Mucosal memory T cells in breastmilk are modulated by SARS-CoV-2 mRNA vaccination

Blair Armistead[1]†, Yonghou Jiang[1]†, Marc Carlson[2], Emily S Ford[3,4], Saumya Jani[4,5], John Houck[1], Xia Wu[4], Tiffany Pecor[1], Alisa Kachikis[6], Winnie Yeung[1], Tina Nguyen[1], Nana Minkah[1,8], Sasha E Larsen[1], Rhea N Coler[1,7,8], David M Koelle[3,4,5,7,9], Whitney E Harrington[1,7,8]*

Affiliations:

[1]Center for Global Infectious Disease Research, Seattle Children's Research Institute; Seattle, WA, USA

[2]Research Scientific Computing, Enterprise Analytics, Seattle Children's Research Institute; Seattle, WA, USA

[3]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center; Seattle, WA, USA

[4]Department of Medicine, University of Washington; Seattle, WA, USA

[5]Department of Laboratory Medicine & Pathology, University of Washington; Seattle, WA, USA

[6]Department of Obstetrics & Gynecology, University of Washington; Seattle, WA, USA

[7]Department of Global Health, University of Washington; Seattle, WA, USA

[8]Department of Pediatrics, University of Washington; Seattle, WA, USA

[9]Benaroya Research Institute; Seattle, WA, USA

*Corresponding author. Email: whitney.harrington@seattlechildrens.org

†Co-first authors

Abstract: We compared the phenotype, diversity, and antigen specificity of T cells in the breastmilk and peripheral blood of lactating individuals who received SARS-CoV-2 mRNA vaccination. Relative to blood, breastmilk contained higher frequencies of T effector and central memory populations that expressed mucosal-homing markers. T cell receptor (**TCR**) sequence overlap was limited between blood and breastmilk. Overabundant breastmilk clones were observed in all individuals, were structurally diverse, and contained CDR3 sequences with known epitope specificity including to SARS-CoV-2 Spike. Spike-specific TCRs were more frequent in breastmilk compared to blood and expanded in breastmilk following a third mRNA vaccine dose. Our observations indicate that the lactating breast contains a distinct T cell population that can be modulated by maternal vaccination with potential implications for infant passive protection.

Summary: The breastmilk T cell repertoire is distinct and enriched for SARS-CoV-2 Spike-specificity after maternal mRNA vaccination.

Introduction:

The breastfed human infant consumes up to 750,000 maternal leukocytes per day, 5-10% of which are T cells whose function is poorly understood (*1, 2*). Breastmilk lymphocytes are most abundant at delivery and decline over the first month post-partum to a steady state that persists for up to two years (*1-4*). Breastmilk T cells are phenotypically distinct from peripheral blood T cells, with higher expression of mucosal and effector memory markers (*5, 6*). Cytometaglovirus, Epstein-Barr virus (**EBV**), influenza, and HIV-specific T cell responses have been detected in breastmilk cells (**BMC**) at higher frequencies than in peripheral blood mononuclear cells (**PBMC**) (*6-10*), and breastmilk T cells expand in the setting of maternal or infant infection (*2, 11, 12*).

The infant stomach pH (*13, 14*) and intestinal permeability (*13, 15, 16*) are also highest in the first few weeks of life, and evidence from animal models demonstrates that breastmilk T cells can survive the offspring gastrointestinal tract and traffic into the mesenteric lymph nodes, liver, spleen, and lung as a form of maternal microchimerism (*17-19*). In mice, breastmilk-derived helminth-specific T cells were protective in the offspring upon challenge with the same helminth (*17*), and in lambs, breastmilk-derived tetanus-specific T cells enhanced the response to tetanus vaccination in the offspring (*20*). Human breastmilk maternal microchimerism has not been conclusively demonstrated, although we recently found in a cohort of infants that maternal microchimerism increased up to three months of age and was positively correlated with breastfeeding (*21*). These data emphasize the potential for breastmilk-derived maternal T cells to become resident in the infant and provide passive protection.

The full repertoire of breastmilk T cells has not been described, however, and it is unknown whether maternal vaccine-specific T cells are present in human breastmilk. In the setting of the ongoing pandemic, pregnant people are now widely receiving SARS-CoV-2 (**SARS2**) vaccines including the Spike protein mRNA-based vaccines mRNA1273 (Moderna) (*22*) and BNT162b2 (Pfizer-BioNTech) (*23*). SARS2 mRNA vaccines generate a robust T cell response in peripheral blood (*22-24*), yet their impact on mucosal T cell responses in breastmilk is unknown. To understand the breadth of maternal T cells consumed by the infant, we characterized the phenotype and diversity of paired breastmilk and peripheral T cells. We further investigated the hypothesis that Spike-specific T cells are present in the breastmilk of vaccinated individuals.

Results:

Breastmilk is enriched for antigen-experienced T cells

We collected paired blood and breastmilk from lactating people who had received two doses of BNT162b2 or mRNA1273 and had no history of SARS2 infection. Breastmilk contained a low but detectible frequency of T cells, and the distribution of CD4+ and CD8+ T cells was similar in BMC and PBMC (**Fig. S1**). However, in contrast to PBMC, nearly all BMC CD4+ T cells were CD45RO+ antigen-experienced (97% vs. 51% in PBMC, $p<0.001$), and there was a significant enrichment of CD45RO+/CCR7- effector memory (**T_{EM}**) and CD45RO+/CCR7+ central memory (**T_{CM}**) populations (**Fig. 1A, B**). The frequency of CD4+ T_{EM} and T_{CM} populations did not vary in BMC or PBMC by time since delivery or 2nd vaccine dose (**Fig. 1C, D, Fig. S2**). Within the CD8+ population, there was also a higher frequency of CD45RO+ T cells in BMC versus PBMC (80% vs. 42%, $p<0.001$), with an enrichment of the T_{EM} but not the T_{CM} population (**Fig. 1A, B**). There was no difference in the frequency of the CD8+ T_{EM} population by time since delivery or 2nd vaccine dose (**Fig. 1C, D**). However, within the T_{CM} population, the frequency of T_{CM} increased in PBMC only as a function of time since 2nd dose (**Fig. S2**). These data emphasize that breastmilk is highly enriched for memory T cell populations relative to peripheral blood.

Mucosal-homing markers are highly expressed on breastmilk T cells

We next investigated the expression of mucosal-homing markers on T cells within BMC and PBMC. The CD4+ population within BMC versus PBMC had a higher frequency of CCR9+ and CD103+ cells, and a higher frequency of double positive CCR9+/CD103+ cells (**Fig. 2A, B**). The frequency of CCR9+ cells increased as a function of time since 2nd dose in the BMC compartment only (**Fig. 2C, D**). In contrast, there was an increase in the frequency of the CCR9+/CD103+ population in BMC but a decrease in the PBMC by time since delivery (**Fig. S3**). Similarly, the CD8+ population within BMC versus PBMC had a higher frequency of CCR9+ and CD103+ cells, as well as a higher frequency of CCR9+/CD103+ cells (**Fig. 2A, B**). The frequency of CCR9+ cells in BMC increased as a function of both time since delivery and time since 2nd dose (**Fig. 2C, D**), whereas the frequency of CCR9+/CD103+ cells in BMC increased as a function of time since 2nd dose only (**Fig. S3**). These data emphasize that the T cells in breastmilk express high levels of mucosal-homing markers which evolve as a function of time.

Restricted T cell receptor repertoire overlap between peripheral blood and breastmilk

We next investigated the diversity of the T cell receptor (**TCR**) repertoire in BMC, as well as the degree of overlap with the repertoire of peripheral blood T cells, using bulk TCR beta chain

(**TCRβ**) sequencing. We first compared TCR repertoire overlap between the BMC and PBMC within each participant using the Morisita Index (*25*) and surprisingly observed relatively low overlap in all pairs (**Fig. 3; Table S1**). As a comparison, we considered the TCRβ repertoire in PBMC from one individual at two time points (9 and 17 days post-2nd vaccine dose); in this case there was a high degree of overlap. The Morisita Index between paired BMC and PBMC was not related to the number of productive templates in the BMC (R^2=0.25, p=0.5), suggesting that the low overlap was independent of sampling depth of the BMC and unrelated to time since delivery or 2nd dose. Within each individual, we compared the frequency of clonotypes across the two compartments using the ImmunoSEQ® Differential Abundance tool (*26*). In all individuals, there were select clonotypes that were statistically significantly overabundant in BMC relative PBMC (**Fig. 3; Table S1**). In contrast, the Simpson Clonality and maximum clone frequency—two metrics of absolute clonality —did not differ in BMC and PBMC. These data indicate that the TCR repertoire of BMC is highly diverse with select overabundant clonotypes but with limited overlap with PBMC.

Characterization of breastmilk overabundant T cell clonotypes

We next explored the structural diversity of the overabundant T cell clonotypes in breastmilk. We analyzed each participant's BMC T cell repertoire using a TCR distance metric, TCRdist3 (*27, 28*), which clusters TCRβs based on structural and functional similarities of amino acids within complementarity-determining regions (**CDR**). In all participants, overabundant clones were broadly distributed across the BMC TCRβ repertoire, suggesting structural (and likely epitope) diversity (**Fig. 4A, Fig. S4**). We additionally evaluated how overabundant clones clustered across all individuals. Most clones were unique to a single individual (i.e. they were "private"), and the clonotypes did not segregate by individual (**Fig. 4B**), reflecting the structural diversity of each individual's overabundant clonotypes.

To determine potential antigen specificity, we compared CDR3 amino acid sequences of the overabundant breastmilk clones from all individuals to TCRβ sequence databases populated by validated epitope-specific TCRs. We identified five direct matches from two participants with CDR3 amino acid sequences and identical V gene usage, two of which were reported to bind SARS2 Spike epitopes (**Fig. 4B, Table S2**). Notably, the two Spike-specific clones had also previously been reported to bind influenza M1 (*29, 30*). Other direct matches included epitopes derived from influenza, *Mycobacterium tuberculosis* lysate, and EBV (**Fig. 4B, Table S2**). Additional clones with identical CDR3 sequences but non-identical V gene usage were reported to bind epitopes from SARS2, EBV, and *M. tuberculosis* (**Fig. 4B, Table S2**). CDR3 sequences

with previously published specificity clustered tightly together, irrespective of individual (**Fig 4**). Finally, we used the IEDB TCRMatch Tool (*31*) to predict TCR-epitope specificity based on sequence similarity to published TCR sequences, which identified TCR clones with potential specificity to a variety of viral epitopes (**Table S3**). These data indicate that overabundant breastmilk T cell clones are structurally diverse and respond to a range of pathogen-specific epitopes.

SARS2 Spike-restricted TCRβs are present in breastmilk T cells

We next investigated the presence of Spike-specific clones in BMC T cells more broadly. We utilized the ImmunoSEQ® COVID Search Tool (*32*) to identify candidate Spike-specific TCRβ in BMC and PBMC samples, irrespective of sampling depth for the BMC samples or availability of paired PBMC for comparison. All PBMC contained candidate Spike-specific TCRβ, though their predicted epitope specificity was distributed across the entire Spike protein with low frequency, suggesting that some of the TCRβ may represent clones in the naïve repertoire rather than expanded vaccine-specific populations (**Fig. 5A**). Thirteen of the 16 BMC samples contained TCRβ predicted to be Spike-specific. In the pairs where both BMC and PBMC were sampled and at least one clone in each compartment was predicted to be SARS2-specific, Spike-specific TCRβs were nearly 2-fold enriched in BMC versus PBMC (**Fig. 5B**). Spike-specific TCRβ decayed slightly faster as a function of time since delivery in BMC than in PBMC, whereas neither decreased as a function of time since 2nd dose (**Fig. S5**).

To further validate the presence of Spike-specific T cells in breastmilk, we cross-referenced each participant's BMC and PBMC TCRβ CDR3 amino acid sequences and V gene usage against publicly available TCRβ datasets from Spike-epitope-loaded tetramer or multimer experiments (*33-37*). We identified high quality CDR3 sequence matches in half of all breastmilk samples, with four samples also containing hits with identical V gene usage (**Table S4**). All PBMC contained clones with identical CDR3 sequence and V gene usage to those published previously (**Table S5**). Consistent with prior studies of PBMC (*33-35, 38*), sequences specific to the Spike peptide YLQPRTFLL were prominent in the breastmilk and blood of individuals known or presumed to be HLA A*02:01 positive. In addition, clones specific to the A*01:01-restricted Spike peptide LTDEMIAQY and the B*15:01-restricted Spike peptide NQKLIANQF were present in the breastmilk and blood of HLA concordant participants (**Table S4, Table S5**), suggesting shared epitope specificity following vaccination. Finally, we utilized the list of Spike-specific TCRβs in conjunction with the TCRdist3 algorithm (*27, 28*) to identify novel potential Spike-specific T cell

clones in the BMC (**Table S5**). These observations demonstrate the presence of Spike-specific T cells in breastmilk following mRNA vaccination.

Spike-specific T cells in breastmilk expand after SARS2 mRNA vaccine

To understand whether the Spike-specific clones in breastmilk were responsive, we took advantage of a natural re-stimulation experiment in which select individuals donated additional breastmilk pre- and ~1-week post-3rd vaccine dose for in-depth phenotyping and tetramer staining. Of five individuals, three were HLA-A*02:01 positive, and we identified HLA-A*02:01-YLQPRTFL tetramer-positive cells in all BMC post-3rd dose. Further, the frequency of tetramer-positive CD8+ cells increased from a mean of 0.3% to 1.6% between the pre- and post-3rd dose samples (**Fig. 6A, B**) in BMC, whereas paired PBMC showed minimal response (**Fig. 6B**). When restricted to CD45RO+/CD8+ cells to account for the naïve cells in PBMC, the fold change appeared greater in BMC versus PBMC (**Fig. 6C**). In the post-3rd dose samples, a higher proportion of tetramer positive, CD45RO+/CD8+ cells in the BMC versus PBMC expressed the activation marker CCR5, suggesting recent activation and/or tissue recruitment (**Fig. 6D**). Interestingly, tetramer positive cells in both the BMC and PBMC had low expression of CCR9 and/or CD103. These data demonstrate that SARS2 Spike-specific cells in breastmilk respond *in vivo* upon antigen restimulation.

Discussion:

We present a comprehensive comparison of the T cells present in breastmilk relative to peripheral blood, considering phenotype, diversity, and antigen specificity. We find that T cells in breastmilk are nearly uniformly memory populations and have high expression of mucosal-homing markers. Their TCRβ repertoire is diverse, yet distinct from paired PBMC, and with select overabundant clones. The overabundant clones are structurally diverse but cluster across individuals suggesting a potential for a shared breastmilk mucosal repertoire. Further, we identify SARS2 Spike-specific clones in the breastmilk of vaccinated individuals, emphasizing that vaccine-specific T cells are present at mucosal sites such as the breast, with important implications for both maternal and infant health.

In bulk analysis, breastmilk T cells were enriched for effector memory populations that displayed high levels of mucosal-homing markers, consistent with earlier reports in individuals living with or without HIV (*6*). These data suggest that breastmilk T cells may be derived from a tissue-resident population in the breast (*39*), rather than due to vessel microtrauma and contamination by

peripheral blood. In addition, the expression of CCR9 increased as a function of both time since delivery and time since 2nd vaccine dose emphasizing that the breast as a mucosal site develops over time (*1*). The increase in mucosal T cell populations may reflect ongoing trafficking into the breast or expansion of breast-resident populations in response to antigenic stimulation (e.g. from infant saliva). The high expression of both CCR9 and CD103 by breastmilk T cells also supports the notion of an entero-mammary axis (*40*). Though not examined in the present study, these observations raise the possibility that breastmilk T cells may traffic to the infant respiratory and gastrointestinal tracts when consumed, as has been shown in animals (*17-19*).

The TCRβ repertoire in BMC was diverse and had uniquely expanded clonotypes relative to paired PBMC. To date, studies have focused on T cell responses to specific pathogens (*6-11*), rather than capturing the full diversity of the compartment. The low degree of TCRβ repertoire overlap between the BMC and PBMC may reflect a difference in the distribution of naïve versus antigen-experienced T cells. However, the absolute clonality was similar and there was evidence of high frequency clonotypes, likely memory cells, that were differentially abundant, suggesting distinct responses in the two compartments independent of the naïve population. The observation of overabundant clones in the BMC is consistent with prior reports of an enrichment of virus-specific responses in breastmilk relative to PBMC (*6*), although in our paired TCRβ analysis only one individual had CMV-specific TCRβ and none reported HIV infection, emphasizing that this breast-specific enrichment is not restricted virus-specific clones. In addition, within each individual, overabundant clones were highly structurally diverse, suggesting breadth of response.

To identify antigen specificity of BMC T cells, we used a combination of prior published TCRβ specificities and predictive algorithms, which most notably allowed for the high-confidence detection of SARS2 Spike-specific clones in most individuals. Consistent with previous observations of convergent epitope specificity across HLA-concordant people following SARS2 mRNA vaccination (*33*), we found identical Spike-specific CDR3 sequences in BMC of several individuals. For instance, BMC sequences restricted to the Spike epitope YLQPRTFLL were shared across several HLA-A*02:01 positive individuals. This observation, along with the approval of 3rd mRNA vaccine doses, provided a unique opportunity to observe the dynamics of Spike-specific T cells after antigen re-encounter using the well validated HLA*A:02:01-YLQ tetramer (*33-35, 37*). We observed expansion of YLQ-specific T cells in BMC with a magnitude greater than in PBMC. Tetramer positive cells in the BMC expressed CCR5, but were CD103 negative, similar to a recent study of nasal CD8+ T cells following SARS2 mRNA vaccination (*41*). The lack of mucosal markers on these cells may reflect timing of sample collection or recent trafficking of

cells from blood. Enrichment of Spike-specific T cells in breastmilk may suggest homing of T cells to the breast or other mucosal sites following vaccination.

In addition to the potential benefit provided to the infant, the recognition of the breast as a site of mucosal immunity distinct from peripheral immunity has important implications for the study of vaccine responses. For example, prior work on response to SARS2 mRNA vaccines has primarily focused on both CD4+ and CD8+ T cell responses present in peripheral blood (*22, 23*), whereas the site of primary infectious challenges is the respiratory tract. Tissue resident cells in the lung are difficult to access in human populations, whereas the collection of breastmilk in lactating individuals is non-invasive. While the association between respiratory tract and breast resident T cell responses will need further investigation, it is possible that measuring immune responses in breastmilk will allow for the characterization of mucosal immunity more broadly following vaccination.

Our study had several limitations. BMC T cells were low frequency, and although we designed our experimental approach to maximize the information obtained from each sample, the sampling depth of BMC versus PBMC was different. However, we took advantage of several computational solutions to overcome this challenge. We were further limited to bulk TCRβ sequencing of combined CD4+ and CD8+ populations, as the number of each sub-population was too small to meet technical requirements to analyze separately. However, the CD4+ and CD8+ frequencies in BMC were similar to PBMC, suggesting that differences in the TCRβ repertoire we observed cannot be explained by a difference in these populations. The number of T cells recovered from each breastmilk limited our ability to conduct functional analyses. However, we took advantage of individuals who received a 3rd dose of vaccine to demonstrate expansion of Spike-specific T cells with tetramer staining.

We demonstrate that breastmilk T cells are highly diverse and enriched for mucosal memory populations, emphasizing that the breast represents a underrecognized site of mucosal immunity. Further, we identify SARS2 Spike-specific T cells in mRNA vaccinated individuals, the first demonstration of vaccine-specific T cells in breastmilk. This observation may have important

implications for both the study of vaccine-induced T cell responses in the vaccinated individual, as well as for infant passive protection.

Materials and Methods:

Cohort

Lactating individuals (n = 24) were recruited as part of the Center for Global Infectious Disease Research Biorespository, approved by the Seattle Children's Institutional Review Board (**IRB**) (STUDY00002048). All participants self-reported as healthy, not pregnant, and weighing > 110 pounds. Demographic data was collected including age, history of SARS or other serious infections, SARS2 vaccination (brand of vaccine received, dates of doses), and age of children. All participants reported receipt of two doses of a SARS2 Spike mRNA vaccine. Breastmilk was pumped by participants within four hours of blood draw, except for three participants for whom blood could not be obtained (n = 3). Of the 5 participants who contributed pre- and post-3rd vaccination samples, 4 had previously contributed to the study and 1 was newly enrolled. The volume of breastmilk donated ranged from 7 mL to 100 mL (median = 60 mL), and the estimated frequency of live T cells in breastmilk ranged from 4.3 cells/mL to 2,198 cells/mL (median = 106 cells/mL).

Isolation of PBMCs from whole blood

Whole blood was collected in EDTA Vacutainer tubes (BD). Within 4 hours of blood collection, tubes were centrifuged at 400 x g for 10 minutes. The plasma fraction was removed, centrifuged at 800 x g for 15 minutes, aliquoted into cryovials, and stored at -80 °C. The remaining blood was diluted in sterile phosphate buffered saline (**PBS**), layered onto Lymphocyte Separation Medium (Corning), and centrifuged at 800 x *g* for 20 minutes at room temperature with no brake. The resulting buffy coat layer was removed and washed two to three times in PBS. Cells were counted using a C-Chip hemocytometer (INCYTO), resuspended in freezing medium (50% fetal bovine serum [FBS], 40% RPMI with L-glutamine, 10% DMSO (Millipore Sigma)) at 5 to 10 million cells/mL, and aliquoted into cryovials. Cryovials were immediately placed into a 1 °C cryogenic

freezing container (Nalgene), which was stored at -80 °C overnight. Cryovials were then moved to liquid nitrogen for long-term storage.

Isolation of BMC

Milk was centrifuged at 400 x *g* for 15 minutes at 4 °C, and the aqueous fraction was aliquoted into cryovials and stored at -80 °C. The cell pellet was washed three times in 40 mL sterile PBS with 2% FBS. As above, cells were counted using a C-Chip hemocytometer (INCYTO), resuspended in freezing medium at 1 to 3 million cells/mL, aliquoted into cryovials, placed in a 1 °C cryogenic freezing container at -80 °C overnight, and then transferred to liquid nitrogen. Breastmilk samples with less than ~10^6 cells were not utilized for further analysis.

Thawing PBMC and BMC

Prior to use in assays, PBMC or BMC were thawed in a 37 °C water bath until a small ice crystal remained. Resuspensions were transferred into a tube containing pre-warmed thaw medium (RPMI with L-glutamine, 20% FBS, 2 mM EDTA) and centrifuged at 400 x *g* for 5 minutes. Cell pellets were resuspended in approximately 5 mL complete medium (RPMI with L-glutamine, 10% FBS, 100 U/mL penicillin, 100 µg/mL streptomycin) and then counted and assessed for viability before final resuspension in complete medium.

Phenotyping and cell sorting by flow cytometry

PBMC and BMC were thawed as described and washed with FACS buffer (PBS, 2% BSA). Cells were stained with 100 µL of master mix comprised of the optimal dilutions of fluorophore-conjugated antibodies and viability dye (**Table S6; Panel 1**) in Brilliant Stain Buffer (BD) for 30 minutes at room temperature, protected from light. Stained cells were washed with 4 mL FACS buffer and resuspended in 200 to 300 µL FACS buffer. Cells were run on a FACSMelody Cell Sorter (BD) or a FACSAria II Cell Sorter (BD), and CD45+ cells were collected. Single stained CompBeads (ThermoFisher) were used as compensation controls, and fluorescence-minus-one stained cells were used to set fluorescence gates. Data were analyzed on FlowJo version 10 (BD), and included analysis of memory markers and mucosal markers. CD45RO+/CCR7- T cells were designated as effector memory (**T_{EM}**), CD45RO+/CCR7+ T cells were designated as central memory (**T_{CM}**), and CD45RO-/CCR7+ T cells were designated as naïve-like (**$T_{N\text{-}like}$**). Gating strategy is shown in Fig. S7. In addition, data were used to anticipate the frequency of T cells

(CD3+) in the collected CD45+ population from BMC to optimize subsequent genomic DNA extraction.

YLQ tetramer generation

YLQPRTFLL (i.e. YLQ peptide) was synthesized by GenScript Biotech (Piscataway, NJ). YLQ peptide (400 µM) was mixed 1:1 (v/v) with 200 µg/mL Flex-T HLA-A*02:01 ultraviolet exchange UVX monomer (BioLegend) and treated with UV irradiation (368 nm) for 30 minutes using a UV crosslinker (Fisher Scientific) to remove the UV-liable peptide. The mixture was incubated at 37 °C for 30 minutes to form YLQ monomers, which were then tetramerized through the addition of 200 µg/mL streptavidin-APC (BD) and incubation on ice for 30 minutes. Excess streptavidin was blocked with PBS + 0.4 µM D-Biotin + 0.3% (w/v) NaN_3 overnight at 4 °C.

YLQ tetramer staining and flow cytometry analysis

Cells were thawed and washed as described above and then stained with YLQ tetramer (1:50) for 30 minutes at 4 ˚C, protected from light. Cells were washed with FACS buffer and then stained with antibody master mix containing the optimal dilutions of all antibodies and viability dye in Brilliant Staining Buffer (**Table S7; Panel 2**) for 30 minutes at room temperature, protected from light. As above, stained cells were washed in FACS buffer and run on a FACSAria II Cell Sorter (BD) with single stained CompBeads (ThermoFisher) as compensation controls. Data were analyzed on FlowJo version 10 (BD). YLQ tetramer performance was validated with YLQ-specific CD8+ T cells expanded from an HLA-A*02:01 SARS2 convalescent donor spiked into an HLA-A*02:01 negative donor with PBMC collected prior to 2019 (Fig. S8) (University of Washington IRB: STUDY00004312). Tetramer positive cells from BMC and PBMC were gated as follows: lymphocyte cloud (FSC-A by SSC-A) → single cells (FSC-A by FSC-H) → live cells (live/dead Aqua by FSC-H) → CD45+/CD14- → CD3+ → CD8+ → YLQ+.

HLA typing

Genomic DNA was extracted from participants' PBMC (where available) with the QIAamp® DNA Blood Mini Kit per manufacturer instructions, and HLA class I genotyping was performed via direct sequencing (Sisco Genetics, Seattle, WA).

TCRβ sequencing

Flow-sorted BMCs (CD45+) expected to contain T cells underwent genomic DNA extraction using a protocol modified from that recommended by Qiagen to recover low-yield DNA. Briefly, sorted BMCs were pelleted in their collection tubes at 400 x g for 10 minutes. Cell pellets were extracted

directly in the collection tubes using 30 µL of Qiagen Protease (Qiagen) and incubated at 70 °C for 20 minutes in a heat block with periodic vortexing. Approximately 1.5 µg of carrier RNA (Qiagen) and 300 µL of Buffer AL (Qiagen) was added. Samples were incubated for another 20 minutes at 70°C in a heat block with periodic vortexing and centrifugation. To recover DNA, 300 µL of 100% ethanol was added to the sample followed by vortexing. The sample was transferred to a QIAamp® Mini spin column and underwent the manufacturer's standard protocol from the QIAamp® DNA Blood Mini Kit for column washing. DNA was eluted in pre-warmed (56 °C) sterile water twice in 50 µL volumes. To assess DNA yield, a qPCR assay targeting the β-globin gene against a standard curve was performed, and only samples anticipated to contain at least 1,000 T cells were sent for TCRβ sequencing. Maternal PBMCs underwent genomic DNA extraction using the standard protocol provided in the QIAamp® DNA Blood Mini Kit (Qiagen), and 3.4 ug of total genomic DNA was sent for TCRβ sequencing. Samples were sent in batches to Adaptive Biotechnologies (Seattle, WA, USA) and assayed using their ImmunoSEQ® hsTCRBB (*42*) service pipeline. Sequencing was performed at a survey level following Adaptive Biotechnologies' custom protocol. Quality control of sequencing data was performed by Adaptive Biotechnologies. Total productive templates from PBMC ranged from 64,403 to 118,002 (median 94,389), whereas total productive templates from BMC ranged from 936 to 17,085 (median 2,976) (**Table S4**).

TCR sequence analysis

TCRβ sequence data were analyzed using the ImmunoSEQ® Analyzer software (Adaptive Biotechnologies) and/or exported to R for analysis with the package immunoArch (*43*). Repertoire overlap between blood and breastmilk was assessed using the Morisita Index, which is relatively protected from differences in sampling depth (*25*). To identify overabundant clones in the breastmilk, frequencies of TCR clonotype nucleotide sequences in the breastmilk were compared to those in peripheral blood using the Differential Abundance tool in ImmunoSEQ® Analyzer using the binomial statistical method with Benjamini Hochberg correction and a lower limit of detection of 10.

The CDR3β amino acid sequence of every overabundant clone was compared to that of all other overabundant clones within each participant using TCRdist3 (*27, 28*). Similarly, CDR3β amino acid sequences of overabundant clones across all participants were compared to one another using TCRdist3 (*27, 28*). To identify epitope specificity of overabundant breastmilk clones, TCRβ sequences were matched by CDR3 amino acid and V gene identification against several public databases of TCRβ epitope specificity, namely ImmuneCODE (*32*) (available online at https://clients.adaptivebiotech.com/pub/covid-2020), VDJdb (available at https://vdjdb.cdr3.net)

(*44*), TCRBdb (available online athttp://bioinfo.life.hust.edu.cn/TCRdb/#/) (*45*), McPAS-TCRB (available online at http://friedmanlab.weizmann.ac.il/McPAS-TCRB/) (*46*), and Immune Epitope Database and analysis Resource (IEDB) (available online at http://www.iedb.org/home_v3.php) (*47*). Epitope matches from any of these databases were considered a direct match if the CDR3 amino acid and TRBV gene were identical and were considered a predicted match if the CDR3 amino acid sequences were identical, but V gene usage was mismatched. All overabundant breastmilk clones were also queried using the IEDB TCRMatch Tool (available online at http://tools.iedb.org/tcrmatch/) (*31*) with a score threshold of 0.97 to identify closely related epitope restrictions.

TCRβ sequences from both breastmilk and PBMC were evaluated for candidate Spike-specific restriction using the COVID Search Tool in ImmunoSEQ® Analyzer, which utilizes TCRβ sequences assigned as specific for SARS-CoV-2 from the ImmuneCODE database (*32, 38*). We compared the frequency of Spike-specific clones out of the number of non-Spike SARS2-specific clones in each sample, where non-Spike SARS2-specific clones were used to represent the "background" response noted in PBMC to identify a vaccine-specific response. Frequencies of assigned Spike-specific TCRβ sequences were compared with negative binomial models accounting for the total productive templates in each sample. For comparison of clonality metrics only (Simpson Clonality and maximum productive frequency), the full dataset of paired breastmilk and PBMC TCRβ sequences was down-sampled to the lowest productive template frequency, and metrics were calculated and compared using the down-sampled dataset.

Candidate spike-specific TCRβ in breastmilk and PBMC were further validated by comparing CDR3β amino acid sequences, V gene usage, epitope HLA restriction, and participant HLA concordance to an internally vetted dataset of published TCRβ sequences obtained via Spike-epitope-loaded tetramer or multimer-based experiments (*33-37*). An exact match at the TCRβ CDR3 sequence was required for these analyses. The internally vetted dataset (*33-37*) was used to train the TCRdist3 algorithm (*27, 28*) with a distance unit threshold of 10 to identify additional potential Spike-specific TCRβ in breastmilk.

ELISA

Wells of a sterile high-binding 384-well plate (Corning) were coated with 11.8 µM full-length trimeric SARS-CoV-2 spike protein (Institute for Protein Design, University of Washington) or 1 µg/mL RBD (Institute for Protein Design, University of Washington) for at least 2 hours at room temperature or up to 3 days at 4 °C. Each well was washed three times with 100 µL 1X Wash Buffer A (Teknova) using a BioTek EL406 plate washer (BioTek). Wells were blocked with with

100 µL blocking buffer (PBS, 1% BSA w/v, 0.05% Tween-20) at least two hours at room temperature or at 4 °C overnight. Wells were then washed three times with 1X Wash Buffer A, and 50 µL assay diluent (50% 1X Wash Buffer A, 50% PBS, 0.1% BSA) was added to each well. Aqueous breast milk fractions were diluted 1:2 and plasma samples were diluted 1:20 in assay diluent in a separate 96-well plate and then added to the first column of the 384-well plate, further diluting the sample by a factor of 5. Each sample was then serially diluted 1:5 in wells from left to right across the plate. After incubating overnight at 4 °C, plates were brought to room temperature, and wells were rinsed five times with 1X Wash Buffer A. HRP-conjugated immunoglobulins were diluted in assay diluent as shown in Table S8, and then 50 µL was added to the appropriate wells. Wells incubated with secondary antibodies for 1 hour protected from light at room temperature and were washed five times with 1X Wash Buffer A followed by one wash with PBS. Wells were then treated with 50 µL Tetramethylbenzidine (TMB, SeraCare) for 2 to 10 minutes, depending on the secondary antibody (Table S8), when the reaction was stopped with the addition of 25 µL 1N H_2SO_4 (Millipore Sigma). Plates were read at 450 nm with a reference filter set at 570 nm using a SpectraMax 13x plate reader (Molecular Devices) and SoftMax Pro 6.4.2 analysis software. The average optical density (OD) value for all dilutions of the negative sample was used to set a minimum cutoff value for each plate. Plate cutoff values were then used to calculate each sample's end point titer (EPT) using a 4-parameter logistic model in XL-fit software (model 208, IDBS). The EPT value for each duplicate plate was averaged for a final EPT run value for each sample.

Statistical approach

To assess the primary difference between cellular phenotypic frequency (e.g. % CCR9+ of CD4+ T cells) in BMC versus PBMC, we built a linear regression model for each outcome with sample type as the predictor and clustering by individual to account for the correlation of individuals contributing paired BMC and PBMC. To subsequently consider whether the difference between BMC and PBMC was modified by time since delivery or time since 2nd dose, we built a second model that also included primary covariates for time since delivery and time since 2nd dose as well as the interaction terms between sample type and these variables (e.g. type*time since delivery). Effect modification was considered significant when the relevant interaction term had a $p<0.1$. In addition, this model allowed us to calculate the effect of time since delivery and time since 2nd dose in each sample type. Due to the large difference in sampling depth in the two compartments, negative binomial models were used to compare the frequency of Spike-specific templates in BMC versus PBMC, accounting for both the number of total SARS2-specific templates identified

(viral genome wide) and the number of Spike-specific templates identified to determine the enrichment of Spike above background reactivity to non-SARS2 coronaviruses and/or TCRs present in the naïve repertoire. The negative binomial model generates an incidence rate ratio (IRR) which represents the number of Spike-specific templates found in the experimental group (e.g. BMC) for every one template identified in the control group (e.g. PBMC). The frequency of tetramer positive T cells in the BMC of individuals pre- and post-3rd dose was compared with paired t tests.

References:

1. F. Hassiotou, D. T. Geddes, P. E. Hartmann, Cells in human milk: state of the science. *J Hum Lact* **29**, 171-182 (2013).
2. F. Hassiotou *et al.*, Maternal and infant infections stimulate a rapid leukocyte response in breastmilk. *Clin Transl Immunology* **2**, e3 (2013).
3. A. S. Goldman, C. Garza, B. L. Nichols, R. M. Goldblum, Immunologic factors in human milk during the first year of lactation. *J Pediatr* **100**, 563-567 (1982).
4. S. Trend *et al.*, Leukocyte Populations in Human Preterm and Term Breast Milk Identified by Multicolour Flow Cytometry. *PLoS ONE* **10**, e0135580 (2015).
5. A. P. Kourtis *et al.*, Breast milk CD4+ T cells express high levels of C chemokine receptor 5 and CXC chemokine receptor 4 and are preserved in HIV-infected mothers receiving highly active antiretroviral therapy. *The Journal of infectious diseases* **195**, 965-972 (2007).
6. S. Sabbaj *et al.*, Breast milk-derived antigen-specific CD8+ T cells: an extralymphoid effector memory cell population in humans. *J Immunol* **174**, 2951-2956 (2005).
7. S. Sabbaj *et al.*, Human immunodeficiency virus-specific CD8(+) T cells in human breast milk. *J Virol* **76**, 7365-7373 (2002).
8. D. C. Moylan *et al.*, Breast Milk Human Cytomegalovirus (CMV) Viral Load and the Establishment of Breast Milk CMV-pp65-Specific CD8 T Cells in Human CMV Infected Mothers. *The Journal of infectious diseases* **216**, 1176-1179 (2017).
9. B. L. Lohman *et al.*, Prevalence and magnitude of human immunodeficiency virus (HIV) type 1-specific lymphocyte responses in breast milk from HIV-1-seropositive women. *The Journal of infectious diseases* **188**, 1666-1674 (2003).
10. F. L. Ruben, I. R. Holzman, P. Fireman, Responses of lymphocytes from human colostrum or milk to influenza antigens. *Am J Obstet Gynecol* **143**, 518-522 (1982).
11. D. L. Bryan, P. H. Hart, K. D. Forsyth, R. A. Gibson, Immunomodulatory constituents of human milk change in response to infant bronchiolitis. *Pediatr Allergy Immunol* **18**, 495-502 (2007).
12. A. Riskin *et al.*, Changes in immunomodulatory constituents of human milk in response to active infection in the nursing infant. *Pediatr Res* **71**, 220-225 (2012).
13. J. P. Moles *et al.*, Breastmilk cell trafficking induces microchimerism-mediated immune system maturation in the infant. *Pediatr Allergy Immunol* **29**, 133-143 (2018).
14. R. A. Miller, Observations on the gastric acidity during the first month of life. *Arch Dis Child* **16**, 22-30 (1941).

15. C. Catassi, A. Bonucci, G. V. Coppa, A. Carlucci, P. L. Giorgi, Intestinal permeability changes during the first month: effect of natural versus artificial feeding. *J Pediatr Gastroenterol Nutr* **21**, 383-386 (1995).
16. N. Kalach, F. Rocchiccioli, D. de Boissieu, P. H. Benhamou, C. Dupont, Intestinal permeability in children: variation with age and reliability in the diagnosis of cow's milk allergy. *Acta Paediatr* **90**, 499-504 (2001).
17. M. G. Darby *et al.*, Pre-conception maternal helminth infection transfers via nursing long-lasting cellular immunity against helminths to offspring. *Sci Adv* **5**, eaav3058 (2019).
18. L. J. Ma, B. Walter, A. Deguzman, H. K. Muller, A. M. Walker, Trans-epithelial immune cell transfer during suckling modulates delayed-type hypersensitivity in recipients as a function of gender. *PLoS ONE* **3**, e3562 (2008).
19. L. Zhou *et al.*, Two independent pathways of maternal cell transmission to offspring: through placenta during pregnancy and by breast-feeding after birth. *Immunology* **101**, 570-580 (2000).
20. S. Tuboly, S. Bernath, R. Glavits, A. Kovacs, Z. Megyeri, Intestinal absorption of colostral lymphocytes in newborn lambs and their role in the development of immune status. *Acta Vet Hung* **43**, 105-115 (1995).
21. C. Balle *et al.*, Factors influencing maternal microchimerism throughout infancy and its impact on infant T cell immunity. *BioRxiv 432825 [**Preprint**]* **March 2nd, 2021**, https://doi.org/10.1101/2021.1103.1102.432825 (2021).
22. L. A. Jackson *et al.*, An mRNA Vaccine against SARS-CoV-2 - Preliminary Report. *N Engl J Med* **383**, 1920-1931 (2020).
23. F. P. Polack *et al.*, Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N Engl J Med* **383**, 2603-2615 (2020).
24. A. Y. Collier *et al.*, Immunogenicity of COVID-19 mRNA Vaccines in Pregnant and Lactating Women. *JAMA*, (2021).
25. G. A. Rempala, M. Seweryn, Methods for diversity and overlap analysis in T-cell receptor populations. *J Math Biol* **67**, 1339-1368 (2013).
26. J. Rytlewski *et al.*, Model to improve specificity for identification of clinically-relevant expanded T cells in peripheral blood. *PLoS ONE* **14**, e0213684 (2019).
27. P. Dash *et al.*, Quantifiable predictive features define epitope-specific T cell receptor repertoires. *Nature* **547**, 89-93 (2017).
28. K. Mayer-Blackwell *et al.*, TCR meta-clonotypes for biomarker discovery with tcrdist3: identification of public, HLA-restricted SARS-CoV-2 associated TCR features. *bioRxiv*, (2021).
29. A. Gil, M. B. Yassai, Y. N. Naumov, L. K. Selin, Narrowing of human influenza A virus-specific T cell receptor alpha and beta repertoires with increasing age. *J Virol* **89**, 4102-4116 (2015).
30. G. Chen *et al.*, Sequence and Structural Analyses Reveal Distinct and Highly Diverse Human CD8(+) TCR Repertoires to Immunodominant Viral Antigens. *Cell Rep* **19**, 569-583 (2017).
31. W. D. Chronister *et al.*, TCRMatch: Predicting T-Cell Receptor Specificity Based on Sequence Similarity to Previously Characterized Receptors. *Front Immunol* **12**, 640725 (2021).
32. S. Nolan *et al.*, A large-scale database of T-cell receptor beta (TCRbeta) sequences and binding associations from natural and synthetic exposure to SARS-CoV-2. *Res Sq*, (2020).
33. A. A. Minervina *et al.*, Convergent epitope-specific T cell responses after SARS-CoV-2 infection and vaccination. *medRxiv*, (2021).
34. A. S. Shomuradova *et al.*, SARS-CoV-2 Epitopes Are Recognized by a Public and Diverse Repertoire of Human T Cell Receptors. *Immunity* **53**, 1245-1257 e1245 (2020).

35. B. Agerer *et al.*, SARS-CoV-2 mutations in MHC-I-restricted epitopes evade CD8(+) T cell responses. *Sci Immunol* **6**, (2021).
36. A. G. Dykema *et al.*, Functional characterization of CD4+ T cell receptors crossreactive for SARS-CoV-2 and endemic coronaviruses. *J Clin Invest* **131**, (2021).
37. T. H. O. Nguyen *et al.*, CD8(+) T cells specific for an immunodominant SARS-CoV-2 nucleocapsid epitope display high naive precursor frequency and TCR promiscuity. *Immunity* **54**, 1066-1082 e1065 (2021).
38. T. M. Snyder *et al.*, Magnitude and Dynamics of the T-Cell Response to SARS-CoV-2 Infection at Both Individual and Population Levels. *medRxiv*, (2020).
39. S. L. Goff, D. N. Danforth, The Role of Immune Cells in Breast Tissue and Immunotherapy for the Treatment of Breast Cancer. *Clin Breast Cancer* **21**, e63-e73 (2021).
40. D. Ramanan *et al.*, An Immunologic Mode of Multigenerational Transmission Governs a Gut Treg Setpoint. *Cell* **181**, 1276-1290 e1213 (2020).
41. A. Ssemaganda *et al.*, Expansion of tissue-resident CD8+ T cells and CD4+ Th17 cells in the nasal mucosa following mRNA COVID-19 vaccination. *bioRxiv*, (2021).
42. H. S. Robins *et al.*, Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood* **114**, 4099-4107 (2009).
43. ImmunoMind Team. (2019). immunarch: An R Package for Painless Bioinformatics Analysis of T-Cell and B-Cell Immune Repertoires. Zenodo. http://doi.org/10.5281/zenodo.3367200.
44. D. V. Bagaev *et al.*, VDJdb in 2019: database extension, new analysis infrastructure and a T-cell receptor motif compendium. *Nucleic Acids Res* **48**, D1057-D1062 (2020).
45. S. Y. Chen, T. Yue, Q. Lei, A. Y. Guo, TCRdb: a comprehensive database for T-cell receptor sequences with powerful search function. *Nucleic Acids Res* **49**, D468-D474 (2021).
46. N. Tickotsky, T. Sagiv, J. Prilusky, E. Shifrut, N. Friedman, McPAS-TCR: a manually curated catalogue of pathology-associated T cell receptor sequences. *Bioinformatics* **33**, 2924-2929 (2017).
47. R. Vita *et al.*, The immune epitope database (IEDB) 3.0. *Nucleic Acids Res* **43**, D405-412 (2015).

Acknowledgements:

We thank all those who donated blood and breastmilk for this work as well as Drs. Christine Johnston and Anna Wald for providing control specimens.

Funding:

NIH grant T32AI007509 (BA)

NIH grant K08AI148588 (ESF)

NIH grant T32CA080416 (SJ)

NIH grant K23AI153390-01 (AK)

NIH grant 1UM1AI148373-01 (RNC)

NIH grant 3UM1AI148373-01S1 (RNC)

NIH grant R21AI163999 (DMK)

NIH grant K08AI135072 (WEH)

NIH contract 75N93019C00063 (DMK)
Burroughs Wellcome Fund CAMS grant 1017213 (WEH)
University of Washington and Seattle Children's (WEH)

Author contributions:

Conceptualization: BA, YJ, WEH
Methodology: YJ, MC, ESF, SJ, JH, XW, TP, SEL, RNC, DMK, WEH
Participant enrollment: AK, WY, TN, WEH
Investigation: BA, YJ, MC, ESF, SJ, JH, XW, TP, DMK, WEH
Visualization: BA, MC, WEH
Funding acquisition: BA, ESF, SJ, AK, RNC, DMK, WEH
Supervision: NM, RNC, DMK, WEH
Writing – draft: BA, YJ, WEH
Writing – critical review/editing: BA, YJ, ESF, SJ, XW, NM, SEL, DMK, WEH

Competing interests: The authors have no competing interests.

Data and materials availability: All flow cytometry data are available in the main manuscript or supplementary files, and all TCR sequencing data will be deposited in immuneACCESS®.

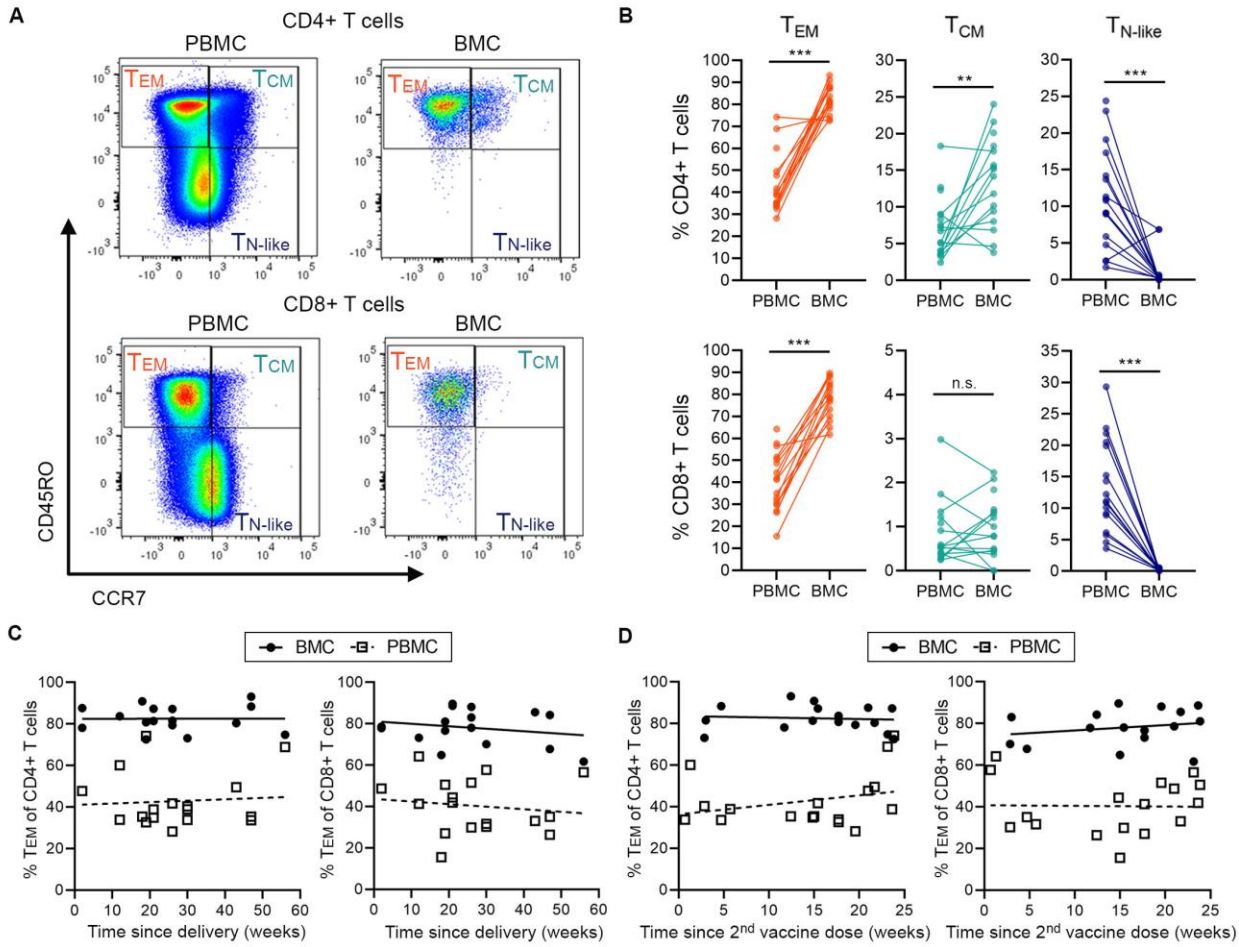

Fig. 1. Antigen-experienced T cells are enriched in breastmilk. The distribution of memory and naïve populations of CD4+ and CD8+ T cells within PBMC and BMC (n=16 BMC, n=17 PBMC; n=14 PBMC/BMC pairs). Comparisons made with linear regression and clustering by individual, ***p< 0.001, **p< 0.01, *p< 0.05, n.s. (not significant). (**A**) Scatter plots from one representative participant. (**B**) Frequencies of T_{EM}, T_{CM}, and $T_{N\text{-like}}$ CD4+ and CD8+ T cells in PBMC and BMC: CD4+ T_{EM}: 43% vs. 83%, p<0.001, CD4+ T_{CM}: 7% vs. 14%, p=0.003, CD4+ $T_{N\text{-like}}$: 11% vs. 1%, p<0.001; CD8+ T_{EM}: 40% vs.78%, p< 0.001, CD8+ T_{CM}: 0.8% vs. 1%, p=n.s., CD8+ 13% vs. 0.2%, $T_{N\text{-like}}$: p<0.001. (**C**) Frequencies of CD4+ and CD8+ T_{EM} cells in BM and PBMC did not vary by time since delivery except PBMC CD4+ T_{EM}: delta=0.09%/week, p=0.08. (**D**) Frequencies of CD4+ and CD8+ T_{EM} cells in BM and PBMC did not vary by time since 2nd vaccine dose.

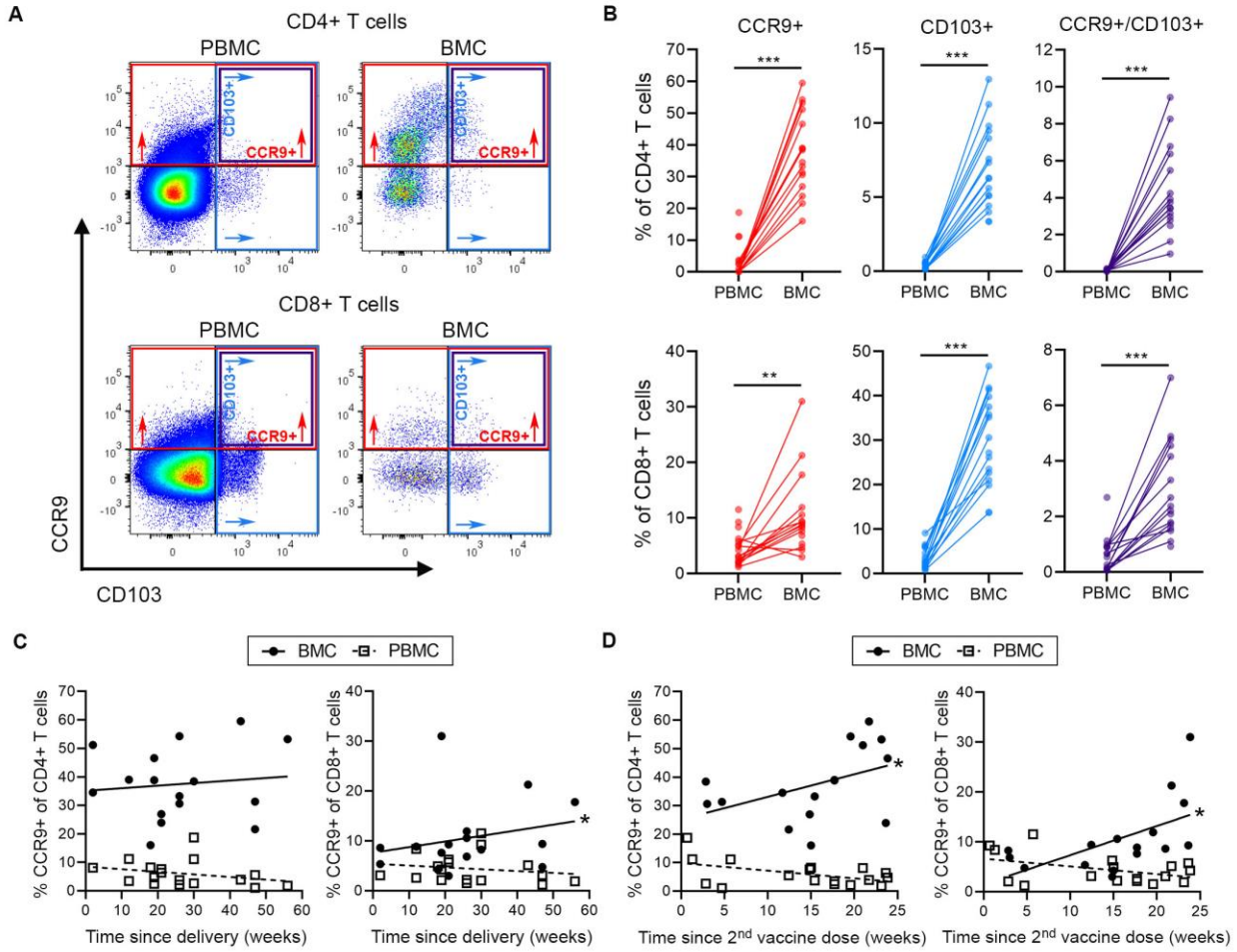

Fig. 2. **Breastmilk T cells display high levels of mucosal-homing markers.** CCR9 and CD103 expression in CD4+ and CD8+ T cell populations from PBMC and BMC (n=16 BMC, n=17 PBMC; n=14 PBMC/BMC pairs). Comparisons made with linear regression and clustering by individual, ***p<0.001, **p<0.01, *p<0.05. (**A**) Scatter plots from one representative participant. (**B**) Expression of CCR9 and CD103 within CD4+ and CD8+ T cell populations in PBMC and BMC: CD4+/CCR9+: 6% vs. 37%, p<0.001, CD4+/CD103+: 0.4% vs. 7%p, <0.001, CD4+/CCR9+/CD103+: 0.07 vs. 4%, p<0.001; CD8+/CCR9+: 4% vs. 11%, p=0.006, CD8+/CD103+: 4% vs. 30%, p<0.001, CD8+/CCR9+/CD103+: 0.6% vs. 3% p<0.001. (**C**) CCR9 expression in CD4+ and CD8+ T cells as a function of time since delivery: BMC CD8/CCR9+: delta=+0.1%/week, p=0.01. PBMC CD4+/CCR9+: delta=-0.1%/week, p =0.04. (**D**) CCR9 expression in CD4+ and CD8+ T cells as a function of time since 2nd vaccine. BMC CD4+/CCR9+:

delta=+0.8%/week, p=0.05; BMC CD8/CCR9+: delta=+0.6%/week, p=0.02. PBMC CD4+/CCR9+: delta=-0.3%/week, p=0.1; PBMC CD8/CCR9+: delta=-0.1%/week, p=0.1.

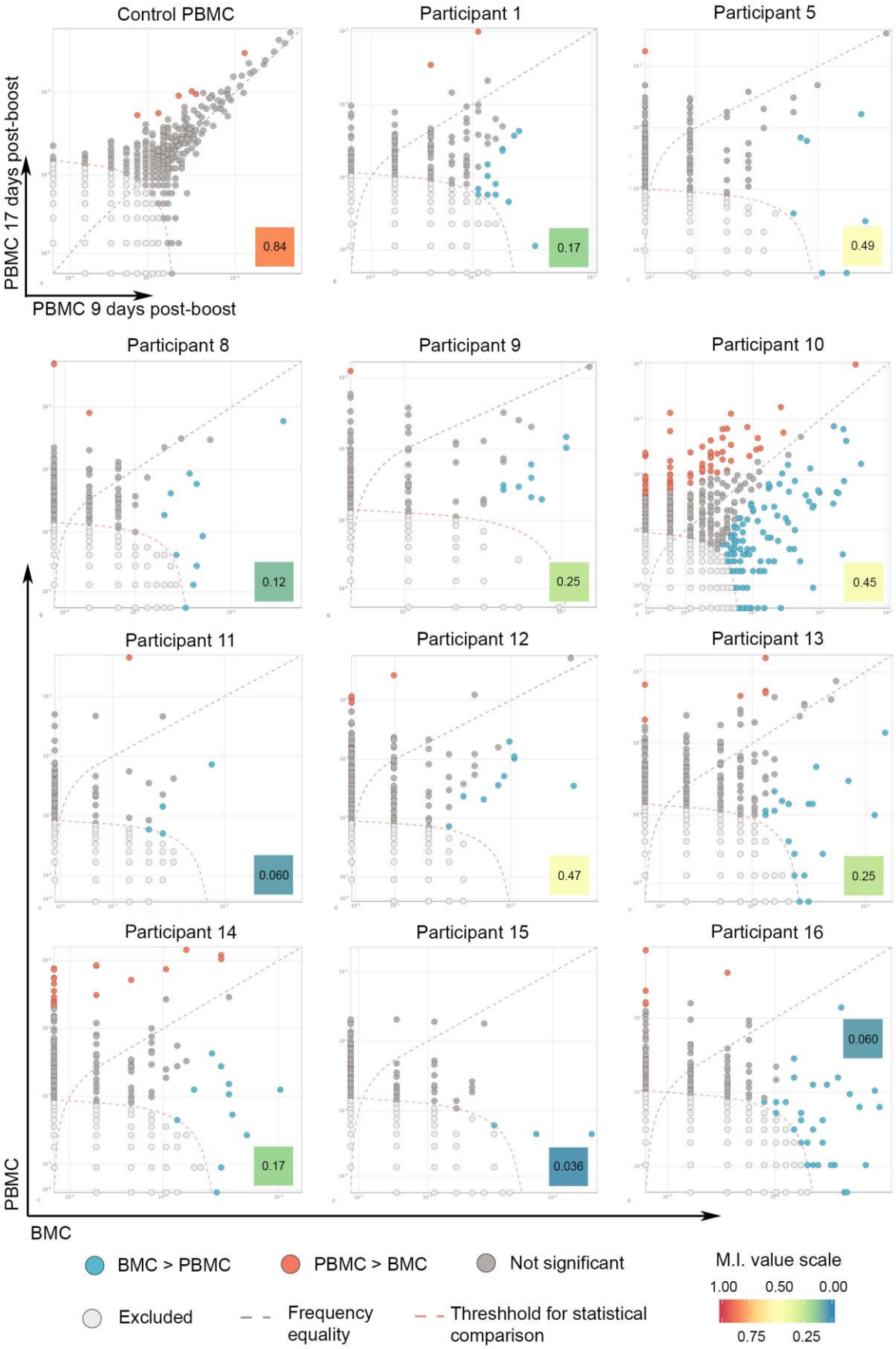
Control PBMC
Participant 1
Participant 5
PBMC 17 days post-boost
PBMC 9 days post-boost
0.84
0.17
0.49
Participant 8
Participant 9
Participant 10
0.12
0.25
0.45
Participant 11
Participant 12
Participant 13
0.060
0.47
0.25
Participant 14
Participant 15
Participant 16
0.060
0.17
0.036
PBMC
BMC
BMC > PBMC
PBMC > BMC
Not significant
M.I. value scale
1.00
0.50
0.00
0.75
0.25
Excluded
Frequency equality
Threshhold for statistical comparison

Fig. 3. The T cell receptor (TCR) repertoires in breastmilk and peripheral blood are distinct. Bulk TCRβ sequencing from BMC and PBMC individuals who had paired samples available and at least 1,000 TCRβ templates in the BMC sample (n = 11). The frequencies of TCRβ clonotypes in the two compartments were compared using the ImmunoSEQ® Differential Abundance Tool. TCRβ repertoire overlap was analyzed using the Morisita Index (M.I., value inset). As a control, TCRβ clonotypes from an individual's PBMCs obtained 9 days and 17 days after 2nd mRNA vaccine dose were compared (upper left plot).

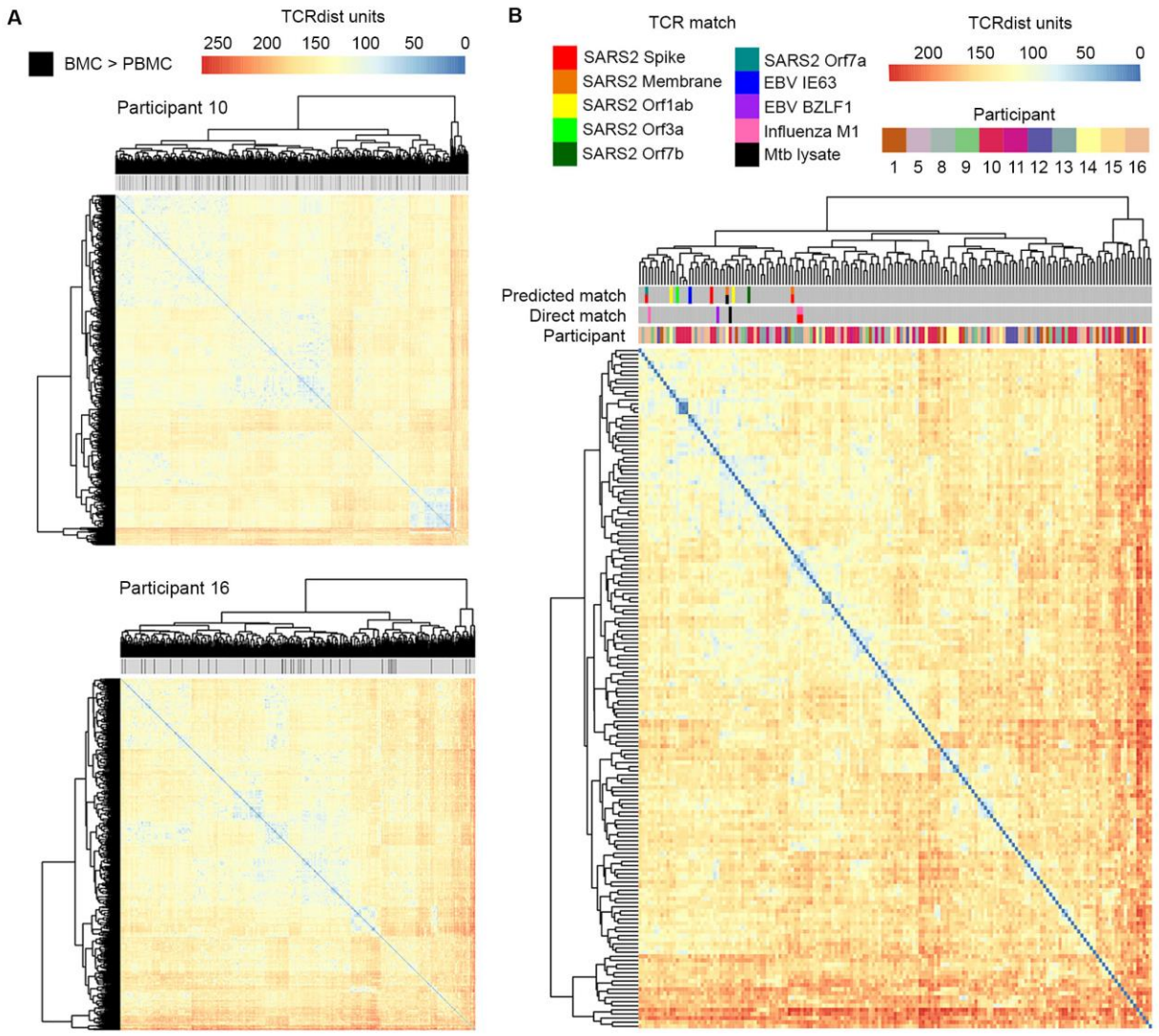

Fig. 4. Overabundant TCR clones in breastmilk are structurally diverse. (A) All TCRβ CDR3 amino acid sequences obtained from the BMC of each participant were compared using TCRdist3; representative plots from two individuals are displayed. Black ticks denote TCRβ sequences significantly overabundant in BMC relative to PBMC. (**B**) Overabundant BMC TCRβ CDR3 amino acid sequences from 11 participants were compared to one another using TCRdist3. Overabundant TCRβ clones were compared by CDR3 amino acid sequence and V gene usage against available public databases of known TCR epitope specificity. Clones matching pathogen-specific epitopes are marked with colored ticks. For Participant 10, only CDR3 amino acid

sequences enriched by a factor ≥50 relative to PBMC or with epitope specificity were included to reduce data skewing from this participant.

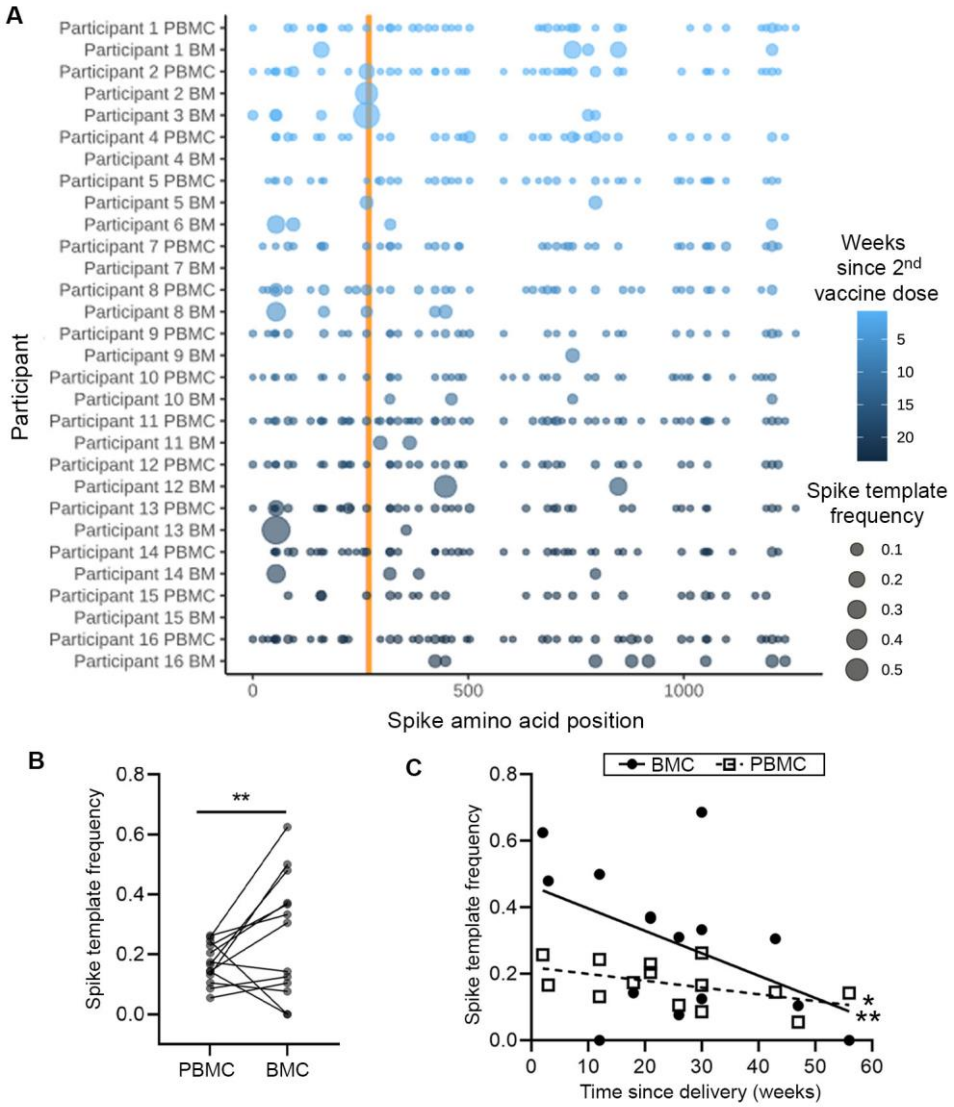

Fig. 5. Candidate SARS2 spike-specific T cells are enriched in the breastmilk relative to the peripheral blood of vaccinated individuals. Data from all sequenced PBMC and BMC (n=16 BMC, n=14 PBMC). The frequency of candidate Spike-specific clones is expressed relative to all clones predicted to bind to SARS2 antigens. Relative abundance compared using negative binomial models, **p<0.01, *p<0.05. (**A**) TCRβ sequences predicted to bind to SARS2 were identified in BMC and PBMC using the ImmunoSEQ® T-MAP COVID Search Tool and are mapped by their epitope binding location on the Spike protein. Orange line indicates the amino acid position of the Spike peptide pool that includes the YLQPRTFLL epitope. (**B**) Spike-specific TCRβ templates are enriched in BMC relative to PBMC, incident rate ratio (IRR)=1.73, p=0.001.

(**C**) Frequency of Spike-specific TCRβ templates as a function of time since delivery: BMC: IRR=0.98/week, p=0.005; PBMC: IRR=0.99/week, p=0.04.

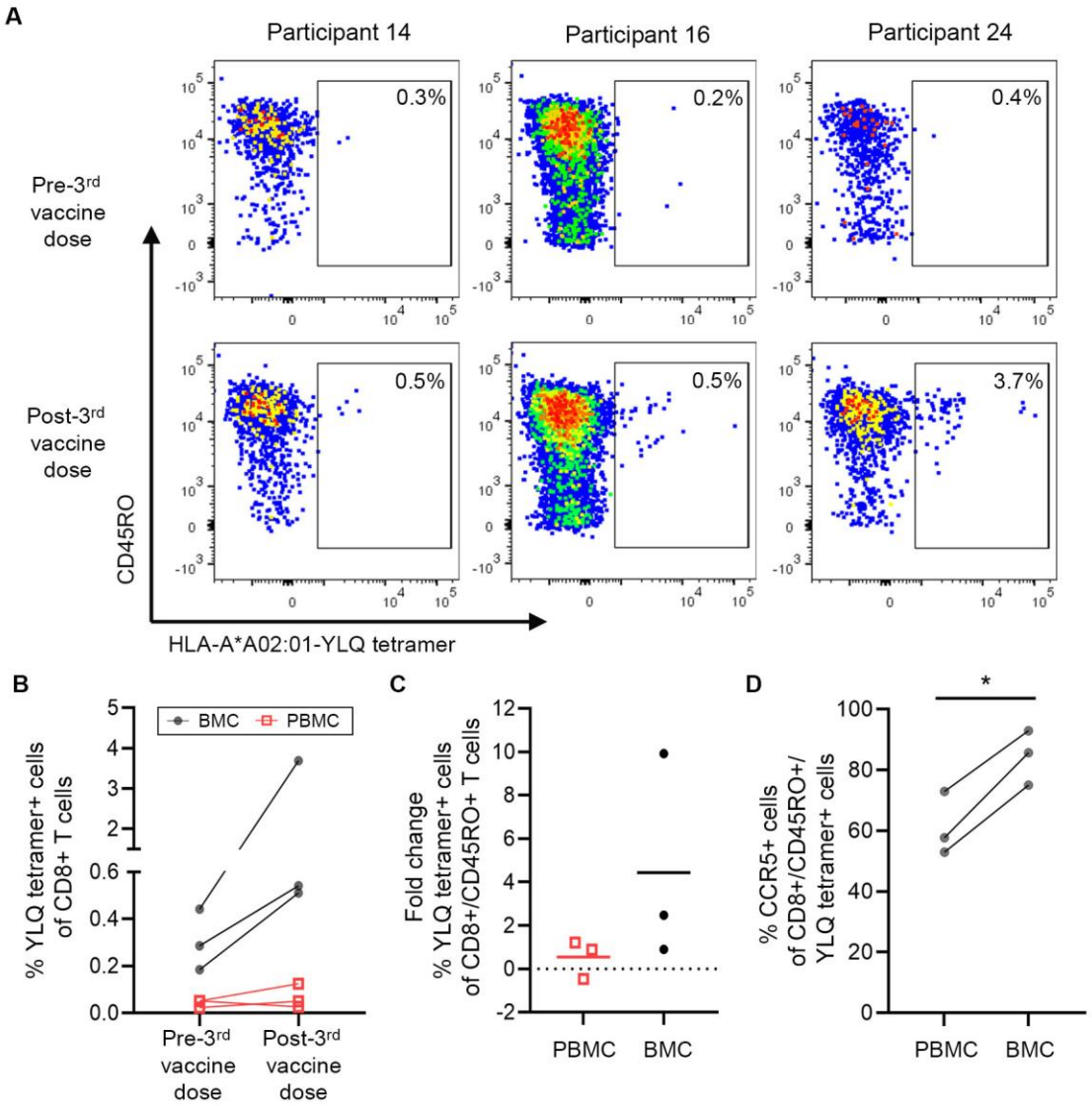

Fig. 6. SARS2 Spike-specific cells expand in breastmilk following 3rd mRNA vaccine dose. The frequencies of HLA-A2*02:01-YLQPRTFLL (YLQ) tetramer positive cells before and after 3rd dose of mRNA vaccine in BMC and PBMC from HLA-A*02:01 positive individuals. Comparisons made with paired t tests, *p<0.05. (**A**) Flow plots comparing frequency of tetramer + cells out of total CD8+ cells in BMC from three individuals. (**B**) Frequency of tetramer + cells out of CD8+ cells in BMC and PBMC. BMC: 0.3% vs. 1.6%, p=0.3, PBMC: 0.04% vs. 0.07% p=0.5. (**C**) Fold change pre to post-3rd dose in frequency of tetramer + cells out of CD8+/CD45RO+ cells in PBMC versus BMC (0.5 vs. 4.4, p=0.3). (**D**) Frequency of CCR5+ cells out of tetramer+ /CD45RO+/CD8+ T cells post-3rd dose in PBMC and BMC (p=0.01).