



# Antibody Repertoire Analysis of Tumor-Infiltrating B Cells Reveals Distinct Signatures and Distributions Across Tissues

Ligal Aizik<sup>1</sup>, Yael Dror<sup>1</sup>, David Taussig<sup>1</sup>, Adi Barzel<sup>2</sup>, Yaron Carmi<sup>3</sup> and Yariv Wine<sup>1\*</sup>

<sup>1</sup> The Shmunis School of Biomedicine and Cancer Research, The George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel, <sup>2</sup> The School of Neurobiology, Biochemistry and Biophysics, The George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel, <sup>3</sup> Department of Pathology, Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel

## OPEN ACCESS

### Edited by:

Peter Sims,  
Columbia University, United States

### Reviewed by:

Victor Greiff,  
University of Oslo, Norway  
Li Tang,  
École Polytechnique Fédérale de  
Lausanne, Switzerland

### \*Correspondence:

Yariv Wine  
yarivwine@tauex.tau.ac.il

### Specialty section:

This article was submitted to  
Systems Immunology,  
a section of the journal  
Frontiers in Immunology

**Received:** 05 May 2021

**Accepted:** 29 June 2021

**Published:** 19 July 2021

### Citation:

Aizik L, Dror Y, Taussig D,  
Barzel A, Carmi Y and Wine Y  
(2021) Antibody Repertoire  
Analysis of Tumor-Infiltrating B Cells  
Reveals Distinct Signatures and  
Distributions Across Tissues.  
*Front. Immunol.* 12:705381.  
doi: 10.3389/fimmu.2021.705381

The role of B cells in the tumor microenvironment (TME) has largely been under investigated, and data regarding the antibody repertoire encoded by B cells in the TME and the adjacent lymphoid organs are scarce. Here, we utilized B cell receptor high-throughput sequencing (BCR-Seq) to profile the antibody repertoire signature of tumor-infiltrating lymphocyte B cells (TIL-Bs) in comparison to B cells from three anatomic compartments in a mouse model of triple-negative breast cancer. We found that TIL-Bs exhibit distinct antibody repertoire measures, including high clonal polarization and elevated somatic hypermutation rates, suggesting a local antigen-driven B-cell response. Importantly, TIL-Bs were highly mutated but non-class switched, suggesting that class-switch recombination may be inhibited in the TME. Tracing the distribution of TIL-B clones across various compartments indicated that they migrate to and from the TME. The data thus suggests that antibody repertoire signatures can serve as indicators for identifying tumor-reactive B cells.

**Keywords:** antibody repertoire, BCR-Seq, tumor infiltrating lymphocytes, next generation sequencing, VDJ recombination, B cell, triple negative breast cancer, AIRR-seq

## INTRODUCTION

B cells and T cells are distinct effector cells in the adaptive arm of the immune system; yet, they often act in synergy to eradicate pathological processes. For more than two decades, research on the immune response to tumors has focused mainly on T cells, with B cells being under investigated and their roles in the tumor microenvironment (TME) remaining controversial (1, 2). Recent studies have, however, revealed that B cells may play a crucial role in tumor immunity, as these cells consistently comprise a substantial cellular component of the TME. Indeed, it has been reported that tumor-infiltrating lymphocyte B cells (TIL-Bs) can reach up to 40% of all TILs in different types of cancer (3, 4). Moreover, several reports have described tertiary lymphoid structures (TLS) that contain a relatively high portion of TIL-Bs (5–7).

Several possible functions have been attributed to TIL-Bs. Among these functions, they can serve as antibody-secreting cells, namely, as plasma cells producing antibodies that inhibit tumor growth by targeting tumor-associated antigens (TAAs) and inducing antibody-dependent cell cytotoxicity or complement-dependent cytotoxicity (8). TIL-Bs can also function as antigen-presenting cells that trigger an anti-tumor T cell response (9, 10). Of particular importance, TIL-Bs can exhibit direct cytotoxic activity, killing tumor cells *via* the Fas/FasL pathway (11): Activated B cells in tumor-draining lymph nodes (DLNs) have been shown to express the Fas ligand, to be upregulated upon engagement with cells of the 4T1 triple-negative breast cancer (TNBC) cell line, and, in turn, to exhibit cytotoxic activity (12). The discovery of regulatory B cells that secrete cytokines suggests an additional possible role of B cells in the TME (13, 14). It has been reported, for example, that the frequency of interleukin (IL)-10-producing B cells in the TME of some cancer types is negatively correlated with that of interferon (IFN)- $\gamma$ -producing CD8<sup>+</sup> T cells but positively correlated with that of regulatory Foxp3<sup>+</sup> CD4<sup>+</sup> T cells. These data thus demonstrate a possible regulatory role for B cells in the TME (15). The above research notwithstanding, reports describing TIL-B activity in cancer remain contradictory, with some studies claiming TIL-B activity to be immunosuppressive (16–18) and others assigning protective roles to these cells (19–21). These contradictory pieces of evidence have led cancer scientists to hold differing views about whether immunotherapies should be aimed to enhance or inhibit the activity of the B cells in the TME.

The analysis of antibodies that become anchored to the membrane of the B cell as B cell receptors (BCRs) was made possible by virtue of advances in the next-generation sequencing (NGS) of BCRs (BCR-Seq), which has revolutionized our understanding of B-cell-mediated immunity. This technology, coupled with the development of advance computational tools, has enabled the elucidation of the enormous antibody diversity, which, in turn, provides important information about the nature of the B cell response following a challenge (22).

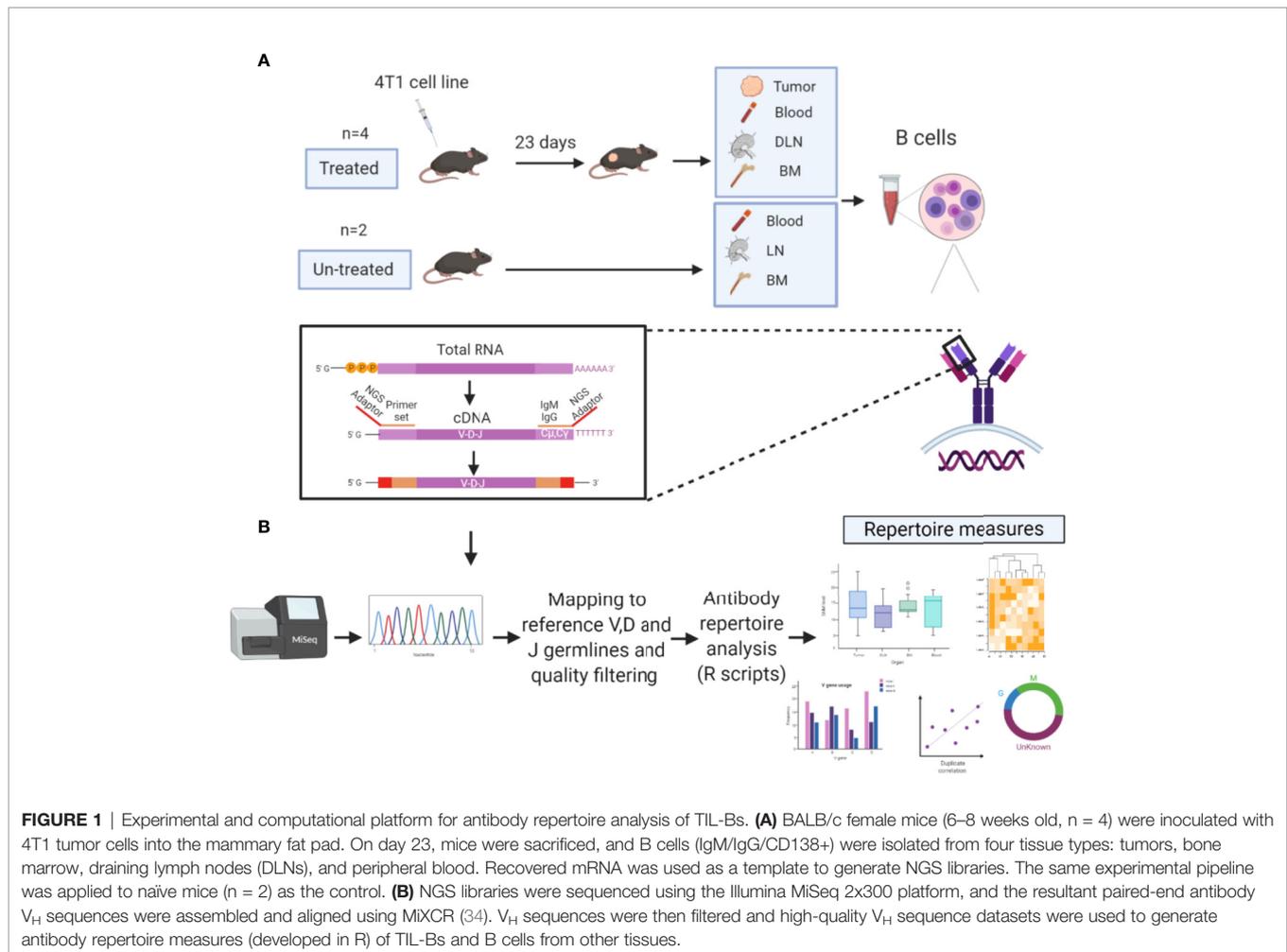
BCR-Seq enables the exploration of antibody repertoire measures that are associated with the unique diversification mechanisms that B cells undergo during their development. Initially, B cells in the bone marrow are subjected to a V(D)J chromosomal rearrangement process, followed by N/P non-templated nucleotide deletion/insertion (23). Subsequently, in response to an antigenic challenge, B cells are clonally expanded and diversify further in a process known as affinity maturation, which includes somatic hypermutation (SHM), followed by clonal selection for B cells that encode productive high-affinity antibodies (24). In addition, class-switch recombination (CSR), in which the antibody isotype is altered (e.g., IgM to IgG), contributes to the effector function that is attributed to antibodies. The above-mentioned mechanisms facilitate the generation of an enormous diversity of antibodies that can exceed the theoretical diversity of  $10^{13}$  (25). Thus, repertoire measures, such as B cell clonal architecture, SHM rate and isotype distribution, can be exploited to facilitate our understanding of the antigen-driven B cell response in cancer (26–29).

To provide new insights into the functional role of B cells in the TME, we performed BCR-Seq on B lymphocytes from four tissue types, namely, tumors, DLNs, blood, and bone marrow, derived from TNBC-induced mice. We tracked the antibody repertoire measures of the TIL-B clones compared to the measures of the B cell clones in the other three tissue types with the aim to identify the signatures that characterize B cells engaged with tumor cells and/or TAAs and to determine whether they are subjected to positive selection processes following an antigen-driven response in the TME and the adjacent DLNs. We found that the TIL-B compartment is dominated by a restricted number of highly expanded clones exhibiting high rates of SHM. Interestingly, these clones were mostly IgM<sup>+</sup> clones, suggesting that the CSR of TIL-Bs is impaired. Moreover, a particular subset of TIL-Bs was found in all compartments, suggesting that this specific subset of B cells migrates from and to the TME.

## RESULTS

### Experimental and Computational Study Design

Tracking B cell clones, either in a temporal manner or between tissues, can be achieved by identifying their antibody variable regions and particularly their complementarity-determining region 3 (CDR3) of the heavy chain (CDRH3), which is used as a B cell clonal “barcode.” Since the CDRH3 is a product of the V(D) J rearrangement, it exhibits the highest diversity in the antibody heavy chain variable region ( $V_H$ ) and is thought to play a key role in antigen recognition (25, 30–33). Since it is generally accepted that the CSR from IgM to IgG represents progression of the immune response (34, 35), we thus focused our analysis on the antibody  $V_H$  of both the IgG and IgM isotypes. To study the repertoire measures of TIL-Bs and to track their clonal distribution across tissues, we established an experimental and computational platform for the analysis of the  $V_H$  sequences obtained from four mice treated with 4T1 cells (vs two untreated mice as control) (**Figure 1A**). At 23 days post tumor-cell inoculation, B cells were collected from four tissue types (tumor, blood, DLNs and bone marrow) and lysed for purification of total mRNA. Recovered mRNA was reverse transcribed to generate cDNA, and the  $V_H$  sequences were amplified by using a primer set specific for the  $V_H$  and the constant region 1 of the heavy chain ( $C_{H1}$ ) of the *IGHM* and *IGHG* genes (**Supplementary Table S1**). Generated  $V_H$  amplicons were prepared for multiplex sequencing on a MiSeq Illumina platform (2×300 bp) by the addition of adaptors and sample-specific barcodes. For the untreated mice, B cells from three tissue types were isolated and subjected to the same process as that for the treated mice. Noteworthy, the general notion behind the present experimental design was to allow sufficient time for the immune system to generate a wide variety of B cells, including memory and plasma cells. While most vaccination protocols in mice spans over 4 weeks, we were limited by a process of necrosis and spontaneous induction of lymph node and lung metastases.

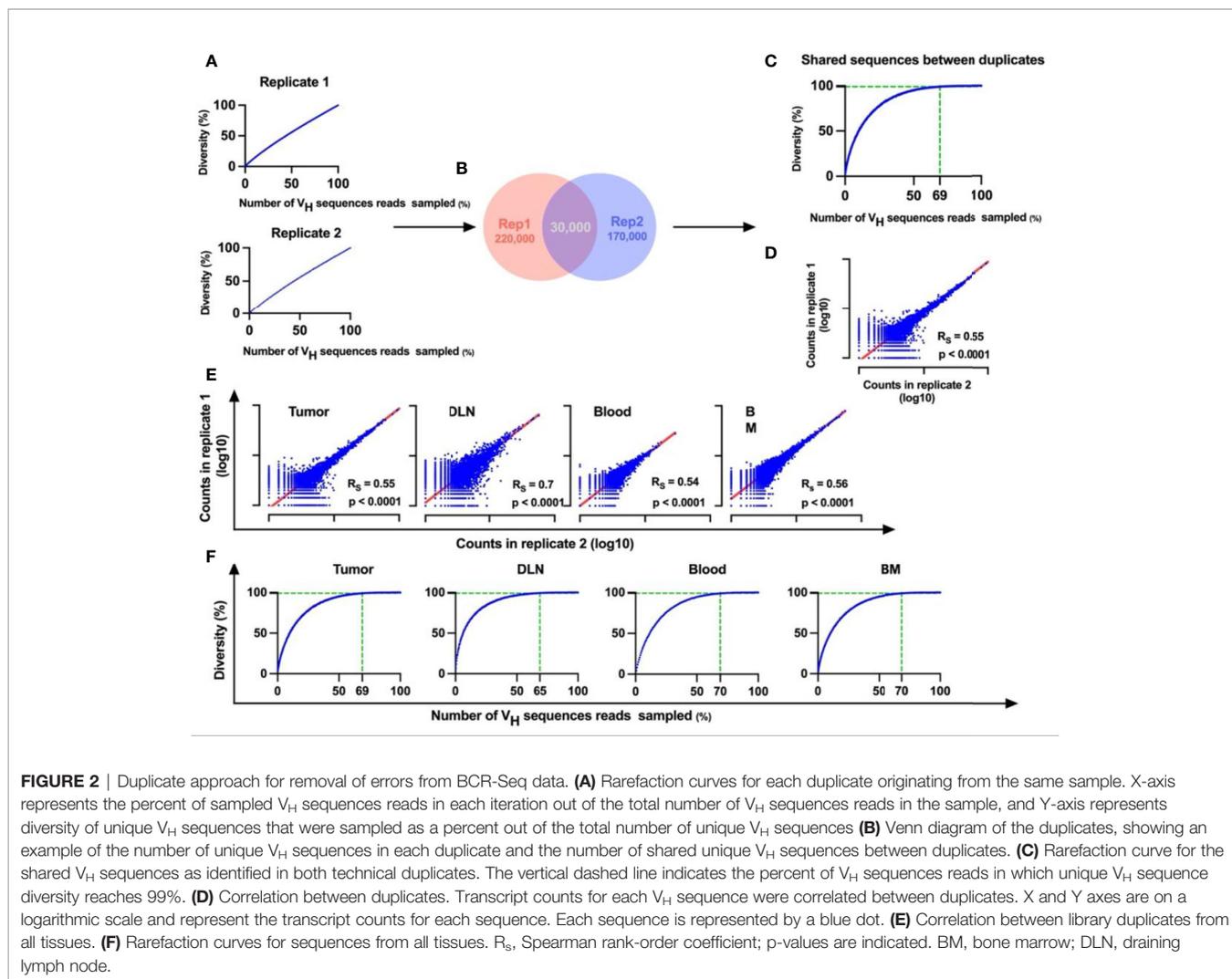


BCR-Seq of B cells from tissues of all four mice yielded a total of  $5.3 \times 10^7$  sequences.  $V_H$  sequences were paired-end aligned and annotated by alignment to reference germlines, as found in the ImMunoGeneTics server [IMGT (36)] using MiXCR (37), to obtain annotation by regions including the frameworks (FRs), the CDRs, and the full-length VDJ region. Aligned  $V_H$  sequences were filtered according to chosen criteria (see *Methods*) before establishing the antibody sequence dataset to be used for generating the repertoire measures (Figure 1B).

## BCR-Seq Analysis

To reduce the errors that are introduced during PCR amplification and sequencing, we carried out BCR-Seq on technical duplicates from each sample. Using the technical duplicates approach was shown to work well in reducing noise derived from error as previously described (38, 39). Only  $V_H$  sequences that appeared in both replicates were considered valid. To estimate whether the sequencing depth was adequate, the BCR-Seq data for each sample was subjected to rarefaction analysis (a method used routinely in ecology to quantify species diversity (40); see *Methods*). When applying the rarefaction analysis on a single duplicate, we found that the

rarefaction curve did not tend to an asymptote (Figure 2A), meaning that a continuous increase of sequence depth increases unique  $V_H$  sequence diversity. This is probably a result of the errors that are continually introduced during BCR-Seq, which precludes our ability to reach saturation of the rarefaction curve. In contrast, with the technical duplicate approach, the rarefaction curve of the  $V_H$  sequences reads (that are shared between the two duplicates) did saturate, thereby indicating that the sequencing depth did indeed suffice and that errors were sufficiently removed from the datasets (Figures 2B, C and Supplementary Figure S1). Subsequent application of the Spearman rank-order correlation on the BCR-Seq duplicates showed that unique antibody  $V_H$  sequence frequencies were correlated between duplicates (Figures 2D, E and Supplementary Figure S2). Rarefaction analysis of the validated sequences that were shared between the duplicates, from all tissues, demonstrated that the curve reached an asymptote, with 60–80% of the reads accounting for 99% of the possible diversity in the sample (Figure 2F). We thus showed that the technical duplicate approach was efficient in reducing the rates of errors introduced during BCR-Seq and that sufficient sequencing coverage was reached with high reproducibility.



Details regarding the number of  $V_H$  sequences reads obtained in each computational step are summarized in **Supplementary Table S2**.

It should be noted that for all the repertoire analyses  $V_H$  sequences reads from the four treated mice were pooled, as were the  $V_H$  sequences reads from the two untreated mice, and that individual datasets for each mouse were kept separate when describing common clones between tissues.

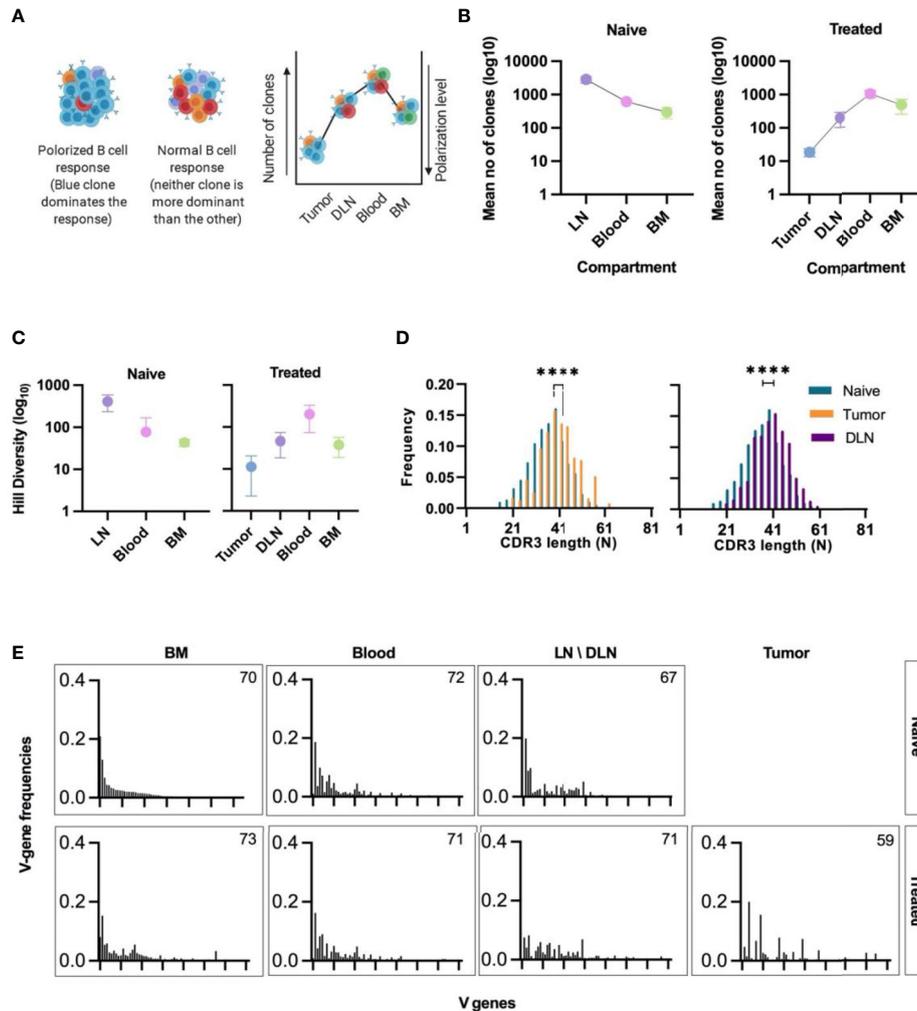
## TIL-Bs Exhibit High Clonal Polarization and Low Clonal Diversity

To determine the clonality of B cells in all tissues, we clustered the B cells on the basis of their identical V and J gene segments and 100% identity of the CDRH3 regions (nucleotide identity). The B cell clonal architecture was described in terms of two parameters: clonal polarization and clonal diversity. For the first parameter, high clonal polarization was defined as a state in which a few B cell clones contribute to the vast majority of  $V_H$  sequences reads in a particular sample (**Figure 3A**). Thus, to

quantify the clonal polarization level in each tissue, we calculated the number of B cell clones that gave rise to 80% of all  $V_H$  sequences reads in a specific tissue. We found that, on average, 20 intra-tumoral immunoglobulin clonotypes dominated the B cell response in the TME, while in the other tissues the number of clones that contributed to 80% of the reads was higher than that in the TME by 2-3 orders of magnitude (**Figure 3B**).

The second parameter that we calculated was clonal diversity, which reflects the number of unique B cell clones present in each tissue. This parameter can be used as an indicative measure to determine whether the nature of the response is oligoclonal or polyclonal. The Hill diversity (41) indicated that the clones in the tumors were significantly less diverse than those in other tissues, a finding that is in keeping with the data for clonal polarization (**Figure 3C**).

The low diversity and the high polarization level in the TIL-B response suggests that B cell clonal proliferation occurs within the tumor tissue in our TNBC model. As expected, the B cell response in the DLNs, which are adjacent to the primary tumor, also showed a high clonal polarization level



**FIGURE 3** | Clonal analysis, CDRH3 length distribution, and V gene usage of B cells isolated from different tissues. **(A)** Schematic representation of clonal polarization. Each color represents a single clone. As the number of clones increases, the polarization decreases. **(B)** Polarization curve; the mean number of clones ( $\log_{10}$ ) that contribute to 80% of all  $V_H$  sequences reads in each tissue is shown on the y-axis. The lower the value, the higher the polarization level of the B cell response. Error bars represent the standard error of the mean (SEM). **(C)** Hill diversity ( $\log_{10}$ ) – used as a measure of clonal diversity – was calculated for all clones in each tissue. Error bars represent standard deviation. **(D)** CDRH3 length distribution of intra-tumoral and DLN unique  $V_H$  sequences. The median CDRH3 length of clones from the tumor or the DLNs of treated mice (orange/purple bars) was compared to that of naïve clones (red bars). Statistical significance was determined by using an nonparametric, unpaired, two-sided, Mann-Whitney t-test. For both the DLNs and the tumor, the p-value  $<0.05$  was considered statistically significant, \*\*\*\* $p \leq 0.0001$ . **(E)** Bar plots of V gene usage frequency in each tissue. All V genes were sorted according to the frequency of V genes observed in the naïve B cells subset (IgM<sup>+</sup> B cells from the bone marrow of naïve mice). Number of V genes used in each tissue is detailed on the bar plot. Naive - naïve mice, Treated - tumor-bearing mice.

(compared to the “naïve” lymph nodes) (**Figure 3B**) and a low Hill diversity (**Figure 3C**).

We then analyzed the length distribution of the CDRH3 to be used as an indicative measure for the positive clonal selection process that may take place in the tumor or the adjacent DLNs. The average CDRH3 length in the TIL-B clones was significantly longer than the CDRH3 length of B cell clones in the bone marrow of naïve mice. The same bias in the CDRH3 length was observed in DLN-derived B cells (**Figure 3D**). Bias in the CDRH3 length distribution may suggest that TIL-Bs are subjected to a positive selection pressure as part of the affinity

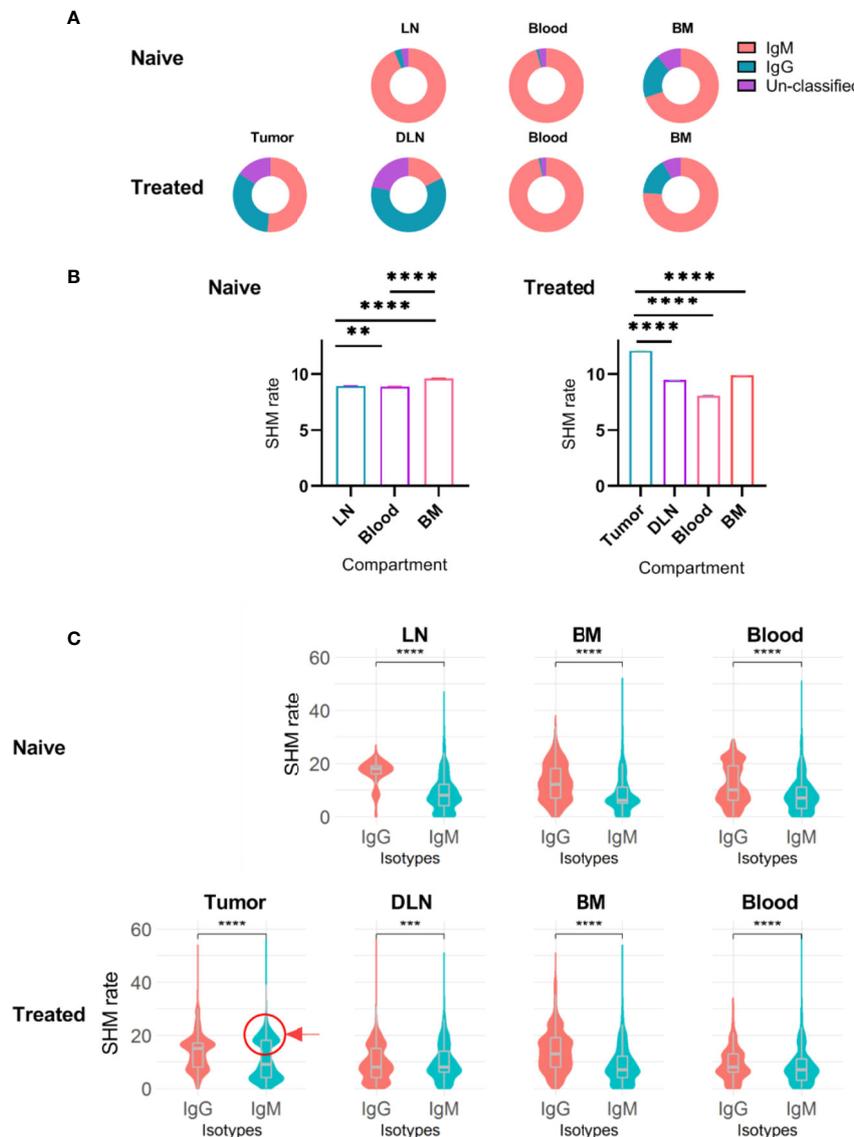
maturation stage. Thus, subsets of B cells in the TME and DLNs may be considered as tumor-reactive B cell clonotypes that were possibly activated following engagement with the tumor cells or TAAs.

An antibody repertoire can be described in terms of the frequencies with which it uses the gene segments [V(D)J], particularly the V gene segment, as it is the longest and most diverse segment (42). V gene usage in naïve B cell subsets is determined by a semi-random process, resulting in the association of the V-D-J segments during gene rearrangement as part of B cell development (43). Thus, we posit that the V gene

usage of IgM<sup>+</sup> B cells isolated from the bone marrow of naïve mice can be used to determine the “basal” V gene usage, as previously shown (44). The V gene usage in all tissues was sorted according to the descending frequencies vs the “basal” V gene usage. We found that the V gene usage signatures in the TIL-B and DLN clones were substantially different to the “basal” signature (Figure 3E and Supplementary Figure S3). This perturbation of V gene usage can be considered as an indicator of the clonal selection process that occurs following an antigenic encounter.

## TIL-Bs Are Dominated by IgM That Exhibits High SHM Rates

Our data revealed that in both naïve and treated mice, IgM was the dominant isotype in the bone marrow B cells, while a small fraction of B cells was IgG<sup>+</sup> (75% IgM and 20% IgG, Figure 4A). We observed a similar distribution when analyzing murine B cells derived from the blood; in all the mice, treated and non-treated, the dominant isotype was IgM (96%), reflecting antibodies encoded by either naïve B cells or non-CSR memory B cells (Figure 4A). Upon reaching the lymph nodes, B cells encounter



**FIGURE 4** | Isotype distribution and SHM rate in B cells isolated from different tissues. **(A)** Doughnut chart for the distribution of immunoglobulin isotypes encoded by B cells isolated from different tissues in treated and untreated mice. **(B)** Bar plots of SHM rate identified in the V<sub>H</sub> regions of B cells isolated from each tissue. Error bars represent the standard error of the mean (SEM). **(C)** Violin charts showing the distributions of the SHM rate within each immunoglobulin isotype across tissues. The violin curves show the density of values; median, lower and upper quartile values are indicated. The red circle and arrow highlight the IgM<sup>+</sup> B cell population that exhibits high SHM rates. Statistical significance for **(B, C)** was evaluated with the nonparametric, unpaired Mann-Whitney t-test (two-sided); \*\*\*\*p ≤ 0.0001, \*\*\*p ≤ 0.001, \*\*p ≤ 0.01. Naïve - naïve mice, Treated - tumor-bearing mice.

antigens and undergo affinity maturation and CSR in germinal centers (45–47). As expected, the IgG/IgM ratio encoded by B cells in the DLNs of the treated mice was higher than that in the lymph nodes of the naïve mice. Finally, TIL-Bs were mostly dominated by IgM and to a lesser extent by IgG (**Figure 4A**).

To further explore repertoire measures that are indicative of the activation state of B cells, we analyzed the SHM rates within each tissue for the different isotypes. We found that the SHM rate in the TIL-B compartment (mean of 12 mutations) was significantly higher than that in the DLNs, bone marrow and blood. Surprisingly, although the SHM rate was high, TIL-Bs were dominated by IgM (**Figures 4A, B**). To further investigate this observation, we compared the SHM rates between isotypes in all tissues. As expected, IgG<sup>+</sup> B cells exhibited a significantly higher SHM rate than IgM<sup>+</sup> B cells, since IgG<sup>+</sup> B cells are considered to be activated B cells and have thus already gone through the process of affinity maturation and CSR. In addition, examining the SHM rate distribution in the IgG<sup>+</sup> B cell subset revealed several levels of SHM, representing B cells that are at different affinity maturation stages (**Figure 4C**). Further examination of the SHM rate distribution in IgM<sup>+</sup> B cells revealed a unimodal distribution (single dominant peak) in all tissues, except in the tumor, where the distribution was bimodal (one peak with a low SHM rate and the other with a high SHM rate, **Figure 4C**). The second peak in the IgM<sup>+</sup> TIL-B subset (highlighted with a circle in **Figure 4C**) represents B cells that have undergone affinity maturation, but without CSR. A possible explanation for this interesting observation is that we harvested the TIL-Bs before they had had sufficient time to complete the CSR process. Another possible explanation is that CSR signaling was impaired due to tumorigenic activity in the TME.

## TIL-B Clonal Distribution Across Tissues

We identified, on average, 1250 B cell clones in the tumor, while the other tissues exhibited higher numbers of B cell clones (DLNs = 4700, blood = 5550, bone marrow = 12,110, **Supplementary Table S3**). To track the TIL-B clonal distribution in DLNs, blood and bone marrow, we identified clones that are common to the tumor and to any one of the three other tissue types. We thus defined clones that appear in the tumor and in at least one additional tissue as common clones and examined their frequency in each tissue (**Figure 5A**; details regarding the number of common clones are listed in **Supplementary Table S4**). The analysis of common clones was carried out for each mouse separately, because it is rare to find the common clones across different individuals. Thus, to obtain insights as to whether the B cell repertoire has convergence attributes, namely, similar repertoire measures of common clones between different mice, we looked at each mouse as a separate individual.

Among the B cell clones observed in the tumor, 20% were designated as common clones shared between the tumor and other tissues, suggesting that these clones had migrated to/from the TME. Although those common clones comprised only 20% of total clones in the TME, their V<sub>H</sub> sequences reads accounted for approximately 80% of all V<sub>H</sub> sequences reads in the TME.

We then set out to identify whether there were over-represented V genes in those common clones, namely, the clones that had been subjected to positive selection during affinity maturation, which skewed the V gene usage signature. A comparison of the semi-random V gene usage, as observed in the naïve tissue, with the V gene usage observed in the common clonotypes (**Figure 5B**) revealed 13 V genes that were significantly over-represented in the treated mice, with 8 V genes being common to at least two mice. The *IGHV3-1*, *IGHV1-7*, and *IGHV2-2* V gene was over-represented in three mice. These over-represented V genes may suggest that a converging process may have occurred in the common clones as a result of engagement with the tumor or TAAs in the TME. Interestingly, 4 V genes were found to be significantly under-represented in the common clonotypes in at least 3 mice (*IGHV7-1*, *IGHV5-6*, *IGHV1-9*, and *IGHV5-2*).

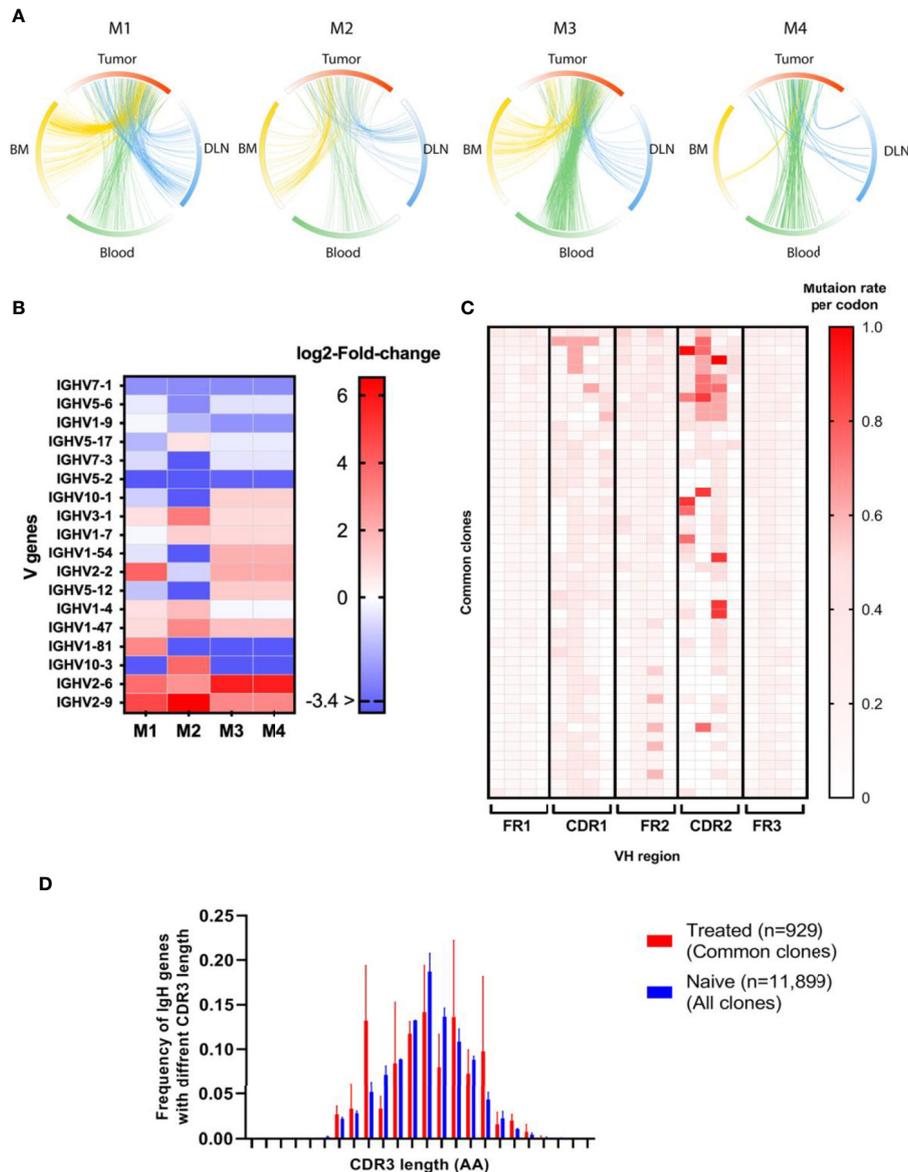
In addition, we examined whether SHM patterns were introduced preferentially along the V<sub>H</sub> regions (FRs and CDRs) within the common clonotypes. As expected, CDRH1-2 exhibited increased rates of SHM compared to the adjacent FRs (**Figure 5C**).

Next, we examined whether the distribution of CDRH3 length was skewed in the common clones versus all clones identified in the naïve mice (**Figure 5D**). While the median CDRH3 length of both the common clones and the naïve clones was 13 AA (n = 930, n = 11995, respectively), the CDRH3 length distribution of the common clonotype did indeed show a distribution pattern that was different to that of the naïve B cells.

Lastly, mice produce four IgG subclasses, namely, IgG1, IgG2A, IgG2B and IgG3. Specifically, it has been suggested that a low IgG1/IgG2A frequency ratio reflects the dominance of Th1 responses over Th2 responses. The production of IgG1 rises in response to the Th2 cytokine IL-4, and the production of IgG2A rises in response to the Th1 cytokine IFN- $\gamma$  (45, 46). To examine any alterations in the ratio of IgG1 to IgG2A frequencies, as found in the common clones, we calculated the IgG1/IgG2A ratio in B cells from the bone marrow, blood and lymph nodes of naïve mice and compared it to the ratio as found in the tumor common clones in each of the above tissue types in the tumor-bearing mice (i.e., the clones common to the tumor and the bone marrow, blood or DLN). We found that IgG1/IgG2A ratio in the common clones was, on average, twofold lower than the ratio in the naïve mice (**Supplementary Figure S4**).

## DISCUSSION

Reports on the role of TIL-Bs in breast cancer are contradictory, with some studies suggesting that the infiltration of B cells into the TME is associated with a poor prognosis (47, 48), but others suggesting that they correlate with improved prognosis (49). Regardless of the specific role of B cells in the TME, it remains an open question whether B cells in the TME are reactive towards tumor cells and whether they traffic to various body compartments. To address this question, we used antibody repertoire measures to profile and track TIL-Bs in a TNBC mouse model.



**FIGURE 5** | Antibody repertoire measures for the common clones. **(A)** Circos plots showing the clonal distribution of TIL-Bs between tissues. M1-4 indicate a circos plot for each mouse. The circos plot arcs are colored by gradients that are in accordance with the frequency of the clone within each tissue. Each connecting line represents a clone that is common between two tissues. Red - tumor, blue - DLN, green - blood, yellow - bone marrow. Details regarding the number of common clones are listed in **Supplementary Table S4** **(B)** Heat map showing  $\log_2$  fold-change of V gene frequencies between clone as identified in the bone marrow of naïve mice and common clones as identified in treated mice. Each column represents a mouse (M1-4) and each row, a V gene. Red gradient represents significantly higher frequency of V gene in the treated mice, while the blue gradient represents a significantly lower frequency of the V gene in the treated mice. Adjusted p values were considered significant using multiple correction test using Dunn-Šidák correction method **(C)** Heat map of the SHM rate per codon for each  $V_H$  region in the common clones. The red color indicates a higher SHM rate; each row represents one common clone; and each column represents a mouse. The heat map includes only the first 50 dominant common clones from each mouse. **(D)** The distribution of  $V_H$  genes with different CDRH3 lengths is shown in the bar graph. The average CDRH3 length of all the common clones for the treated mice (red bars) was compared to that for the naïve clones (blue bars). Statistical significance was determined by using an unpaired t-test. The p-value is  $p = 0.4939$ , which is considered to be not statistically significant (not shown). Error bars represent the standard deviation.

We utilized an experimental and computational platform to elucidate repertoire measures in four tissues of TNBC mice. First, we performed several quality assurance steps to make sure that the antibody sequences were of high quality and that the sequences obtained were reproducible. This was achieved by

using experimental duplicates for all NGS libraries. This approach, accompanied by applying additional filters, resulted in high-quality antibody  $V_H$  sequence datasets. The resultant high-quality datasets were then further analyzed to identify repertoire measures with distinct signatures suggesting that

TIL-Bs did indeed engage with tumor cells and/or TAAs in the TME.

This study thus provides evidence that B cell populations in the tumors and the DLNs of TNBC mice were enriched with clones that exhibit the characteristics of activated, affinity-matured B cells, as indicated by increased SHM rates and skewed V gene usage and CDRH3 length. Moreover, by comparing the clonality of TIL-Bs with that of B cells in the DLNs, bone marrow and blood, we found that TIL-Bs exhibited low clonal diversity and high clonal polarization, namely, a limited number of distinct B cell clones were prevalent and these accounted for the greater part of the B cell response in the TME. Of note, this observation of polarized clonality in the TME has been described in other breast cancer studies (50, 51) as well as in studies of ovarian cancer (52) and melanoma (26, 53), which also reported that a small number of clones dominate the B cell response in the TME. These studies taken together suggest that the B cells in the TME are constantly challenged by TAAs or by the tumor cells themselves. Moreover, the clonality of the B cells in the adjacent DLNs was also polarized, although to a lesser extent. With regard to isotype distribution, IgM<sup>+</sup> B cells in the bone marrow are immature and/or naïve-mature B cells that have developed from hematopoietic stem cells, and therefore predominantly express the IgM isotype on their surface, whereas bone marrow IgG<sup>+</sup> B cells can be considered as antigen-specific, terminally differentiated B cells, which home to the bone marrow following affinity maturation (54, 55). IgM<sup>+</sup> naïve B cells migrate from the bone marrow to secondary lymphoid organs *via* the blood circulation (54), and it has been reported that, in humans, circulating B cells in the blood comprise mostly of naïve B cells (IgM<sup>+</sup>, 70%) and memory B cells that are either class-switched or non-class switched (30%) (56). As expected, in our study, the dominant isotype in the blood was found to be IgM, but the dominant isotype in the DLNs was skewed towards IgG. Surprisingly, the dominant isotype in the TME was found to be IgM, rather than IgG, and more importantly, we identified a population of non-class switched IgM<sup>+</sup> B cells in the TME that exhibited high SHM rates; this population may represent B cells that have undergone affinity maturation even though their ability to carry out CSR has been inhibited. It is generally accepted that CSR and SHM take place in germinal centers and that these events are independent, with neither one being a prerequisite for the other (57). However, recent work re-visiting this question showed that CSR is triggered before differentiation into germinal centers and that the vast majority of CSR events occur before the onset of SHM (58, 59). Nonetheless, whether CSR starts early in the process of B cell activation or not, the finding of a subset of highly mutated IgM<sup>+</sup> B cells is intriguing in that it gives rise to the possibility that the TME generates inhibitory signals that prevent the full development of TIL-Bs and that hamper their ability to differentiate and encode for antibodies with appropriate effector functions.

We also analyzed the trafficking of TIL-Bs to other body tissues (DLNs, blood and bone marrow) by identifying common clonotypes. We found that an average of 20% of TIL-B clones are common to the tumor and at least one other tissue and that these common clones are highly prevalent in the TME. Moreover, V gene

usage of the common clones was skewed towards several dominant V genes (some of them found in two of the four mice), suggesting that these clones had been subject to positive selection. Furthermore, as expected, the SHM in the common B cell clones occurred preferentially in the CDRs, which are unstructured loops in antibody binding sites (60). Interestingly, the decrease in the IgG1/IgG2A ratio in the IgG subset of the common clones suggests that TIL-B cells encode mostly to IgG2A antibodies that fix complement (61) and bind to all activating FcγRs, rather than to the IgG1 subclass that does not fix complement and that binds well only to inhibitory FcγRIIb (62).

The identification of TIL-Bs exhibiting reactive repertoire measures (tumor-reactive B cells or TRBCs) holds great potential for application in future studies. For example, the combination of antibody repertoire analysis and global gene expression profiling will allow further explorations of the possible cross-talk between TRBCs and the TME, thus facilitating an understanding of the role of B cells in the TME, with direct implications for the development of advanced immunotherapies.

## METHODS

### Ethical Statement of Animal Studies

All mice were housed in an American Association for the Accreditation of Laboratory Animal Care-accredited animal facility and maintained under specific pathogen-free conditions. Animal experiments were approved and conducted in accordance with Tel-Aviv University Laboratory Accreditation #01-16-095. Female 8-12 weeks old mice were used in all experiments.

### In Vivo Tumor Models and Control Mice

TNBC 4T1 cells were purchased from the ATCC. BALB/C mice were injected subcutaneously into mammary fat-pad number five with  $2 \times 10^5$  4T1 cells suspended in 30  $\mu$ l of DMEM (Gibco, Thermo Fisher Scientific). The induced tumors were monitored by measuring their size twice a week using calipers. After 23 days, mice were sacrificed, and B cells were isolated from four tissue types, namely, bone marrow, blood, DLNs, and tumors. For the control mice, three tissue types were collected - bone marrow, blood, and lymph nodes. For our analyses, we chose tumors that were less than 50 mm<sup>2</sup> in size, without macroscopic necrotic areas and with a clear spatial separation from the draining lymph node.

### B Cell Isolation

All tissue preparations were taken at the same time from each mouse (after euthanasia by CO<sub>2</sub> inhalation). For isolation of B cells from the lymph nodes, the lymph nodes were removed from the euthanized mice and mashed through a 70- $\mu$ m cell strainer. Cells were then washed by centrifugation at 600 rcf for 5 min at 4–8°C. For isolation of TIL-Bs, tumors were enzymatically digested with 2,000 U/mL of DNase I and 2 mg/mL of collagenase IV (both from Sigma Aldrich, Merck, Israel) in HBSS for 30 min 37°C with a magnetic stirrer (200 rpm). Cells were then washed by centrifugation at 600 rcf for 5 min at 4–8°C. Approximately,

100K B cells were isolated following the above-mentioned approach. For obtaining B cells from peripheral blood, peripheral blood was collected by perfusion into sodium heparin-coated vacuum tubes before 1:1 dilution in HBSS supplemented with 2% FBS and 5 mM EDTA. For isolation of B cells from bone marrow, the femurs and tibias were removed from the tumor-bearing mice, ground using pestle and mortar in RPMI 1640 (Gibco, Thermo Fisher Scientific), filtered through a 70- $\mu$ M cell strainer, and washed twice with RPMI 1640. Lymphocytes from all tissues were enriched on a Histopaque-1077 Hybrid-Max (Sigma Aldrich, Merck, Israel) density gradient medium, and the collected cells were washed twice with a complete RPMI 1640. For all tissues, cells were then incubated with a mixture of anti-IgG, anti-IgM, anti-CD138 magnetic beads (MojoSort™ Nanobeads, BioLegend, Carlsbad, CA) according to the manufacturer's instructions.

### Amplification of $V_H$ Repertoires From B Cells (BCR-Seq)

Total RNA was extracted from enriched B cells using the RNeasy Micro Kit (Qiagen, cat no. 74004) according to the manufacturer's protocol. RNA concentration was determined on Epoch™ Microplate Spectrophotometer (BioTek). RNA was stored at  $-80^\circ\text{C}$ . Later, first-strand cDNA was synthesized using SuperScript™ III First-Strand Synthesis System (Thermo Scientific Cat no. 18080051) with 200 ng RNA as the template and Oligo (dT) primers (Thermo Scientific), according to the manufacturer's protocol. After cDNA synthesis, PCR amplification of the variable heavy Ig genes was performed using a set of 19 forward primers with the gene-specific regions annealing to framework 1 of the VDJ-region and two reverse primers with the gene-specific region binding to the IgG and IgM constant regions (see primer list in **Supplementary Table S3**). Primers had overhang nucleotides to facilitate Illumina adaptor addition during the second PCR. PCR reactions were carried out using FastStart™ High Fidelity DNA polymerase (Sigma-Aldrich Cat no. 3553400001) in a reaction volume of 250  $\mu$ l using 10  $\mu$ l of the cDNA product as a template, and setting the following conditions as standard:  $95^\circ\text{C}$  for 3 min; 4 cycles of  $95^\circ\text{C}$  for 30 s,  $50^\circ\text{C}$  for 30 s,  $68^\circ\text{C}$  for 1 min; 4 cycles of  $95^\circ\text{C}$  for 30 s,  $55^\circ\text{C}$  for 30 s,  $68^\circ\text{C}$  for 1 min; 20 cycles of  $95^\circ\text{C}$  for 30 s,  $63^\circ\text{C}$  for 30 s,  $68^\circ\text{C}$  for 1 min;  $68^\circ\text{C}$  for 7 min;  $4^\circ\text{C}$  storage. PCR products were purified using Agencourt AMPure XP beads (Beckman Coulter, Cat no. A63881), according to the manufacturer's protocol (ratio  $\times$  1.8 in favor of the beads). Recovered DNA products from the first PCR was applied to a second PCR amplification to attach Illumina adaptors to the amplified  $V_H$  genes using the primer extension method, as described previously (60). Second PCR reactions were carried out using Phusion® High-Fidelity DNA Polymerase (NEB) with the following cycling conditions:  $95^\circ\text{C}$  denaturations for 3 min;  $98^\circ\text{C}$  for 10 s,  $40^\circ\text{C}$  for 30 s, and  $72^\circ\text{C}$  for 30 s for two cycles;  $98^\circ\text{C}$  for 10 s,  $62^\circ\text{C}$  for 30 s, and  $72^\circ\text{C}$  for 30 s for 7 cycles; and a final extension at  $72^\circ\text{C}$  for 5 min. PCR products were subjected to 1% agarose DNA gel electrophoresis and gel-purified with Zymoclean™ Gel DNA Recovery Kit (Zymo Research) according

to the manufacturer's instructions.  $V_H$  library concentrations were measured using Qubit Flex Fluorometer (Thermo Fisher Scientific), and library quality was assessed using the 4200 TapeStation system (Agilent). DNA length of all libraries was between 540–570 bp.

### Replicate Sequencing

Technical replicates (two per sample) of BCR-Seq libraries were prepared based on cDNA from each mouse/tissue. cDNA was split, and library preparation was performed in parallel with different Illumina indices as described above.

### Illumina Sequencing, Data Analysis, and Statistics

$V_H$  libraries from sorted B cells were subjected to NGS on the MiSeq platform with the reagent kit V3  $2 \times 300$  bp paired-end (Illumina), using an input concentration of 16 pM with 5% PhiX. Forward and reverse raw fastq files were paired and CDR3 and full-length VDJ regions of successfully paired sequences were annotated using MiXCR (37) with default parameters. For downstream analyses, sequences were pre-processed and read-only retained based on the following criteria: (i)  $V_H$  sequences reads were shared between 2 replicates of the same library; (ii) both CDR3 and VDJ regions could be detected by IMGT; (iii) VDJ regions were present with a minimum abundance of 2 reads.

### Computational Analysis

**Terminology.** We use two terms for describing the  $V_H$  sequence repertoire: 1) " $V_H$  sequences reads" - this term reflects the number of reads as observed in the NGS data; 2) "Unique  $V_H$  sequence" - a result of clustering the sequences reads and creating a dataset of unique  $V_H$  sequences only. For example, a unique  $V_H$  sequence can appear 50 times in the dataset thus, have 50  $V_H$  sequences reads originating from the same sequence. However, this  $V_H$  sequence will appear only one time after clustering by  $V_H$  sequences thus, will be termed unique  $V_H$  sequence.

**Rarefaction analysis.** For the rarefaction analysis,  $V_H$  sequences reads were randomly and repetitively sampled, and the number of species (unique  $V_H$  sequences) was evaluated as a function of sample size (40, 63).

**Duplicate's correlation.** For the correlation between the replicates of the same library (duplicates), the count of shared unique  $V_H$  sequences between the two replicates was compared using Spearman's rank correlation. Statistical significance was evaluated using GraphPad prism V9.0.2.

**Repertoire measures analysis.** All repertoire measures described in this study are a result of pooling of measures from all four treated mice, and for the naïve cohort, pooling of two naïve mice, apart from the circus plots and the V gene usage in **Figure 5**.

For the clonal analysis, we defined clonally related B cells as those sharing the same  $V_H$  and  $J_H$  gene groups and the same CDR3 nucleotide sequence. Their abundance (clonal polarization) is based on  $V_H$  sequences reads frequency, for each clone. To assess clonal diversity, we applied the Hill diversity, based on Rényi's definition (41).

$$D^q = \left( \sum_{i=1}^S p_i^q \right)^{1/(1-q)}$$

where  $p_i$  is the frequency of each clone and  $S$  is the total number of clones. The  $q$  values represent weights; thus, as  $q$  increases, clones with higher frequency have a stronger impact on the obtained Hill diversity. Particularly, there are three significant and relevant  $q$  values, which has close connection to popular diversity indices:  $q = 0$ ,  $q = 1$  and  $q = 2$ .

We chose to apply  $q$  value of 2, so the abundance of the clones will have a stronger effect on the Hill diversity.

For the CDRH3 length analysis, we measured the number of nucleotides/amino acid starting at the end of FR3 (CAR amino acids) and ending with Y amino acid at the beginning of the J region.

For the V gene usage analysis, the number of unique  $V_H$  sequences annotated to each V gene was quantified in each tissue. V gene usage of IgM-expressing B cell  $V_H$  sequences reads from the bone marrow of the two naïve mice was averaged, and then sorted from the most abundant V gene to the rarest V gene. Next, all V gene frequencies were sorted according to the naïve-bone marrow V gene order.

Isotype distribution plots are based on the number of unique  $V_H$  sequences annotated to each isotype.

To examine the mutation level in each B cell repertoire, primer trimming was performed on all unique  $V_H$  sequences before including nucleotide deletions, insertions, and substitutions.

The computational platform for data validation and the generation of repertoire measures was developed in R using the following R packages: Sjmisc (64), Stringdist (65), Plyr (66), Dplyr (67), Stringr (68), Biostrings (69), Devtools (70), Rcpp (71). The majority of the graphs were generated using GraphPad prism V9.0.2 excluding violin and circus plots that were generated using R packages Tidyverse (ggplot) (72) and circlize (73) respectively. All statistical tests in the manuscript were carried using R/GraphPad prism V9.0.2.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and

accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/>, PRJNA699402. The developed R codes for antibody repertoire analysis can be accessed upon request from: <https://github.com/ligalaizik/Antibody-repertoire-analysis-R-code->.

## ETHICS STATEMENT

The animal study was reviewed and approved by Tel Aviv University.

## AUTHOR CONTRIBUTIONS

LA, AB, YC, and YW: study conceptualization and design. LA, YD, YC, and YW performed experiments. LA performed data analysis. LA and YW wrote the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

The work was partially supported by the Israel Ministry of Health (MOH) grant #3-17162.

## ACKNOWLEDGMENTS

We would like to extend our special thanks to Prof. Adit Ben-Baruch for scientific discussions and critical reading of the manuscript. Schematics in **Figures 1, 3A**, were created with Biorender.com.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2021.705381/full#supplementary-material>

## REFERENCES

1. Yuen GJ, Demissie E, Pillai S. B Lymphocytes and Cancer: A Love-Hate Relationship. *Trends Cancer* (2016) 2:747–57. doi: 10.1016/j.trecan.2016.10.010
2. Wang SS, Liu W, Ly D, Xu H, Qu L, Zhang L. Tumor-Infiltrating B Cells: Their Role and Application in Anti-Tumor Immunity in Lung Cancer. *Cell Mol Immunol* (2019) 16:6–18. doi: 10.1038/s41423-018-0027-x
3. Chung W, Eum HH, Lee HO, Lee KM, Lee HB, Kim KT, et al. Single-Cell RNA-Seq Enables Comprehensive Tumour and Immune Cell Profiling in Primary Breast Cancer. *Nat Commun* (2017) 8:15081. doi: 10.1038/ncomms15081
4. Chevrier S, Levine JH, Zanotelli VRT, Silina K, Schulz D, Bacac M, et al. An Immune Atlas of Clear Cell Renal Cell Carcinoma. *Cell* (2017) 169:736–49. doi: 10.1016/j.cell.2017.04.016
5. Cabrita R, Lauss M, Sanna A, Donia M, Skaarup Larsen M, Mitra S, et al. Tertiary Lymphoid Structures Improve Immunotherapy and Survival in Melanoma. *Nature* (2020) 577:561–5. doi: 10.1038/s41586-019-1914-8
6. Zhu G, Falahat R, Wang K, Mailloux A, Artzi N, Mulé JJ. Tumor-Associated Tertiary Lymphoid Structures: Gene-Expression Profiling and Their Bioengineering. *Front Immunol* (2017) 8:767. doi: 10.3389/fimmu.2017.00767
7. Lavin Y, Kobayashi S, Leader A, Amir ED, Elefant N, Bigenwald C, et al. Innate Immune Landscape in Early Lung Adenocarcinoma by Paired Single-Cell Analyses. *Cell* (2017) 169:750–65. doi: 10.1016/j.cell.2017.04.014
8. Mullins CS, Gock M, Krohn M, Linnebacher M. Human Colorectal Carcinoma Infiltrating B Lymphocytes Are Active Secretors of the Immunoglobulin Isotypes A, G and M. *Cancers (Basel)* (2019) 11. doi: 10.3390/cancers11060776
9. Bruno TC, Ebner PJ, Moore BL, Squalls OG, Waugh KA, Eruslanov EB, et al. Antigen-Presenting Intratumoral B Cells Affect CD4(+) TIL Phenotypes in

- Non-Small Cell Lung Cancer Patients. *Cancer Immunol Res* (2017) 5:898–907. doi: 10.1158/2326-6066.Cir-17-0075
10. Rossetti RAM, Lorenzi NPC, Yokochi K, Rosa M, Benevides L, Margarido PFR, et al. B Lymphocytes Can be Activated to Act as Antigen Presenting Cells to Promote Anti-Tumor Responses. *PLoS One* (2018) 13:e0199034. doi: 10.1371/journal.pone.0199034
  11. Tao H, Lu L, Xia Y, Dai F, Wang Y, Bao Y, et al. Antitumor Effector B Cells Directly Kill Tumor Cells via the Fas/FasL Pathway and Are Regulated by IL-10. *Eur J Immunol* (2015) 45:999–1009. doi: 10.1002/eji.201444625
  12. Xia Y, Tao H, Hu Y, Chen Q, Chen X, Xia L, et al. IL-2 Augments the Therapeutic Efficacy of Adoptively Transferred B Cells Which Directly Kill Tumor Cells via the CXCR4/CXCL12 and Perforin Pathways. *Oncotarget* (2016) 7:60461–74. doi: 10.18632/oncotarget.11124
  13. Shen M, Sun Q, Wang J, Pan W, Ren X. Positive and Negative Functions of B Lymphocytes in Tumors. *Oncotarget* (2016) 7:55828–39. doi: 10.18632/oncotarget.10094
  14. Peng B, Ming Y, Yang C. Regulatory B Cells: The Cutting Edge of Immune Tolerance in Kidney Transplantation. *Cell Death Dis* (2018) 9:109. doi: 10.1038/s41419-017-0152-y
  15. Wei X, Jin Y, Tian Y, Zhang H, Wu J, Lu W, et al. Regulatory B Cells Contribute to the Impaired Antitumor Immunity in Ovarian Cancer Patients. *Tumour Biol* (2016) 37:6581–8. doi: 10.1007/s13277-015-4538-0
  16. Inoue S, Leitner WW, Golding B, Scott D. Inhibitory Effects of B Cells on Antitumor Immunity. *Cancer Res* (2006) 66:7741–7. doi: 10.1158/0008-5472.Can-05-3766
  17. Fremd C, Schuetz F, Sohn C, Beckhove P, Domschke B. B Cell-Regulated Immune Responses in Tumor Models and Cancer Patients. *Oncoimmunology* (2013) 2:e25443. doi: 10.4161/onci.25443
  18. Schioppa T, Moore R, Thompson RG, Rosser EC, Kulbe H, Nedospasov S, et al. B Regulatory Cells and the Tumor-Promoting Actions of TNF-Alpha During Squamous Carcinogenesis. *Proc Natl Acad Sci USA* (2011) 108:10662–7. doi: 10.1073/pnas.1100994108
  19. DiLillo DJ, Yanaba K, Tedder TF. B Cells Are Required for Optimal CD4+ and CD8+ T Cell Tumor Immunity: Therapeutic B Cell Depletion Enhances B16 Melanoma Growth in Mice. *J Immunol* (2010) 184:4006–16. doi: 10.4049/jimmunol.0903009
  20. Li Q, Teitz-Tennenbaum S, Donald EJ, Li M, Chang AE. In Vivo Sensitized and *In Vitro* Activated B Cells Mediate Tumor Regression in Cancer Adoptive Immunotherapy. *J Immunol* (2009) 183:3195–203. doi: 10.4049/jimmunol.0803773
  21. Nielsen JS, Sahota RA, Milne K, Kost SE, Nesslinger NJ, Watson PH, et al. CD20+ Tumor-Infiltrating Lymphocytes Have an Atypical CD27- Memory Phenotype and Together With CD8+ T Cells Promote Favorable Prognosis in Ovarian Cancer. *Clin Cancer Res* (2012) 18:3281–92. doi: 10.1158/1078-0432.Ccr-12-0234
  22. Marks C, Deane CM. How Repertoire Data Are Changing Antibody Science. *J Biol Chem* (2020) 295:9823–37. doi: 10.1074/jbc.REV120.010181
  23. Bassing CH, Swat W, Alt FW. The Mechanism and Regulation of Chromosomal V(D)J Recombination. *Cell* (2002) 109 Suppl:S45–55. doi: 10.1016/s0092-8674(02)00675-x
  24. Mishra AK, Mariuzza RA. Insights Into the Structural Basis of Antibody Affinity Maturation From Next-Generation Sequencing. *Front Immunol* (2018) 9:117. doi: 10.3389/fimmu.2018.00117
  25. Georgiou G, Ippolito GC, Beausang J, Busse CE, Wardemann H, Quake SR. The Promise and Challenge of High-Throughput Sequencing of the Antibody Repertoire. *Nat Biotechnol* (2014) 32:158–68. doi: 10.1038/nbt.2782
  26. Cipponi A, Mercier M, Seremet T, Baurain JF, Théate I, van den Oord J, et al. Neogenesis of Lymphoid Structures and Antibody Responses Occur in Human Melanoma Metastases. *Cancer Res* (2012) 72:3997–4007. doi: 10.1158/0008-5472.Can-12-1377
  27. Miho E, Roškar R, Greiff V, Reddy ST. Large-Scale Network Analysis Reveals the Sequence Space Architecture of Antibody Repertoires. *Nat Commun* (2019) 10:1321. doi: 10.1038/s41467-019-09278-8
  28. Bashford-Rogers RJM, Bergamaschi L, McKinney EF, Pombal DC, Mescia F, Lee JC, et al. Analysis of the B Cell Receptor Repertoire in Six Immune-Mediated Diseases. *Nature* (2019) 574:122–6. doi: 10.1038/s41586-019-1595-3
  29. McDaniel JR, Pero SC, Voss WN, Shukla GS, Sun Y, Schaeztle S, et al. Identification of Tumor-Reactive B Cells and Systemic IgG in Breast Cancer Based on Clonal Frequency in the Sentinel Lymph Node. *Cancer Immunol Immunother* (2018) 67:729–38. doi: 10.1007/s00262-018-2123-2
  30. Kotnis A, Du L, Liu C, Popov SW, Pan-Hammarström Q. Non-Homologous End Joining in Class Switch Recombination: The Beginning of the End. *Philos Trans R Soc Lond B Biol Sci* (2009) 364:653–65. doi: 10.1098/rstb.2008.0196
  31. Miqueu P, Guillet M, Degauque N, Doré JC, Soullillou JP, Brouard S. Statistical Analysis of CDR3 Length Distributions for the Assessment of T and B Cell Repertoire Biases. *Mol Immunol* (2007) 44:1057–64. doi: 10.1016/j.molimm.2006.06.026
  32. Xu JL, Davis MM. Diversity in the CDR3 Region of V(H) Is Sufficient for Most Antibody Specificities. *Immunity* (2000) 13:37–45. doi: 10.1016/s1074-7613(00)00006-6
  33. Akbar R, Robert PA, Pavlović M, Jeliazkov JR, Snapkov I, Slabodkin A, et al. A Compact Vocabulary of Paratope-Epitope Interactions Enables Predictability of Antibody-Antigen Binding. *Cell Rep* (2021) 34:108856. doi: 10.1016/j.celrep.2021.108856
  34. Prince HE, Yeh C, Lapé-Nixon M. Utility of IgM/IgG Ratio and IgG Avidity for Distinguishing Primary and Secondary Dengue Virus Infections Using Sera Collected More Than 30 Days After Disease Onset. *Clin Vaccine Immunol* (2011) 18:1951–6. doi: 10.1128/cvi.05278-11
  35. Cucunawangsih, Lugito NP, Kurniawan A, Immunoglobulin G (IgG) to IgM Ratio in Secondary Adult Dengue Infection Using Samples From Early Days of Symptoms Onset. *BMC Infect Dis* (2015) 15:276. doi: 10.1186/s12879-015-1022-9
  36. Brochet X, Lefranc MP, Giudicelli V. IMGT/V-QUEST: The Highly Customized and Integrated System for IG and TR Standardized V-J and V-D-J Sequence Analysis. *Nucleic Acids Res* (2008) 36:W503–508. doi: 10.1093/nar/gkn316
  37. Bolotin DA, Poslavsky S, Mitrophanov I, Shugay M, Mamedov IZ, Putintseva EV, et al. MiXCR: Software for Comprehensive Adaptive Immunity Profiling. *Nat Methods* (2015) 12:380–1. doi: 10.1038/nmeth.3364
  38. Greiff V, Menzel U, Haessler U, Cook SC, Friedensohn S, Khan TA, et al. Quantitative Assessment of the Robustness of Next-Generation Sequencing of Antibody Variable Gene Repertoires From Immunized Mice. *BMC Immunol* (2014) 15:40. doi: 10.1186/s12865-014-0040-5
  39. Avram O, Vaisman-Mentesh A, Yehezkel D, Ashkenazy H, Pupko T, Wine Y. ASAP - A Webserver for Immunoglobulin-Sequencing Analysis Pipeline. *Front Immunol* (2018) 9:1686. doi: 10.3389/fimmu.2018.01686
  40. Willis AD. Rarefaction, Alpha Diversity, and Statistics. *Front Microbiol* (2019) 10:2407. doi: 10.3389/fmicb.2019.02407
  41. Hill MO. Diversity and Evenness: A Unifying Notation and Its Consequences. *Ecology* (1973) 54:427–32. doi: 10.2307/1934352
  42. Chaudhary N, Wesemann DR. Analyzing Immunoglobulin Repertoires. *Front Immunol* (2018) 9:462. doi: 10.3389/fimmu.2018.00462
  43. Kirkham CM, Scott JNF, Wang X, Smith AL, Kupinski AP, Ford AM, et al. Cut-And-Run: A Distinct Mechanism by Which V(D)J Recombination Causes Genome Instability. *Mol Cell* (2019) 74:584–97.e589. doi: 10.1016/j.molcel.2019.02.025
  44. Greiff V, Menzel U, Miho E, Weber C, Riedel R, Cook S, et al. Systems Analysis Reveals High Genetic and Antigen-Driven Predetermination of Antibody Repertoires Throughout B Cell Development. *Cell Rep* (2017) 19:1467–78. doi: 10.1016/j.celrep.2017.04.054
  45. Snapper CM, Mond JJ. Towards a Comprehensive View of Immunoglobulin Class Switching. *Immunol Today* (1993) 14:15–7. doi: 10.1016/0167-5699(93)90318-f
  46. Mosmann TR, Coffman RL. TH1 and TH2 Cells: Different Patterns of Lymphokine Secretion Lead to Different Functional Properties. *Annu Rev Immunol* (1989) 7:145–73. doi: 10.1146/annurev.iy.07.040189.001045
  47. Mohammed ZM, Going JJ, Edwards J, Elsberger B, McMillan DC. The Relationship Between Lymphocyte Subsets and Clinico-Pathological Determinants of Survival in Patients With Primary Operable Invasive Ductal Breast Cancer. *Br J Cancer* (2013) 109:1676–84. doi: 10.1038/bjc.2013.493
  48. Charoentong P, Finotello F, Angelova M, Mayer C, Efreanova M, Rieder D, et al. Pan-Cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep* (2017) 18:248–62. doi: 10.1016/j.celrep.2016.12.019

49. Iglesia MD, Vincent BG, Parker JS, Hoadley KA, Carey LA, Perou CM, et al. Prognostic B-Cell Signatures Using mRNA-Seq in Patients With Subtype-Specific Breast and Ovarian Cancer. *Clin Cancer Res* (2014) 20:3818–29. doi: 10.1158/1078-0432.Ccr-13-3368
50. Nzula S, Going JJ, Stott DI. Antigen-Driven Clonal Proliferation, Somatic Hypermutation, and Selection of B Lymphocytes Infiltrating Human Ductal Breast Carcinomas. *Cancer Res* (2003) 63:3275–80.
51. Coronella JA, Spier C, Welch M, Trevor KT, Stopeck AT, Villar H, et al. Antigen-Driven Oligoclonal Expansion of Tumor-Infiltrating B Cells in Infiltrating Ductal Carcinoma of the Breast. *J Immunol* (2002) 169:1829–36. doi: 10.4049/jimmunol.169.4.1829
52. Kroeger DR, Milne K, Nelson BH. Tumor-Infiltrating Plasma Cells Are Associated With Tertiary Lymphoid Structures, Cytolytic T-Cell Responses, and Superior Prognosis in Ovarian Cancer. *Clin Cancer Res* (2016) 22:3005–15. doi: 10.1158/1078-0432.Ccr-15-2762
53. Bolotin DA, Poslavsky S, Davydov AN, Frenkel FE, Fanchi L, Zolotareva OI, et al. Antigen Receptor Repertoire Profiling from RNA-Seq Data. *Nat Biotechnol* (2017) 35:908–11. doi: 10.1038/nbt.3979
54. Brynjolfsson SF, Persson Berg L, Olsen Ekerhult T, Rimkute I, Wick MJ, Mårtensson IL, et al. Long-Lived Plasma Cells in Mice and Men. *Front Immunol* (2018) 9:2673. doi: 10.3389/fimmu.2018.02673
55. Tokoyoda K, Zehentmeier S, Chang HD, Radbruch A. Organization and Maintenance of Immunological Memory by Stroma Niches. *Eur J Immunol* (2009) 39:2095–9. doi: 10.1002/eji.200939500
56. Perez-Andres M, Paiva B, Nieto WG, Caraux A, Schmitz A, Almeida J, et al. Human Peripheral Blood B-Cell Compartments: A Crossroad in B-Cell Traffic. *Cytometry B Clin Cytom* (2010) 78(Suppl 1):S47–60. doi: 10.1002/cyto.b.20547
57. Honjo T, Kinoshita K, Muramatsu M. Molecular Mechanism of Class Switch Recombination: Linkage With Somatic Hypermutation. *Annu Rev Immunol* (2002) 20:165–96. doi: 10.1146/annurev.immunol.20.090501.112049
58. Roco JA, Mesin L, Binder SC, Nefzger C, Gonzalez-Figueroa P, Canete PF, et al. Class-Switch Recombination Occurs Infrequently in Germinal Centers. *Immunity* (2019) 51:337–50.e337. doi: 10.1016/j.immuni.2019.07.001
59. Zhang Y, Meyer-Hermann M, George LA, Figge MT, Khan M, Goodall M, et al. Germinal Center B Cells Govern Their Own Fate via Antibody Feedback. *J Exp Med* (2013) 210:457–64. doi: 10.1084/jem.20120150
60. Al-Lazikani B, Lesk AM, Chothia C. Standard Conformations for the Canonical Structures of Immunoglobulins. *J Mol Biol* (1997) 273:927–48. doi: 10.1006/jmbi.1997.1354
61. Neuberger MS, Rajewsky K. Activation of Mouse Complement by Monoclonal Mouse Antibodies. *Eur J Immunol* (1981) 11:1012–6. doi: 10.1002/eji.1830111212
62. Nimmerjahn F, Ravetch JV. Divergent Immunoglobulin G Subclass Activity Through Selective Fc Receptor Binding. *Science* (2005) 310:1510–2. doi: 10.1126/science.1118948
63. Chao A, Chiu C-H, Hsieh TC, Davis T, Nipperess DA, Faith DP. Rarefaction and Extrapolation of Phylogenetic Diversity. *Methods Ecol Evol* (2015) 6:380–8. doi: 10.1111/2041-210X.12247(2015)
64. Lüdecke D. Sjmisc: Data and Variable Transformation Functions. *J Open Source Software* (2018) 3:754. doi: 10.21105/joss.00754
65. van der Loo M. The Stringdist Package for Approximate String Matching. *R J* (2014) 6. doi: 10.32614/RJ-2014-011
66. Wickham H. The Split-Apply-Combine Strategy for Data Analysis. *J Stat Software* (2011) 40:29. doi: 10.18637/jss.v040.i01
67. Dplyr: A Grammar of Data Manipulation V. In: *R Package, Version 0.4.3* (2015).
68. *Strings: Simple, Consistent Wrappers for Common String Operations* (2019).
69. Efficient Manipulation of Biological Strings V. In: *R Package, version 2.54.0* (2019).
70. *Devtools: Tools to Make Developing R Packages Easier* (2021).
71. Eddelbuettel D, Francois R. Rcpp: Seamless R and C++ Integration. *2011* (2011) 40:18. doi: 10.18637/jss.v040.i08
72. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the Tidyverse. *J Open Source Software* (2019) 4:1686. doi: 10.21105/joss.01686
73. Gu Z, Gu L, Eils R, Schlesner M, Brors B. Circlize Implements and Enhances Circular Visualization in R. *Bioinformatics* (2014) 30:2811–2. doi: 10.1093/bioinformatics/btu393

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Aizik, Dror, Taussig, Barzel, Carmi and Wine. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.