed by the HTRC biobanking core. Lutheran General samples were leftover surgical tissue samples collected from patients who consented to research under AdvocateAurora IRB# 6793, which is linked to Dr. Bissonnette's 10-209-A study. Additional samples were collected from leftover surgical tissue at UCMC under IRB# 17-0882. Colon tissue was collected into MACS Tissue Storage Solution (Miltenyi 130-100-008) at 4°C. Tissue was then cut into 1-5mm fragments, before transferring up to 4 x 5 mg slices into vials containing 1 mL freezing media (50:40:10 FBS:RPMI:DMSO). To thaw, vials were warmed for minimum necessary time in 37°C bath, then immediately transferred to 50 mL conical tube filled with RPMI, and centrifuged at 400 xg for 5 minutes.

Tissue was dissociated as described[66], briefly each pellet was resuspended in 5 mL of collagenase-containing enzymatic digestion solution prepared according to manufacturer's instructions (Miltenyi Human Tumor Dissociation Kit #30-095-929), prewarmed to 37°C. Tissue was incubated at 37°C for 20 min with gentle vortexing at 2–3-minute intervals. Dissociated cells were transferred to 70 μm nylon mesh filter, and filtered cells were centrifuged and resuspended in 3 mL ACK red blood cell lysis buffer (Gibco #A1049201) for 1 minute at RT, then immediately diluted with 10 mL of PBS with 1% FBS. Cells were centrifuged and subsequently resuspended in Live/dead Fixable Near-IR Dead Cell Stain Kit (Thermo # L34975) viability dye solution (1:1000 in PBS), followed by 15 min incubation at RT in the dark. Following viability dye staining, cells were washed and stained for sorting and single-cell RNA sequencing.

Flow cytometry and sorting for single cell libraries

Cells were washed and resuspended in FACS buffer (2% fetal bovine serum (FBS, Gemini) in sterile Ca2+ and Mg2+ free PBS). To prevent nonspecific Fc binding , cells were blocked with 10% heat-inactivated human serum in FACS buffer. For relevant experiments, cells were stained with a titrated concentration (~100nM, with some batch to batch variation) of tetramer in blocking buffer for 30 minutes at 25ºC. Finally, cells were stained with the indicated antibody panels diluted in FACS buffer for 30 minutes at 4ºC. Data was collected on a BD Fortessa 4-15 or LSR II 4-15. Sorting was done on a BD FACSAria Fusion, MACSQuant Tyto, or BD Symphony.

For the purposes of these studies thymocytes and IELs were stained with a panel of Total-seq-C conjugated antibodies including (CD4, CD8, CD45RA, CCR7, TCRγδ, CD62L, isotype) and PE-conjugated oligonucleotide barcoded tetramers (CD1d-sulfatide, CD1d-PBS57, CD1d-unloaded, CD1d-K152A, CD1d-W160A, MR1-unloaded, MR1-msmeg, MR1-5OPRU). Thymocytes were sorted on Vδ1 antibody (panel: Vδ1 FITC (Invitrogen # TCR2730), CD45 BV510 (BioLegend #304036), CD3 PE-Cy7, Live/Dead Near IR) and colon IELs were sorted using the same antibodies above but with the addition of Vδ2 FITC (BioLegend #331406), CD14 APC/Cy7 (BioLegend #301824) (dump gate) and PE-labeled tetramers. After sorting, cells were resuspended and processed for sequencing.

Library construction and RNA sequencing for single cell libraries

Samples were processed for sequencing using the 10X Genomics 5' Single Cell Immune Profiling workflow. Libraries for surface antibody barcode, αβTCR VDJ, γδTCR VDJ, and gene expression were generated. γδ TCR libraries were generated using the 10X universal forward primers and nested custom reverse primers (TRG outer Reverse: AGTCTTCATGGTGTTCCCCTCC, TRG inner Reverse: AAAATAGTGGGCTTGGGGGA, TRD outer Reverse GTCGTGTTGAACTGAACATGTCAC, TRD inner Reverse CAGACAAGCGACATTTGTTCCA). Libraries for all samples were pooled and sequenced on a NovaSeq6000 or NovaX (Illumina) and resultant FASTQs were demultiplexed by the University of Chicago Genomics Facility.

Single cell data analyses for γ Constant project

Sequences were deconvoluted using Cellranger v7.0 (for thymus datasets) or Cellranger v.3.0 (for colon datasets) and imported into R. Initial QC and gene expression and clustering analysis was conducted using the standard Seurat v5 pipeline[71]. In the thymus dataset, cells were well integrated with the exception of 1 cluster composed entirely of cells from a single sample. This cluster had no notable differentially expressed genes and clustered away from all of the other γδ thymocytes, and was therefore removed from further analysis to avoid confounding the results with cells from a single sample. The datasets were then subsetted to only contain cells expressing productive γδ T cell receptors. TRGC1 single positive cells were defined as having normalized RNA counts (data slot of the Seurat object) > 0 and TRGC2 normalized counts <= 0. Cells expressing both TRGC1 and TRGC2 >0 were classified as double positive. Cells lacking expression of either were classified as NA. Individual cells were then reclassified as TRGC1 if the plurality of cells in the clonotype were classified as TRGC1 by gene expression. TRGC2 clones were corrected the same way, and clones with a plurality of TRGC double positive cells were discarded. Differentially expressed genes were determined using the FindMarkers() function from Seurat with -log2FC cutoff at .1 paired with Benjamini-Hochberg significance testing to correct for false discovery rate.

T cell receptor analysis was performed using the scRepertoire package[72]. Additional analyses and visualizations were made using ggplot2 and base R tools. T cell repertoire and clonal expansion proportion plots were significance tested using chi-squared testing, treating cells as independent.

Single Cell Data Analysis for Thymus Development Project

After initial QC and γδ subsetting as described above, Seurat function FindAllMarkers() was used to identify top differentially expressed genes and label cluster identities. Tetramer positive cells were gated on a per cell basis, thresholded on a minimum of 1.5 normalized counts for the tetramer and at least 2x normalized counts for that tetramer over the others. CD1d multi+ were gated the same way, but only gated over MR-1 and isotype controls. Then pseudotime trajectory analysis was performed using Monocle 3[73], then integrated into Seurat metadata. Gene set scores were calculated using the AddModuleScore() function in Seurat. Graph vs graph analysis was performed independently of the Seurat object and with package specific nearest neighbor analysis in python, using the CoNGA[42] package.

Additional TCR analysis including motif analysis was performed using the scRepertoire package in R.

TRGC transcript PCR

Healthy PBMCs from 2 donors were sorted on Vδ1 or Vδ2 expression and resuspended in 1mL TRI-zol reagent (Thermo Scientific). Previously generated and cryopreserved IEL lines were thawed, rested overnight and resuspended in 1mL TRI-zol and RNA was extracted by chloroform/isopropanol precipitation. After cDNA generation with the Maxima cDNA synthesis kit (Thermo Scientific), TCR γ constant region was amplified using DreamTaq (Thermo Scientific).

Appendix I: Lineage Defining Gene Sets

Egress score (Sanchez et al., Nature Comms)	**Immediate Early Gene (IEG) score (self curated with Jackie Liao, Ciofani Lab)**	**Type 1 score (CTL score (Rat et al. 2020 Science Advances: Table S3)**
KLF2	*JUN*	*CST7*
CORO1A	*JUNB*	*GZMA*
CCR7	*FOS*	*GZMB*
CXCR4	*FOSB*	*IFNG*
CXCR6	*FOSL1*	*NKG7*
FOXO1	*NFKB1*	*PRF1*
CXCR3	*NFKB2*	*TNFSF10*
S1PR1	*REL*	*KLRG1*
S1PR4	*NR4A1*	*FASLG*
S100A4	*NR4A2*	*CCL3*
S100A6	*EGR1*	*GNLY*
EMP3	*EGR2*	*KLRB1*
	EGR3	*GZMM*
	EGR4	*GZMK*
	HLX	*CTSW*
	HOXB7	*NCR3*
	MYC	
	IRF1	
	ID3	
	GEM	
	DUSP2	
	TRAF1	
	CD69	
	IL2RA	
	JAK3	

Bibliography

1. Hayday, A.C. (2019). γδ T Cell Update: Adaptate Orchestrators of Immune Surveillance. J Immunol *203*, 311-320. 10.4049/jimmunol.1800934.

2. Bonneville, M., O'Brien, R.L., and Born, W.K. (2010). Gammadelta T cell effector functions: a blend of innate programming and acquired plasticity. Nat Rev Immunol *10*, 467-478. nri2781 [pii]
10.1038/nri2781.

3. McMurray, J.L., von Borstel, A., Taher, T.E., Syrimi, E., Taylor, G.S., Sharif, M., Rossjohn, J., Remmerswaal, E.B.M., Bemelman, F.J., Vieira Braga, F.A., et al. (2022). Transcriptional profiling of human Vδ1 T cells reveals a pathogen-driven adaptive differentiation program. Cell Rep *39*, 110858. 10.1016/j.celrep.2022.110858.

4. Reis, B.S., Darcy, P.W., Khan, I.Z., Moon, C.S., Kornberg, A.E., Schneider, V.S., Alvarez, Y., Eleso, O., Zhu, C., Schernthanner, M., et al. (2022). TCR-Vγδ usage distinguishes protumor from antitumor intestinal γδ T cell subsets. Science *377*, 276-284. 10.1126/science.abj8695.

5. Harmon, C., Zaborowski, A., Moore, H., St Louis, P., Slattery, K., Duquette, D., Scanlan, J., Kane, H., Kunkemoeller, B., McIntyre, C.L., et al. (2023). γδ T cell dichotomy with opposing cytotoxic and wound healing functions in human solid tumors. Nat Cancer *4*, 1122-1137. 10.1038/s43018-023-00589-w.

6. Fahl, S.P., Contreras, A.V., Verma, A., Qiu, X., Harly, C., Radtke, F., Zúñiga-Pflücker, J.C., Murre, C., Xue, H.H., Sen, J.M., and Wiest, D.L. (2021). The E protein-TCF1 axis controls γδ T cell development and effector fate. Cell Rep *34*, 108716. 10.1016/j.celrep.2021.108716.

7. Parker, M.E., and Ciofani, M. (2020). Regulation of γδ T Cell Effector Diversification in the Thymus. Frontiers in Immunology *11*. 10.3389/fimmu.2020.00042.

8. Castro, C.D., Boughter, C.T., Broughton, A.E., Ramesh, A., and Adams, E.J. (2020). Diversity in recognition and function of human γδ T cells. Immunol Rev *298*, 134-152. 10.1111/imr.12930.

9. Ribot, J.C., Lopes, N., and Silva-Santos, B. (2021). γδ T cells in tissue physiology and surveillance. Nature Reviews Immunology *21*, 221-232. 10.1038/s41577-020-00452-4.

10. Karunakaran, M.M., Willcox, C.R., Salim, M., Paletta, D., Fichtner, A.S., Noll, A., Starick, L., Nöhren, A., Begley, C.R., Berwick, K.A., et al. (2020). Butyrophilin-2A1 Directly Binds Germline-Encoded Regions of the Vγ9Vδ2 TCR and Is Essential for Phosphoantigen Sensing. Immunity *52*, 487-498.e486. 10.1016/j.immuni.2020.02.014.

11. Willcox, C.R., Vantourout, P., Salim, M., Zlatareva, I., Melandri, D., Zanardo, L., George, R., Kjaer, S., Jeeves, M., Mohammed, F., et al. (2019). Butyrophilin-like 3 Directly Binds a Human Vγ4(+) T Cell Receptor Using a Modality Distinct from Clonally-Restricted Antigen. Immunity *51*, 813-825.e814. 10.1016/j.immuni.2019.09.006.

12. Mayassi, T., Ladell, K., Gudjonson, H., McLaren, J.E., Shaw, D.G., Tran, M.T., Rokicka, J.J., Lawrence, I., Grenier, J.C., van Unen, V., et al. (2019). Chronic Inflammation Permanently Reshapes Tissue-Resident Immunity in Celiac Disease. Cell *176*, 967-981.e919. 10.1016/j.cell.2018.12.039.

13. Luoma, A.M.e.a. (2013). Crystal Structure of Vδ1 T Cell Receptor in Complex with CD1d-Sulfatide Shows MHC-like Recognition of a Self-Lipid by Human γδ T Cells Immunity *39*.

14. Bai, L., Picard, D., Anderson, B., Chaudhary, V., Luoma, A., Jabri, B., Adams, E.J., Savage, P.B., and Bendelac, A. (2012). The majority of CD1d-sulfatide-specific T cells in human blood use a semiinvariant Vdelta1 TCR. Eur J Immunol *42*, 2505-2510. 10.1002/eji.201242531.

15. Uldrich, A.P., Le Nours, J., Pellicci, D.G., Gherardin, N.A., McPherson, K.G., Lim, R.T., Patel, O., Beddoe, T., Gras, S., Rossjohn, J., and Godfrey, D.I. (2013). CD1d-lipid antigen recognition by the γδ TCR. Nature Immunology *14*, 1137-1145. 10.1038/ni.2713.

16. Ravens, S., Schultze-Florey, C., Raha, S., Sandrock, I., Drenker, M., Oberdörfer, L., Reinhardt, A., Ravens, I., Beck, M., Geffers, R., et al. (2017). Human γδ T cells are quickly reconstituted after stem-cell transplantation and show adaptive clonal expansion in response to viral infection. Nature Immunology *18*, 393-401. 10.1038/ni.3686.

17. Davey, M.S., Willcox, C.R., Joyce, S.P., Ladell, K., Kasatskaya, S.A., McLaren, J.E., Hunter, S., Salim, M., Mohammed, F., Price, D.A., et al. (2017). Clonal selection in the human Vδ1 T cell repertoire indicates γδ TCR-dependent adaptive immune surveillance. Nature Communications *8*, 14760. 10.1038/ncomms14760.

18. Murphy, K.M., and Weaver, C. (2017). Janeway's immunobiology, 9th Edition (New York : Garland Science).

19. Di Marco Barros, R., Roberts, N.A., Dart, R.J., Vantourout, P., Jandke, A., Nussbaumer, O., Deban, L., Cipolat, S., Hart, R., Iannitto, M.L., et al. (2016). Epithelia Use Butyrophilin-like Molecules to Shape Organ-Specific γδ T Cell Compartments. Cell *167*, 203-218.e217. 10.1016/j.cell.2016.08.030.

20. Boyden, L.M., Lewis, J.M., Barbee, S.D., Bas, A., Girardi, M., Hayday, A.C., Tigelaar, R.E., and Lifton, R.P. (2008). Skint1, the prototype of a newly identified immunoglobulin superfamily gene cluster, positively selects epidermal γδ T cells. Nature Genetics *40*, 656-662. 10.1038/ng.108.

21. Hayes, S.M., Li, L., and Love, P.E. (2005). TCR signal strength influences alphabeta/gammadelta lineage fate. Immunity *22*, 583-593. 10.1016/j.immuni.2005.03.014.

22. Haks, M.C., Lefebvre, J.M., Lauritsen, J.P., Carleton, M., Rhodes, M., Miyazaki, T., Kappes, D.J., and Wiest, D.L. (2005). Attenuation of gammadeltaTCR signaling efficiently diverts thymocytes to the alphabeta lineage. Immunity *22*, 595-606. 10.1016/j.immuni.2005.04.003.

23. Lauritsen, J.P.H., Wong, G.W., Lee, S.-Y., Lefebvre, J.M., Ciofani, M., Rhodes, M., Kappes, D.J., Zúñiga-Pflücker, J.C., and Wiest, D.L. (2009). Marked Induction of the Helix-Loop-Helix Protein Id3 Promotes the Gamma Delta T Cell Fate and Renders Their Functional Maturation Notch Independent. Immunity *31*, 565-575. 10.1016/j.immuni.2009.07.010.

24. Zuberbuehler, M.K., Parker, M.E., Wheaton, J.D., Espinosa, J.R., Salzler, H.R., Park, E., and Ciofani, M. (2019). The transcription factor c-Maf is essential for the commitment of IL-17-producing γδ T cells. Nature Immunology *20*, 73-85. 10.1038/s41590-018-0274-0.

25. Sumaria, N., Grandjean, C.L., Silva-Santos, B., and Pennington, D.J. (2017). Strong TCRgammadelta Signaling Prohibits Thymic Development of IL-17A-Secreting Gamma Delta T Cells. Cell Reports *19*, 2469-2476. 10.1016/j.celrep.2017.05.071.

26. In, T.S.H., Trotman-Grant, A., Fahl, S., Chen, E.L.Y., Zarin, P., Moore, A.J., Wiest, D.L., Zúñiga-Pflücker, J.C., and Anderson, M.K. (2017). HEB is required for the specification of fetal IL-17-producing γδ T cells. Nat Commun *8*, 2004. 10.1038/s41467-017-02225-5.

27. Boehme, L., Roels, J., and Taghon, T. (2022). Development of γδ T cells in the thymus - A human perspective. Semin Immunol *61-64*, 101662. 10.1016/j.smim.2022.101662.

28. Sanchez Sanchez, G., Papadopoulou, M., Azouz, A., Tafesse, Y., Mishra, A., Chan, J.K.Y., Fan, Y., Verdebout, I., Porco, S., Libert, F., et al. (2022). Identification of distinct functional thymic programming of fetal and pediatric human γδ thymocytes via single-cell analysis. Nature Communications *13*, 5842. 10.1038/s41467-022-33488-2.

29. Chodaczek, G., Papanna, V., Zal, M.A., and Zal, T. (2012). Body-barrier surveillance by epidermal γδ TCRs. Nat Immunol *13*, 272-282. 10.1038/ni.2240.

30. Jameson, J., Ugarte, K., Chen, N., Yachi, P., Fuchs, E., Boismenu, R., and Havran, W.L. (2002). A role for skin gammadelta T cells in wound repair. Science *296*, 747-749. 10.1126/science.1069639.

31. McDonald, B.D., Jabri, B., and Bendelac, A. (2018). Diverse developmental pathways of intestinal intraepithelial lymphocytes. Nature Reviews Immunology *18*, 514-525. 10.1038/s41577-018-0013-7.

32. Park, J.H., and Lee, H.K. (2021). Function of γδ T cells in tumor immunology and their application to cancer therapy. Experimental & Molecular Medicine *53*, 318-327. 10.1038/s12276-021-00576-0.

33. de Vries, N.L., van de Haar, J., Veninga, V., Chalabi, M., Ijsselsteijn, M.E., van der Ploeg, M., van den Bulk, J., Ruano, D., van den Berg, J.G., Haanen, J.B., et al. (2023). γδ T cells are effectors of immunotherapy in cancers with HLA class I defects. Nature *613*, 743-750. 10.1038/s41586-022-05593-1.

34. Silva-Santos, B., Mensurado, S., and Coffelt, S.B. (2019). γδ T cells: pleiotropic immune effectors with therapeutic potential in cancer. Nature Reviews Cancer *19*, 392-404. 10.1038/s41568-019-0153-5.

35. Vantourout, P., and Hayday, A. (2013). Six-of-the-best: unique contributions of γδ T cells to immunology. Nat Rev Immunol *13*, 88-100. 10.1038/nri3384.

36. Tieppo, P., Papadopoulou, M., Gatti, D., McGovern, N., Chan, J.K., Gosselin, F., Goetgeluk, G., Weening, K., Ma, L., and Dauby, N. (2019). The human fetal thymus generates invariant effector γδ T cells. Journal of Experimental Medicine *217*, e20190580.

37. Thelen, F., and Witherden, D.A. (2020). Get in Touch With Dendritic Epithelial T Cells! Frontiers in Immunology *11*. 10.3389/fimmu.2020.01656.

38. McKenzie, D.R., Hart, R., Bah, N., Ushakov, D.S., Muñoz-Ruiz, M., Feederle, R., and Hayday, A.C. (2022). Normality sensing licenses local T cells for innate-like tissue surveillance. Nature Immunology *23*, 411-422. 10.1038/s41590-021-01124-8.

39. Hu, T., Gimferrer, I., and Alberola-Ila, J. (2011). Control of early stages in invariant natural killer T-cell development. Immunology *134*, 1-7. 10.1111/j.1365-2567.2011.03463.x.

40. Odagiu, L., May, J., Boulet, S., Baldwin, T.A., and Labrecque, N. (2020). Role of the Orphan Nuclear Receptor NR4A Family in T-Cell Biology. Front Endocrinol (Lausanne) *11*, 624122. 10.3389/fendo.2020.624122.

41. Kelly, K., and Siebenlist, U. (1995). Immediate-early genes induced by antigen receptor stimulation. Current Opinion in Immunology *7*, 327-332. https://doi.org/10.1016/0952-7915(95)80106-5.

42. Schattgen, S.A., Guion, K., Crawford, J.C., Souquette, A., Barrio, A.M., Stubbington, M.J.T., Thomas, P.G., and Bradley, P. (2022). Integrating T cell receptor sequences and

transcriptional profiles by clonotype neighbor graph analysis (CoNGA). Nature Biotechnology *40*, 54-63. 10.1038/s41587-021-00989-2.

43. Dong, D., Zheng, L., Lin, J., Zhang, B., Zhu, Y., Li, N., Xie, S., Wang, Y., Gao, N., and Huang, Z. (2019). Structural basis of assembly of the human T cell receptor–CD3 complex. Nature *573*, 546-552. 10.1038/s41586-019-1537-0.

44. Horna, P., Weybright, M.J., Ferrari, M., Jungherz, D., Peng, Y., Akbar, Z., Tudor Ilca, F., Otteson, G.E., Seheult, J.N., Ortmann, J., et al. (2024). Dual T-cell constant β chain (TRBC)1 and TRBC2 staining for the identification of T-cell neoplasms by flow cytometry. Blood Cancer Journal *14*, 34. 10.1038/s41408-024-01002-0.

45. Zeng, X., Wang, T., Kang, Y., Bai, G., and Ma, B. (2023). Evaluation of Molecular Simulations and Deep Learning Prediction of Antibodies' Recognition of TRBC1 and TRBC2. Antibodies (Basel) *12*. 10.3390/antib12030058.

46. Band, H., Hochstenbach, F., Parker, C.M., McLean, J., Krangel, M.S., and Brenner, M.B. (1989). Expression of human T cell receptor-gamma delta structural forms. J Immunol *142*, 3627-3633.

47. Lefranc, M.-P., and Rabbitts, T.H. (1989). The human T-cell receptor γ (TRG) genes. Trends in Biochemical Sciences *14*, 214-218. https://doi.org/10.1016/0968-0004(89)90029-7.

48. Buresi, C., Ghanem, N., Huck, S., Lefranc, G., and Lefranc, M.-P. (1989). Exon duplication and triplication in the human T-cell receptor gamma constant region genes and RFLP in French, Lebanese, Tunisian, and Black African populations. Immunogenetics *29*, 161-172. 10.1007/BF00373641.

49. Sturm, E., Braakman, E., Bontrop, R.E., Chuchana, P., Van de Griend, R.J., Koning, F., Lefranc, M.P., and Bolhuis, R.L. (1989). Coordinated V gamma and V delta gene segment rearrangements in human T cell receptor gamma/delta+ lymphocytes. Eur J Immunol *19*, 1261-1265. 10.1002/eji.1830190717.

50. Borst, J., Wicherink, A., Van Dongen, J.J., De Vries, E., Comans-Bitter, W.M., Wassenaar, F., and Van Den Elsen, P. (1989). Non-random expression of T cell receptor gamma and delta variable gene segments in functional T lymphocyte clones from human peripheral blood. Eur J Immunol *19*, 1559-1568. 10.1002/eji.1830190907.

51. Roels, J., Van Hulle, J., Lavaert, M., Kuchmiy, A., Strubbe, S., Putteman, T., Vandekerckhove, B., Leclercq, G., Van Nieuwerburgh, F., Boehme, L., and Taghon, T. (2022). Transcriptional dynamics and epigenetic regulation of E and ID protein encoding genes during human T cell development. Front Immunol *13*, 960918. 10.3389/fimmu.2022.960918.

52. Xin, W., Huang, B., Chi, X., Liu, Y., Xu, M., Zhang, Y., Li, X., Su, Q., and Zhou, Q. (2024). Structures of human γδ T cell receptor-CD3 complex. Nature. 10.1038/s41586-024-07439-4.

53. Broughton, A.E. (2023). CD1d-γδ T Cell Receptor Interaction: Adaptations and Function in Tissue Immunity. Doctor of Philosophy (University of Chicago).

54. Harly, C., Guillaume, Y., Nedellec, S., Peigné, C.M., Mönkkönen, H., Mönkkönen, J., Li, J., Kuball, J., Adams, E.J., Netzer, S., et al. (2012). Key implication of CD277/butyrophilin-3 (BTN3A) in cellular stress sensing by a major human γδ T-cell subset. Blood *120*, 2269-2279. 10.1182/blood-2012-05-430470.

55. Fergusson, J.R., Smith, K.E., Fleming, V.M., Rajoriya, N., Newell, E.W., Simmons, R., Marchi, E., Björkander, S., Kang, Y.H., Swadling, L., et al. (2014). CD161 defines a transcriptional and functional phenotype across distinct human T cell lineages. Cell Rep *9*, 1075-1088. 10.1016/j.celrep.2014.09.045.

56. Pizzolato, G., Kaminski, H., Tosolini, M., Franchini, D.M., Pont, F., Martins, F., Valle, C., Labourdette, D., Cadot, S., Quillet-Mary, A., et al. (2019). Single-cell RNA sequencing unveils the shared and the distinct cytotoxic hallmarks of human TCRVδ1 and TCRVδ2 γδ T lymphocytes. Proc Natl Acad Sci U S A *116*, 11906-11915. 10.1073/pnas.1818488116.

57. Palker, T.J., Fong, A.M., Scearce, R.M., Patel, D.D., and Haynes, B.F. (1998). Developmental regulation of lymphocyte-specific protein 1 (LSP1) expression in thymus during human T-cell maturation. Hybridoma *17*, 497-507. 10.1089/hyb.1998.17.497.

58. Zhang, B., Lin, Y.Y., Dai, M., and Zhuang, Y. (2014). Id3 and Id2 act as a dual safety mechanism in regulating the development and population size of innate-like γδ T cells. J Immunol *192*, 1055-1063. 10.4049/jimmunol.1302694.

59. Tan, L., Fichtner, A.S., Bruni, E., Odak, I., Sandrock, I., Bubke, A., Borchers, A., Schultze-Florey, C., Koenecke, C., Förster, R., et al. (2021). A fetal wave of human type 3 effector γδ cells with restricted TCR diversity persists into adulthood. Sci Immunol *6*. 10.1126/sciimmunol.abf0125.

60. Van Dongen, J.J., Wolvers-Tettero, I.L., Seidman, J.G., Ang, S.L., Van de Griend, R.J., De Vries, E.F., and Borst, J. (1987). Two types of gamma T cell receptors expressed by T cell acute lymphoblastic leukemias. Eur J Immunol *17*, 1719-1728. 10.1002/eji.1830171207.

61. Tribel, F., Lefranc, M.P., and Hercend, T. (1988). Further evidence for a sequentially ordered activation of T cell rearranging gamma genes during T lymphocyte differentiation. Eur J Immunol *18*, 789-794. 10.1002/eji.1830180520.

62. Zhao, X., Shan, Q., and Xue, H.-H. (2022). TCF1 in T cell immunity: a broadened frontier. Nature Reviews Immunology *22*, 147-157. 10.1038/s41577-021-00563-6.

63. Xing, S., Li, F., Zeng, Z., Zhao, Y., Yu, S., Shan, Q., Li, Y., Phillips, F.C., Maina, P.K., Qi, H.H., et al. (2016). Tcf1 and Lef1 transcription factors establish CD8+ T cell identity through intrinsic HDAC activity. Nature Immunology *17*, 695-703. 10.1038/ni.3456.

64. Jaccard, A., Wyss, T., Maldonado-Pérez, N., Rath, J.A., Bevilacqua, A., Peng, J.J., Lepez, A., Von Gunten, C., Franco, F., Kao, K.C., et al. (2023). Reductive carboxylation epigenetically instructs T cell differentiation. Nature *621*, 849-856. 10.1038/s41586-023-06546-y.

65. Apostolov, A.K., Hamani, M., Hernandez-Vargas, H., Igalouzene, R., Guyennon, A., Fesneau, O., Marie, J.C., and Soudja, S.M.h. (2022). Common and Exclusive Features of Intestinal Intraepithelial γδ T Cells and Other γδ T Cell Subsets. ImmunoHorizons *6*, 515-527. 10.4049/immunohorizons.2200046.

66. Luoma, A.M., Suo, S., Williams, H.L., Sharova, T., Sullivan, K., Manos, M., Bowling, P., Hodi, F.S., Rahma, O., Sullivan, R.J., et al. (2020). Molecular Pathways of Colon Inflammation Induced by Cancer Immunotherapy. Cell *182*, 655-671.e622. https://doi.org/10.1016/j.cell.2020.06.001.

67. Di Marco Barros, R., Roberts, N.A., Dart, R.J., Vantourout, P., Jandke, A., Nussbaumer, O., Deban, L., Cipolat, S., Hart, R., Iannitto, M.L., et al. (2016). Epithelia Use Butyrophilin-like Molecules to Shape Organ-Specific γδ T Cell Compartments. Cell *167*, 203-218.e217. 10.1016/j.cell.2016.08.030.

68. Malinarich, F.H., Grabski, E., Worbs, T., Chennupati, V., Haas, J.D., Schmitz, S., Candia, E., Quera, R., Malissen, B., Förster, R., et al. (2010). Constant TCR triggering suggests that the TCR expressed on intestinal intraepithelial γδ T cells is functional in vivo. European Journal of Immunology *40*, 3378-3388. https://doi.org/10.1002/eji.201040727.

69. Herzenberg, L.A., and Herzenberg, L.A. (1989). Toward a layered immune system. Cell *59*, 953-954. 10.1016/0092-8674(89)90748-4.

70. Fichtner, A.S., Ravens, S., and Prinz, I. (2020). Human γδ TCR Repertoires in Health and Disease. Cells *9*. 10.3390/cells9040800.

71. Hao, Y., Stuart, T., Kowalski, M.H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., and Satija, R. (2024). Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nature Biotechnology *42*, 293-304. 10.1038/s41587-023-01767-y.

72. Borcherding, N., Bormann, N.L., and Kraus, G. (2020). scRepertoire: An R-based toolkit for single-cell immune receptor analysis. F1000Res *9*, 47. 10.12688/f1000research.22139.2.

73. Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nature Biotechnology *32*, 381-386. 10.1038/nbt.2859.