RESEARCH ARTICLE

Development of the expressed immunoglobulin μ chain repertoire during maturation of mice B cells

Jingwen LIANG [1*], Yingfeng LUO [2*], Yi SUN [3*], Meng LEI [2], Bing ZHANG [2], Songnian HU (✉) [2],

Yaofeng ZHAO (✉) [1]

1 State Key Laboratory of Agrobiotechnology, College of Biological Sciences, China Agricultural University, Beijing 100094, China
2 CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China
3 Laboratory of Animal Molecular Genetics, College of Animal Science and Veterinary Medicine, Shandong Agricultural University,
Taian 271018, China

Abstract In the bone marrow and spleen, the developing B cell populations undergo both negative and positive selections to shape their B cell receptor repertoire. To gain insight into the shift of the immunoglobulin heavy (IgH) chain repertoire during B cell development, we undertook large scale Ig μ chain repertoire analysis of pre-B, immature B and spleen B cell populations. We found that the majority of V_H gene segments, V_H families, J_H and D gene segments, were observed to have significantly different usage frequencies when three B cell populations were compared, but the usage profile of the V_H, D, and J_H genes between different B cell populations showed high correlations. In both productive and nonproductive rearrangements, the length of CDRH3 shortened significantly on average when B cells entered the periphery. However, the CDRH3 length distribution of nonproductive rearrangements did not follow a Gaussian distribution, but decreased successively in the order $3n-2$, $3n-1$ and $3n$, suggesting a direct correlation between mRNA stability and CDRH3 length patterns of nonproductive rearrangements. Further analysis of the individual components comprising CDRH3 of productive rearrangements indicated that the decrease in CDRH3 length was largely due to the reduction of N addition at the 5′ and 3′ junctions. Moreover, with development, the amino acid content of CDRH3 progressed toward fewer positively charged and nonpolar residues but more polar residues. All these data indicated that the expressed Ig μ chain repertoire, especially the repertoire of CDRH3, was fine-tuned when B cells passed through several checkpoints of selection during the process of maturation.

*These authors contribute equally to the work

1 Introduction

In the humoral immune system, a diverse antibody repertoire is essential to produce effective and specific immune responses. During the development of B cells, the antibody (or B cell receptor, BCR) repertoire is generated by somatic joining of the variable (V), diversity (D), and joining (J) gene segments at immunoglobulin heavy (IgH) chain and light (IgL) chain loci [1]. Diversity of each chain is determined by various V_H-D-J_H or $V_{\kappa/\lambda}$-$J_{\kappa/\lambda}$ combinations and unpredictable junctional sequences, which are created *de novo* by the imprecise joining of gene segments and the varied insertion of non-templated (N) and palindromic (P) nucleotides. The most diverse V_H-D-J_H junction encodes the third complementarity-determining region of the IgH chain variable region (CDRH3), in which both length and amino acid composition play a vital role in defining the BCR specificity [2].

In theory, the potential BCR repertoire has been estimated to surpass 10^8 potential sequence variants. However, the majority of the initially generated IgH and/or IgL chains are successively removed from the whole repertoire, when the developing B cells pass through several checkpoints depending on positive and/or negative selections [3,4]. During the first checkpoint, pre-B cells are positively selected for the expression of functional μ heavy chains (μHCs) encoded by the productively rearranged IgH alleles [5]. The μHC has the ability to pair with the surrogate light chain to form the pre-BCR, which is transported to the cell surface and induces the proliferation of the pre-B cells [6,7]. A few studies have indicated that

autoreactivity of the pre-BCR is crucially important for inducing the expansion of pre-B cells with a productive IgH chain rearrangement [8,9]. The second checkpoint is usually referred to as central B cell tolerance, a process that negatively selects the autoreactive immature B cells in the bone marrow environment with high efficiency and stringency [10]. Receptor editing, anergy, and deletion are three known mechanisms utilized to establish the tolerance [11–15]. In wild-type mice, receptor editing at the IgL chain loci play a major role in silencing of autoreactive B cells (~50%), and a small proportion of low avidity autoreactive B cells become anergic, whereas the very few remaining undergo deletion [16,17]. When the newly formed immature B cells migrate to peripheral lymphoid tissues as transitional B cells, they continue to experience multiple checkpoints before finally develop entering the mature B cell pool [18]. In contrast to the detailed investigations of the central B cell tolerance, relatively little is known about the mechanisms underlying the peripheral B cell tolerance. However, several reports suggest that the transitional B cells probably experience both negative and positive selection [19–23].

The effect of positive and/or negative selection could be reflected in differences between BCR repertoires of different B cell populations. Several studies by Schroeder Jr et al. elucidated detailed features of the CDRH3 repertoire in various murine B cell populations, in terms of length, amino acid composition, and average hydrophobicity. Among them, the key feature of CDRH3 repertoire development in both BALB/c and C57BL/6 mice is an increase in average CDRH3 length with B cell maturation [24–26]. However, use of all the above features as accurate reflections of the entire BCR repertoire has been challenged, as they were deduced from the analysis of a limited number of μ chain transcripts containing only V_H7183 family members.

In the last few years, high-throughput sequencing technologies have been widely utilized to describe the repertoire of antibody and T cell receptor (TCR) in humans and zebrafish [27–33]. Here, to comprehensively understand the effects of selections on BCR repertoire during the B cell development, we describe and compare the IgH repertoire of pre-, immature and spleen B cells in C57BL/6 mice using 454 high-throughput pyrosequencing with particular attention to the usage profiles of V_H, D, and J_H genes and the repertoire of CDRH3. From the viewpoint of comparative studies, these results will provide the basis for further investigation of the BCR development in domestic animals.

2   Materials and methods

2.1   Mice

C57BL/6 mice of clean grade were purchased from Vital River Laboratories (Beijing, China) and housed and cared for in a barrier environment according to the national standard (GB14925-2001) issued by General Administration of Quality Supervision, Inspection and Quarantine of China.

2.2   B cell isolation and cell sorting

The pooled bone marrows and spleens were collected from eight 6–8 weeks old mice (four male and four female). Single-cell suspensions were prepared by passing tissue samples through 200 μm nylon mesh and resuspending the filtered cells in phosphate buffered saline (PBS). Erythrocytes were lysed in red-blood-cell (RBC) lysis buffer (eBioscience, San Diego, CA, USA). The mixture of pre-B and immature B (IM-B) cells were isolated from bone marrow single-cell suspension by depletion of CD43$^+$ pro-B cells and non-B cells, using MACS B cell Isolation Kit (Miltenyi Biotec, Bergisch Gladbach, Germany). The IM-B cells were than positively selected from the mixture with the Anti-Mouse IgM MicroBeads (Miltenyi Biotec), and the pre-B cells were the negative fraction. Spleen total B (S-B) cells were positively selected from a spleen single-cell suspension using CD45R(B220) MicroBeads (Miltenyi Biotec). Cells (~10^5) of each B cell subset were stained with PE-Cy5.5-anti-mouse B220, FITC-anti-mouse IgM, and PE-anti-mouse CD43, and the purity of each B cell subset was analyzed on a MoFLo High-performance cell sorter (DakoCytomation, Fort Collins, CO, USA).

2.3   mRNA preparation, cDNA synthesis, and PCR

Total RNA samples were isolated from each B cell subset using *mir*Vana™ miRNA Isolation Kit (Ambion, Austin, TX, USA), and reverse transcription was conducted using mouse C_μ-specific primer (RTC$_\mu$) [34] and M-MLV Reverse Transcriptase (Promega, Madison, WI, USA) following the user's instructions. Recombined IgH VDJ regions were amplified from each cDNA sample using a multiplex of 17 mouse V_H family-specific upstream primers and a mouse C_μ-specific downstream primer ($C_\mu0$) (a 454-adaptor sequence was added at the 5′ end of each primer) [34]. A touchdown PCR was performed with 25 μL 2 × Phusion HF PCR Master Mix (NEB, Beverley, MA, USA) and 10 μL cDNA per sample, using the following protocol: 95°C for 2 min, then touchdown PCR for 6 cycles (94°C for 1 min, from 67°C to 57°C decreasing 2°C per cycle for 1 min and 72°C for 1 min), followed by 14 cycles of 94°C for 1 min, 57°C for 1 min, and 72°C for 1 min, and a final extension step of 68°C for 10 min. To minimize PCR amplification bias, for each B cell subset sample, three C_μ-specific downstream primers with three different barcoding index sequences ($C_\mu1$–$C_\mu3$) were separately utilized to repeat the preceding PCR. All primer sequences are given in Appendix-Table S1.

2.4 Amplicon purification and sequencing

PCR amplicons (350–450 bp) were purified by 1.2% agarose gel electrophoresis and MinElute Gel Extraction Kit (Qiagen, Hilden, Germany). For each B cell subset sample, 4 separate amplicons using $C_\mu 0$–$C_\mu 3$ were mixed in equal amount before sequencing on the 454 GS FLX sequencer (Roche, Basel, Switzerland). For sequencing, the DNA libraries were treated according to the operation manual.

2.5 Sequence analysis

A local BLAST database was constructed using the C57BL/6 reference germline V_H, D, and J_H sequences [35–37], and the BLAST algorithm was prepared according to the NCBI IgBLAST (http://www.ncbi.nlm.nih.gov/igblast/). All reads were aligned by the local BLAST to each germline V_H and J_H gene, and the highest scored V_H and J_H genes were used for the following analysis. The algorithm to determine the V_H-D-J_H junctions and to identify the 3′ V-regions, 5′ J-regions, D-regions, P-regions, and N-regions were designed on the basis of the IMGT/JunctionAnalysis [38]. Reads were considered to be informative antibody sequences if they passed the following quality control criteria: a minimum length of 300 bp; identified V_H and J_H genes; and absence of ambiguous nucleotides in the junction region. A non-redundant rearrangements library was constructed when the clonally related sequences were reduced. From these unique sequences, the $V_H DJ_H$ combination repertoire, antibody repertoire and CDRH3 characteristics were calculated. The definitions of CDRH3 and CDRH3-loop were described in the reference [24,39].

2.6 Statistical analysis

Data were combined and subsequently analyzed in Excel (Microsoft, USA). Differences between B cell populations were assessed where appropriate by two-tailed Student's t-test and χ^2 test. Correlations were performed either by weighting the read-abundance for a particular V_H, V_H family, D, J_H, and $V_H DJ_H$ class or by weighting the total number for each amino acid in the CDRH3 loop.

3 Results

3.1 Deep sequencing of the mouse IgH repertoire in different B cell populations

Pre-B (B220$^+$, CD43$^-$, and IgM$^-$), IM-B (B220$^+$, CD43$^-$, and IgM$^+$) and S-B (B220$^+$) cells from eight C57BL/6 mice were isolated using MACS. FACS analysis of the purified cells demonstrated that more than 90% of cells in each population were consistent with the correspondent phenotypes (Appendix-Fig. S1). A total of 397654, 333360, and 298987 reads identified by barcoding index were obtained from pre-B, I-B, and S-B cells, respectively. After selection according to the quality control criteria, 256532 (64.5%) pre-B cell sequences, 174593 (52.4%) I-B cell sequences, and 145764 (48.8%) S-B cell sequences with a complete and identifiable set of $V_H DJ_H$ gene assignments were utilized for the following analysis.

The number of functional V_H (110), D (10, including two identical DSP2.x segments), and J_H (4) gene segments of the C57BL/6 mice could theoretically produce 3960 (110 × 9 × 4) possible $V_H DJ_H$ combinations. Saturation analysis of the $V_H DJ_H$ combination diversity demonstrated that the $V_H DJ_H$ repertoire was close to saturation (nearly 3000 $V_H DJ_H$ combinations) when the reads increased to as many as 150000 in all three cell populations (Appendix-Fig. S2a), indicating that the data covered almost all actual $V_H DJ_H$ combinations in the C57BL/6 mice repertoire.

In all three B cell populations, the majority of the rearranged IgH sequences were productive (pre-B 219469, 85.6%; IM-B 146227, 83.8%; S-B 117643, 80.7%). Productive rearranged IgH sequences with the same V_H, D, and J_H segments usage and identical CDRH3 nucleotide sequences were identified as clonally related sequences, which probably arose from PCR amplification of a single sequence or multiple cDNA copies from clonal expanded cells. So, clonally related sequences actually encoded almost identical IgH variable regions. The frequency of clonally related sequences gradually decreased with increasing clone size, and unique sequences with no clonally related sequences (clone size is 1) were predominate in all three B cell populations (Appendix-Fig. S3). However, compared to the IM-B and S-B cell populations, the pre-B cell population had significantly smaller clone size, when the reads approached 30000 (Appendix-Fig. S3). As the reads increased, clone size showed a similar trend in all three B cell populations (data not shown). To avoid repeatedly calculating the antibody repertoire, only one sequence was chosen at random from each group of clonally related sequences to represent a unique (unredundant) antibody molecule. More than 84% (99165) of the productive IgH sequences from S-B cell populations were unique IgH rearrangements, compared to 77.9% (170933) in pre-B cells and 75.0% (109741) in IM-B cells. The modest increase of the unique rearrangements percentage in S-B cells is due to the antigen-encountered B cells, which underwent somatic hypermutation (SHM) in CDRH3. Saturation analysis of the antibody diversity with the number of unique rearrangements demonstrated that the antibody repertoire was far from saturation in all three B cell populations (Appendix-Fig. S2b). This indicated that during the VDJ recombination nucleotide addition and deletion in V_H-D and D-J_H junctions greatly enriched the antibody repertoire.

3.2 The usage profile of the V_H, D, and J_H genes was not altered with B cell development

The V_H gene segment usage was highly uneven among the 110 functional V_H genes, with a few V_H segments having greater opportunities to participate in the V_HDJ_H recombination than the majority (Appendix-Fig. 1a). In each B cell population, the top 20 frequently utilized V_H genes occupied more than 60% of the total antibody repertoire (pre-B 63.7%, IM-B 67.6%, and S-B 68.0%). There was a strong overlap (16/20) between the top 20 V_H genes in three B cell populations. Of the overlapped V_H genes, nine were from V_HJ558, two from V_HQ52, two from V_H7183, and the remaining three from V_H36-60, V_HS107, and V_HSM7, respectively. In addition, 10 functional V_H segments ($V_HJ558.9$, $V_HJ558.64$, $V_HQ52.11$, $V_HJ606.2$, $V_HJ606.3$, $V_H3609.1$, $V_H3609.3$, $V_H3609.11$, $V_HVGAM3.8-4$, and $V_H16.1$) were not found to participate in recombination in any of the three B cell populations. It is noteworthy that some pseudo-V_H genes participated in the rearrangement, although the proportion was extremely low (data not shown). In nearly half of the expressed V_H genes (45/100), χ^2 analysis showed a significantly different proportion among three B cell populations ($P < 10^{-5}$, Appendix-Table S2).

The usage proportion of the V_H family J558 was overdominant and proportionally correlated with its germline complexity (52/110), but in fact only nine V_HJ558 genes ($V_HJ558.50$, $V_HJ558.52$, $V_HJ558.53$, $V_HJ558.55$, $V_HJ558.59$, $V_HJ558.67$, $V_HJ558.72$, $V_HJ558.75$, and $V_HJ558.77$) contributed to about 80% of the V_HJ558 repertoire (Fig. 1a and Fig. 1b). Two D-proximal V_H families, V_H7183 and V_HQ52, also showed a usage proportion compatible with their family size (10/110 for V_H7183 and 9/110 for V_HQ52) (Fig. 1b). Both families were found to be overrepresented in adult bone marrow relative to adult spleen, but showed different changing trends in our analysis. In agreement with the previous studies, V_H7183 was significantly less frequently used between pre-B and S-B cells ($P < 10^{-33}$), but for V_HQ52 the situation was just the reverse (pre-B vs. S-B, $P < 10^{-14}$) (Fig. 1b). In contrast to the above three V_H families, the contributions of V_H36-60 and V_H3609 were notably in inverse proportion to their family size; the V_H36-60 had a complexity similar to V_H3609 (6/110 vs. 8/110), yet it was the second most highly expressed family in IM-B and S-B cells, and ten times more frequent than V_H3609 (average 11.6 vs. 1.05%) (Fig. 1b). Notably, the most frequently utilized V_H gene in IM-B and S-B cells was $V_H36-60.6$, which constituted more than 60% of the V_H36-60 repertoire (Fig. 1a).

In all three B cell populations, the J_H genes proportional usage from high to low was J_H2, J_H4, J_H3 and J_H1. Compared with the pre-B cells, the J_H2 usage continuously increased with B cell development (pre-B vs. IM-B, $P < 10^{-6}$; and IM-B vs. S-B, $P < 10^{-14}$), but the J_H3 usage

was significantly decreased (pre-B vs. IM-B, $P < 10^{-11}$, and IM-B vs. S-B, $P < 10^{-11}$) (Fig. 1c). Since the D genes could be greatly shortened by nucleotide deletion at both ends, D-regions could not be identified in about 8% of unique sequences from three B cell populations (Fig. 1d). In the remaining unique sequences, the longest D gene, DFL16.1, was most frequently present in all three B cell populations ($> 30\%$), although its expression was dramatically decreased with B cell development (pre-B vs. IM-B, $P < 10^{-30}$, and IM-B vs. S-B, $P < 10^{-28}$) (Fig. 1d). In terms of the D_H families, the usage proportion was DSP2 > DFL16 > DQ52 > DST4 in all three B cell populations.

Although the majority of V_H genes, V_H families (11/16), J_H (4/4) and D (6/9) genes were observed to have significantly different usage frequencies ($P < 10^{-5}$) between the three B cell populations, the usage profile of the V_H, D, and J_H genes showed a high correlation (correlation coefficient $r > 0.94$ in all pairs) (Fig. 1 and Fig. 2). The V_H- and J_H- profiles of IM-B and S-B cells were more similar than those of pre-B and IM-B cells, and the D-profiles were almost the same among three B cell populations (Fig. 2).

3.3 V_HDJ_H recombination profile was not altered during B cell development

The number of V_HDJ_H combination patterns detected in pre-B, IM-B, and S-B cells was 2999, 2848, and 2791, respectively, which covered more than 70% of the possible V_HDJ_H combinations. A total of 2576 V_HDJ_H combination patterns were shared among all three B cell populations, however, the number expressed specifically in a single B cell population, was only 146, 59, and 59 in pre-B, IM-B, and S-B cells, respectively (Fig. 3a). From the 3D representations of the V_HDJ_H repertoires, it is evident that the overall profile of the V_HDJ_H combinations in the three B cell populations exhibited distinct similarity: the vast majority of the V_HDJ_H repertoires were occupied by only a few V_HDJ_H combinations, which were almost identical among the B cell populations (Fig. 4). The similarity could also be calculated by the high degree of correlation ($r > 0.92$ in all pairs), and the combination profiles of the IM-B and S-B cells, which showed more commonalities (Fig. 2).

The nine combination patterns: $V_HJ558.53/DFL16.1/J_H2$, $V_HJ558.55/DFL16.1/J_H2$, $V_HJ558.67/DFL16.1/J_H2$, $V_HJ558.67/DFL16.1/J_H1$, $V_HJ558.53/DFL16.1/J_H2$, $V_H36-60.6/DFL16.1/J_H2$, $V_H36-60.6/DFL16.1/J_H1$, $V_H36-60.6/DFL16.1/J_H4$ and $V_HJ558.75/DFL16.1/J_H4$, each contributed more than 0.5% of the antibody repertoire in all three B cell populations. Of these, the most diverse recombination was $V_HJ558.55/DFL16.1/J_H2$ (1.06%) identified in pre-B cells, with as many as 1662 unique CDRH3 sequences filtered from 2167 sequences. In contrast to the high similarity of the combination profiles were the dramatic differences between antibody repertoires
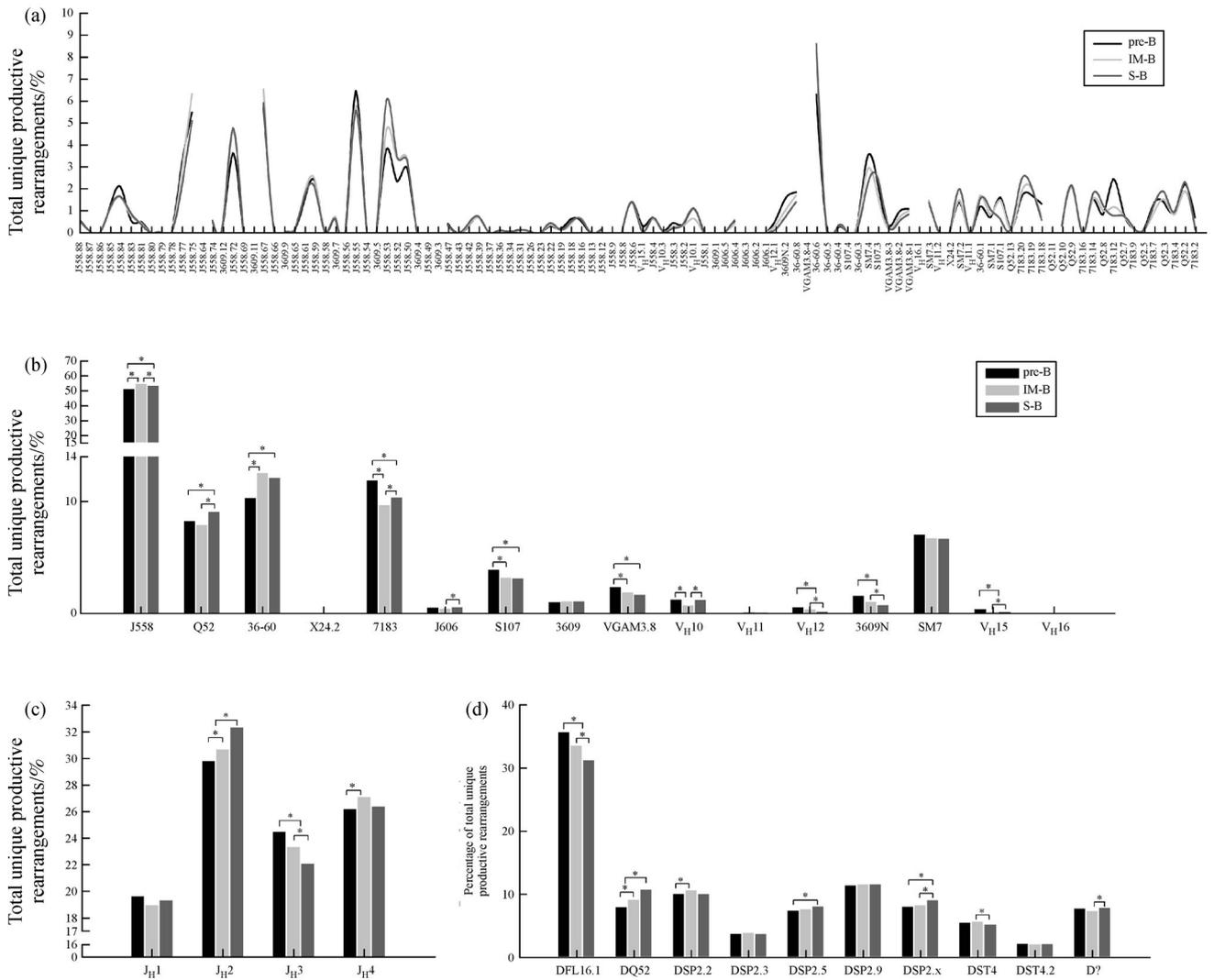
Fig. 1 V_H, V_H family, J_H and D usage of the unique productive rearrangements transcribed from sorted pre-B ($n = 170933$), IM-B ($n = 109741$) and S-B ($n = 99165$) cells. Usage is reported as the percentage of the corresponding segment-containing transcripts. An asterisk indicates significant differences ($P < 10^{-5}$) compared with two corresponding B cell populations. (a) Shows V_H gene segment usage; (b) V_H family usage; (c) J_H gene segment usage; (d) D gene segment usage.

of the three B cell populations. The sequenced antibody repertoires were not saturated, however, and the number of unique antibodies shared among all three B cell populations was only 564, with the vast majority of unique antibodies expressed specifically in a single B cell population (pre-B 158690, IM-B 97769, and S-B 94828) (Fig. 3b).

3.4 CDRH3 length shortened when B cells enter the periphery

To investigate how the size of CDRH3 changed during mouse B cell development, we compared the distribution of CDRH3 lengths in unique productive rearrangements in three B cell populations (Fig. 5a). In general, the average length of CDRH3 showed a small and not significant increment (about 0.13 nt) between pre-B and IM-B cells ($P > 10^{-5}$), but shortened considerably on average (~0.8 nt) during the development from IM-B to S-B cells ($P < 10^{-90}$). When compared to pre-B cells, the variance in CDRH3 length of IM-B cells narrowed. The prevalence of both short ($\leqslant 27$ nt) and long ($\geqslant 45$ nt) CDRH3 decreased, but the prevalence of middle CDRH3 (30–39 nt) increased. When the B cells entered the periphery, the distribution of CDRH3 lengths changed again. During the progression from IM-B to S-B cells, the prevalence of CDRH3 with 33 nt or less increased, but the prevalence of CDRH3 with 39 nt or more decreased.
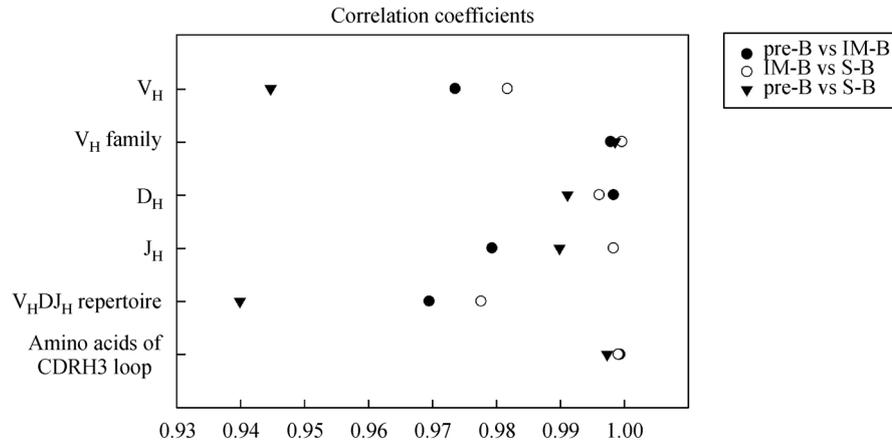
Fig. 2 Usage profile correlation of V_H, V_H family, D, J_H, V_HDJ_H repertoire and amino acids in the CDRH3 loop between pre-B, IM-B and S-B cells
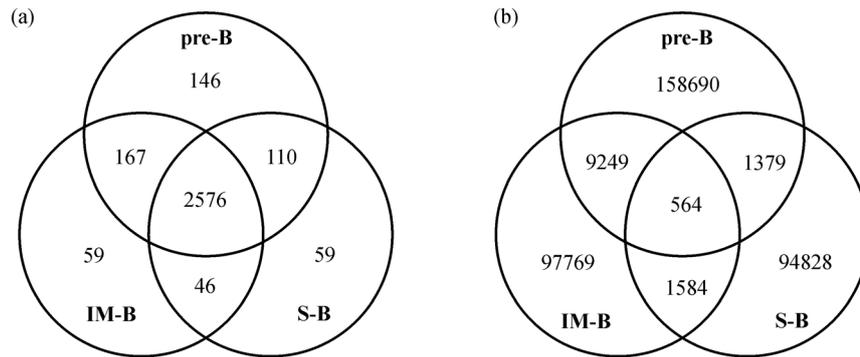
Fig. 3 The number of expressed V_HDJ_H combinations (a) and unique productive rearrangements (b) in pre-B, IM-B, and S-B cells

3.5   Decrease in CDRH3 length reflects deletion in N region and altered usage of shorter D gene segments

Deconstruction of productively rearranged unique CDRH3 sequences that contained identifiable D gene segments (Table 1) allowed further evaluation of the relative contribution of VDJ germline sequence and N (including P) to CDRH3 length. From pre-B to IM-B cells, the average CDRH3 length was unchanged. Minor decreases in the D germline contribution (-0.06 nt) and N addition at the 5′ and 3′ junctions (-0.04 nt and -0.06 nt) were offset by the increase in the J_H germline contribution ($+0.16$ nt). The decrease of N addition at the 3′ junction and the increase of the J_H germline contribution reached a statistically significant level ($P < 10^{-5}$ and $P < 10^{-19}$). From IM-B to S-B cells, the average CDRH3 length significantly decreased by 0.76 nt ($P < 10^{-104}$). The marked decrease in N addition at the 5′ and 3′ junctions (-0.24 nt and -0.27 nt) reflected two-thirds of the total decrease ($P < 10^{-54}$ and $P < 10^{-93}$). Also, the contribution of D germline sequence decreased by 0.11 nt, which was statistically significant ($P < 10^{-7}$). From pre-B to S-B cells,

increased use of DQ52, the shortest D segment, and decreased use of DFL16.1, the longest D segment, was a major factor in the decrease of contribution of D germline sequence (Fig. 1c). However, the V_H and J_H germline sequences did not appear to contribute to the decrease in CDRH3 length during mouse B cell development.

3.6   In unique nonproductive sequences of all three B cell populations, the frequency of three CDRH3 lengths decreased successively in the order $3n-2$, $3n-1$ and $3n$

Because nonproductive sequences are not expressed as functional IgH proteins on the cell surface, they can be used to estimate the CDRH3 length distribution of IgH rearrangements before selection. The number of unique nonproductive rearrangements obtained from pre-B, IM-B and S-B cell populations was 31899, 24979 and 25956, respectively, which is much fewer than the number of productive rearrangements, indicating that the nonproductive rearrangements might be inefficiently transcribed or degraded more rapidly than productive rearrangements. Comparison of the average CDRH3 length between
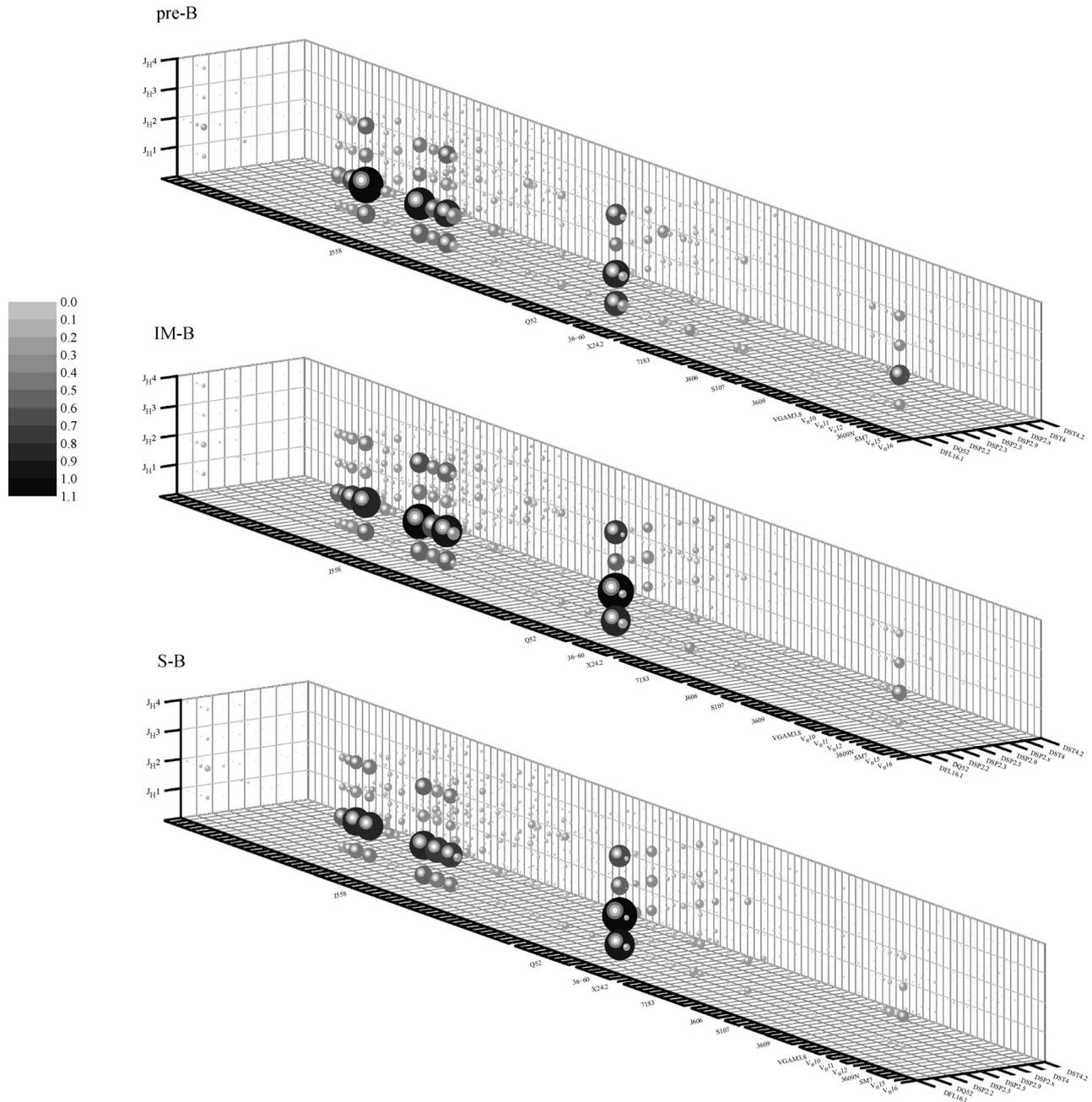
Fig. 4 Three dimensional representations of the entire $V_H DJ_H$ repertoire transcribed from sorted pre-B ($n = 157079$), IM-B ($n = 101258$), and S-B ($n = 90729$) cells. The three axes show all possible V_H, D, and J_H gene segments, and each point in three-space represents a unique $V_H DJ_H$ combination. Both the size and gray scale of each dot refers to the percentage of the unique rearrangements matching a particular combination relative to the total unique productive rearrangements with identifiable V_H, D, and J_H segments.

productive and nonproductive rearrangement in each B cell populations revealed no significant differences (pre-B, 35.44 vs. 35.39, $P = 0.34$; IM-B, 35.57 vs. 35.38, $P = 0.00098$; S-B, 34.77 vs. 34.87, $P = 0.11$). Comparison of the average CDRH3 length of nonproductive rearrangements among three B cell populations revealed no significant difference between pre-B and IM-B cells

(35.39 nt vs. 35.38 nt, $P = 0.88$). However, as with productive rearrangements, nonproductive rearrangements in S-B cells contained shorter CDRH3 sequences than IM-B cells, with mean 34.87 nt vs. 35.38 nt ($P < 10^{-10}$). Unexpectedly, the CDRH3 length distribution of unique nonproductive rearrangements did not follow a Gaussian distribution. In all three B cell populations, the frequency

of CDRH3 length decreased successively in the order $3n-2$, $3n-1$, and $3n$ (Fig. 5b), indicative of a direct correlation between mRNA stability and CDRH3 length patterns of nonproductive rearrangements.

Table 1 Comparison of nucleotide lengths of the CDRH3 components in productively rearranged unique sequences with identifiable D gene segments

Length	pre-B[1]	IM-B[1]	S-B[1]
V_H	5.73 (0.004[3])[a][4]	5.73 (0.005)[a]	5.71 (0.005)[a]
5′ of D[2]	4.57 (0.009)[a]	4.53 (0.011)[a]	4.29 (0.011)[b]
D	11.12 (0.011)[a]	11.06 (0.014)[a]	10.95 (0.014)[b]
3′ of D[2]	3.08 (0.007)[a]	3.02 (0.009)[b]	2.75 (0.009)[c]
J_H	12.56 (0.011)[a]	12.72 (0.013)[b]	12.70 (0.014)[b]
Total CDRH3	36.45 (0.020)[a]	36.45 (0.024)[a]	35.69 (0.025)[b]

Note: [1]The number of sequences analyzed for pre-B, IM-B and S-B cells was 154365, 99629 and 89653, respectively; [2]5′ of D and 3′ of D indicate the N (and including P) nucleotides added between V_H and D and between D and J_H, respectively; [3]Standard errors in parentheses; [4]Different superscript letters (a, b and c) indicate statistically significant difference at $P < 10^{-5}$ within each line.
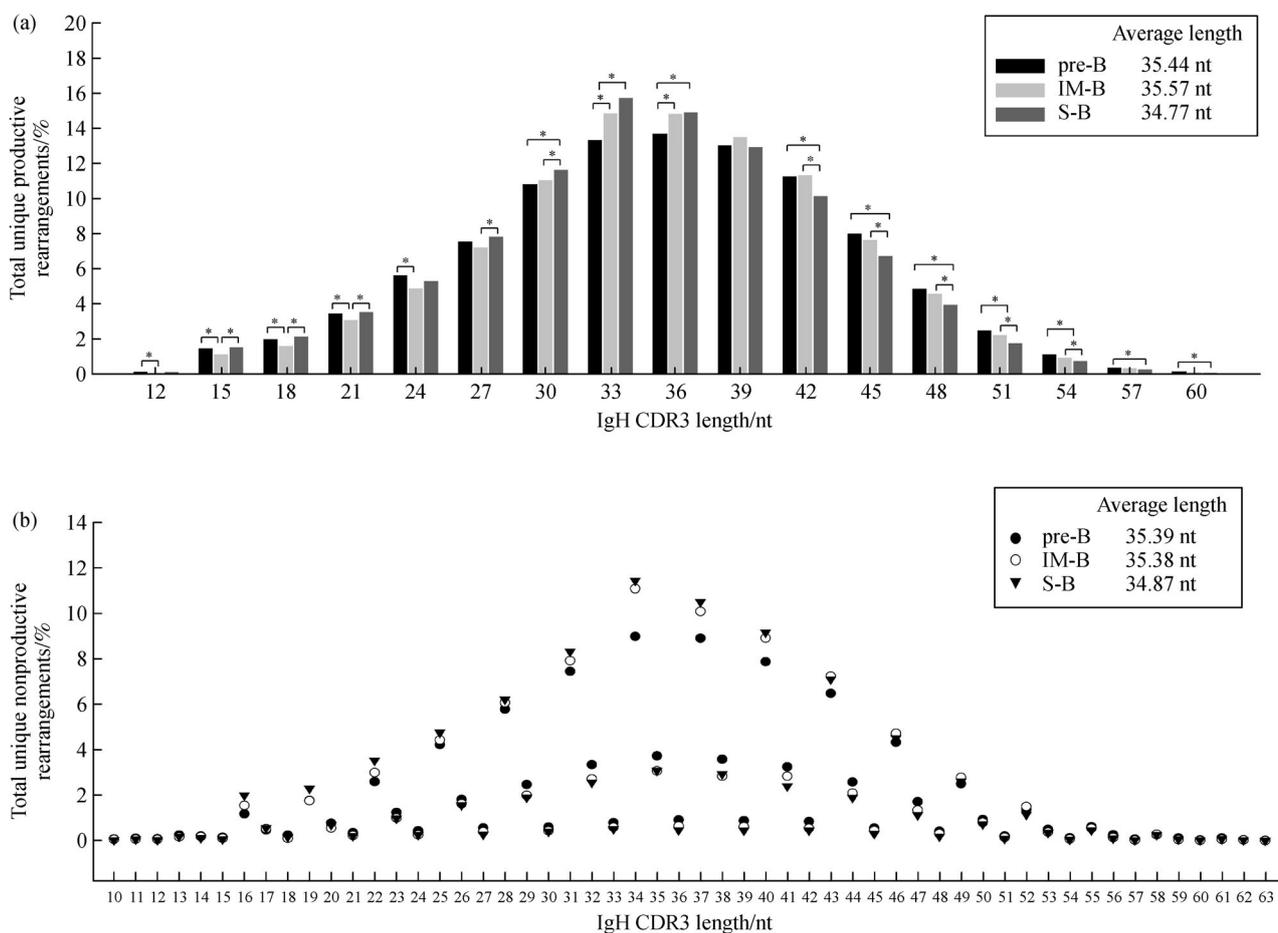

Fig. 5 The CDRH3 length distribution of unique productive (a) and nonproductive (b) rearrangements transcribed from pre-B, IM-B and S-B cells. Each bar in subfigure a and point in subfigure b indicates the percentage of the unique productive (or nonproductive) rearrangements containing the corresponding length of CDRH3 relative to the total unique productive (or nonproductive) rearrangements. In subfigure a, an asterisk indicates significant differences ($P < 10^{-5}$) compared with two corresponding B cell populations.

3.7 Increased use of polar amino acids and decreased use of positively charged and nonpolar amino acids in CDRH3 loops during B cell development

Although the majority of amino acids (except glutamine and methionine) were observed to have significantly different usage proportions ($P < 10^{-6}$) among the three B cell populations, the overall usage profiles of the 20 amino acids showed high correlations (correlation coefficient $r > 0.997$ in all pairs) (Fig. 2). In all three B cell populations, the top five preferred amino acids in CDRH3 loops were tyrosine, glycine, serine, asparagine, and alanine, which contributed to more than 60% of the total amino acids (Fig. 6a). Division of the amino acids into four subsets according to the chemical properties of their side chains and comparison of their proportions in three B cell populations are shown in Fig. 6b. The number of positively charged amino acids was significantly decreased during B cell development (pre-B vs. IM-B, $P < 10^{-48}$; IM-B vs. S-B, $P < 10^{-58}$), and the same was true for nonpolar amino acids (pre-B vs. IM-B, $P < 10^{-43}$; IM-B vs. S-B, $P < 10^{-114}$). Conversely, a significant increase in the proportion of polar amino acids was observed with the development of B cells (pre-B vs. IM-B, $P < 10^{-18}$; IM-B vs. S-B, $P < 10^{-161}$). A greater change in proportion was observed between the repertoires of IM-B and S-B cells compared with those of pre-B and IM-B cells. In addition, the number of negatively charged amino acids increased during the development from pre-B to IM-B cells ($P < 10^{-81}$), but did not change significantly between IM-B and S-B cells ($P = 0.099$).
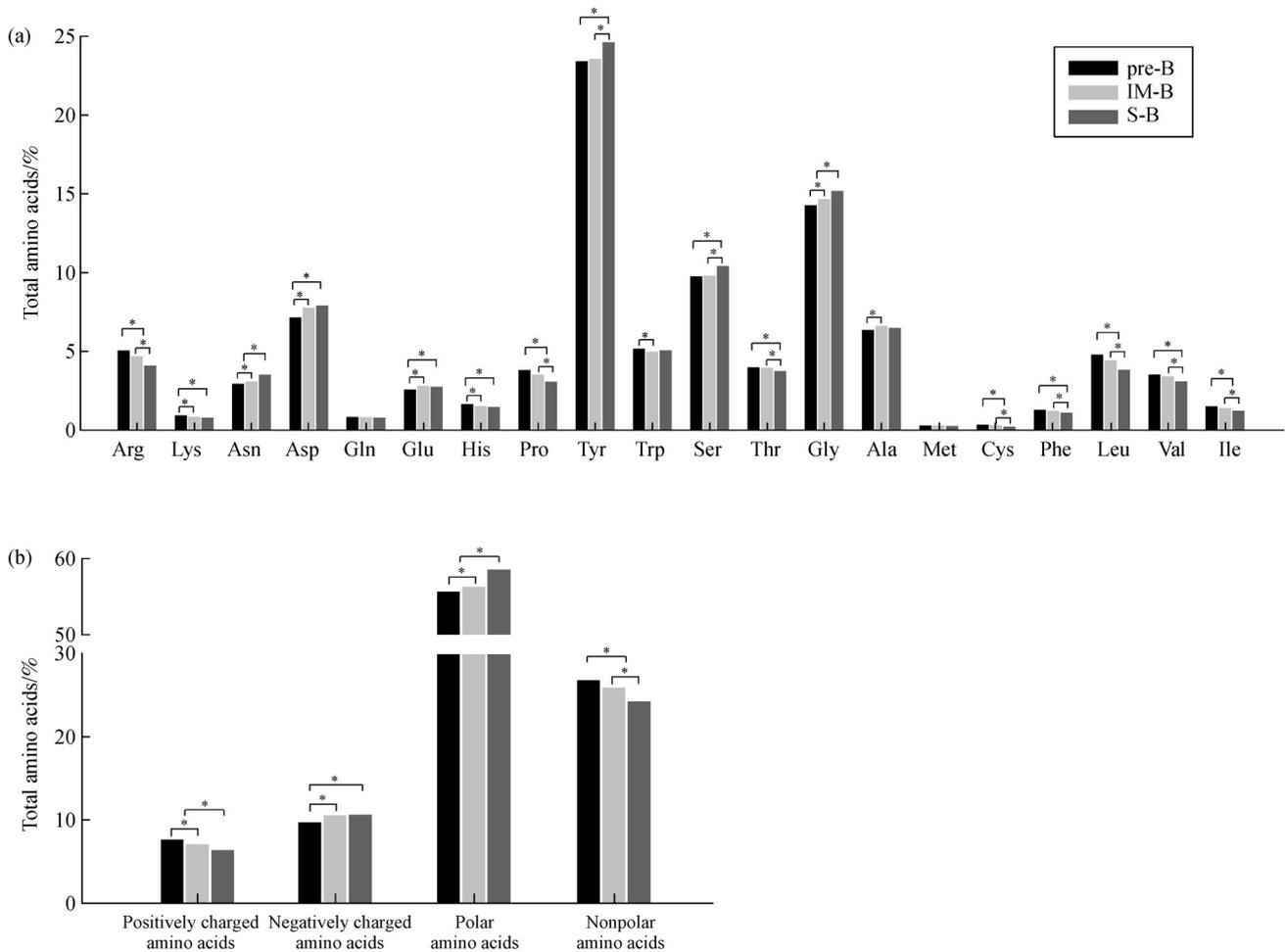
Fig. 6 Amino acid usage in the CDRH3 loop of the unique productive rearrangements transcribed from sorted pre-B, IM-B and S-B cells. Usage of individual amino acids in subfigure a (or usage of four amino acid subsets in subfigure b) is calculated through analysis of 1164987 amino acids from 170933 unique pre-B CDRH3 loops, 752793 amino acids from 109741 unique IM-B CDRH3 loops, and 653640 amino acids from 99165 unique S-B CDRH3 loops and the usage is reported as the percentage of individual amino acids (or each subset of amino acids) relative to the total amino acids. An asterisk indicates significant differences ($P < 10^{-5}$) compared with two corresponding B cell populations.

4   Discussion

A strain-dependent developmental difference in the V_H gene usage was intensively studied in BALB/c and C57BL/6 mice. In BALB/c, the two most D-proximal V_H families, V_H7183 and V_HQ52, were overrepresented in fetal liver B cells and early B cell development in adult bone marrow, compared with their germline complexity, but this preference was lost when the B cells matured and migrated to the periphery [26,40–42]. In contrast, usage bias of the 3′ V_H genes in C57BL/6 strain was clearly observed in fetal liver B cells [42], but in our research was not apparent in the precursor B cells during B cell development in adult bone marrow. The mouse pre-B could be further divided into three subclasses, pre-B-I (c-kit$^+$, CD25$^-$ and CD43$^+$), large pre-B-II (c-kit$^-$, CD25$^+$ and partially CD43$^+$), and small pre-B-II (c-kit$^-$, CD25$^+$ and CD43$^-$), within which the latter is predominant (60%–70%) [43,44]. The research of ten Boekel et al. demonstrated that in normal C57BL/6 mice a V_H repertoire shift happens as cells matured from the pre-B-I to the pre-B-II, since some V_H domains encoded by the V_H7183 and V_HQ52 families are not able to form preBCR with the surrogate light chain or the preBCR formed is incapable of efficiently inducing pre-B cell clonal proliferation [45]. In our present study, the overall expression profile of the V_H families did not show significant changes during B cell maturation in adult C57BL/6 mice. The detection of the unbiased V_H repertoire in our study is probably due to the pre-B cell population we used, which mainly belonged to the pre-B-II subclass (CD43$^-$), which developes from preBCR$^+$ cells and has already passed the positive selection step.

Due to the death of the $V_HDJ_H^-$/$V_HDJ_H^-$ B cells, nonproductive rearrangements could only be detected from the $V_HDJ_H^+$/$V_HDJ_H^-$ B cells. Thus the presence of a given V_H gene in the nonproductive rearrangements cannot contribute positively or negatively to selection. In contrast, the proportion of a given V_H gene in the productive rearrangements could be determined by the efficiency of participation in rearrangement as well as the property to be selected. Therefore, in a previous study, the selection of the V_H repertoire during B cell development was measured by the shift in the proportion of productive to nonproductive rearrangements or the shift in the fractions of in-frame rearrangements (IF fractions) of a given V_H gene (or V_H family) in various B cell populations [46]. Alternatively, whatever the stage of development, a B cell could only express one kind of H chain with unique specificity. Consequently, in our study, the effect of selection, either to a V_H gene or to a V_HDJ_H combination, was determined by the diversity change, which was calculated from the substantial sequence data composed of the expressed productive V_HDJ_H. Although most of the pre-B cells utilized here are pre-B-II, compared with the other V_H

families, the combination repertoire of V_H7183 family does show more obvious variation between the pre-B and IM-B cells than that between IM-B and S-B cells ($r = 0.79$ vs. $r = 0.95$). This finding is consistent with previous report that the V_H7183 family was more susceptible to positive selection in early B cell development, especially the most D-proximal $V_H7183.2$ (equivalent to V_H81x in BALB/c strain) [45]. It is noteworthy that, the combination repertoire of V_H36-60 family seemed not to be selected against during the overall development of the adult B cells ($r = 0.995$ vs. $r = 0.999$). This is consistent with the finding of Decker et al. (1991), who observed an unexpectedly high representation (more than 80%) of the productive rearrangements of the V_HM315, a member of V_H36-60 family in BALB/c strain [47]. In the sequence of the entire Igh V locus of strain C57BL/6, the V_H gene segment corresponding to V_HM315 is the $V_H36-60.6$, was always predominantly utilized in the combination repertoire of V_H36-60 family in our investigation. Meng et al. suggested that the IF fraction of a rearranged V_H gene could be used as a measure of the V_H selection at the pro-B to pre-B cell transition [46]. However, for a V_H gene, a high IF fraction does not always mean the high diversity contributed to the whole repertoire, nearly all of the V_HJ558 genes had a high IF fraction.

Unlike the suggestions resulting from previous observations, which used a limited number of $V_H7183DJC\mu$ transcripts, the average CDRH3 length generated by the rearrangement machinery was found to be reduced when B cells entered the periphery, and similar results have been observed in the human [29,48]. Of the individual components of CDRH3, the D and J_H elements comprised the bulk of the length of the region by contributing more than 10 nt each, whereas the N (both 5′ and 3′ of D segment) and V_H elements added about 7 nt and 5 nt each. Notably, the reduction in average CDRH3 length was largely due to the fewer N additions but not the decrease of the germline contribution, indicating that the selection for B cells might favor more frequently the short CDRH3 encoded by the nonrandom germline sequences. Furthermore, there were no significant difference in the length of CDRH3 between productive and nonproductive rearrangements in each B cell populations, and the CDRH3 of nonproductive rearrangements was also significantly shorter in S-B cells versus IM-B cells. These results indicated that the longer CDRH3 seemed to be removed ahead of the transcription process.

It is well known that somatic rearrangement of V_H, D, and J_H genes usually results in the acquisition of premature translation-termination codons (PTCs) in the Ig genes. In our cDNA database, the nonproductive rearrangements were much fewer than the productive rearrangements. This result is in line with previous publications indicating that the PTC-containing mRNAs of TCR and Ig genes are degraded more efficiently by the process of nonsense-mediated mRNA decay (NMD) [49,50]. In mammalian

cells, the widely accepted mechanism for PTC recognition is the exon junction complex (EJC) model, which proposed that only a PTC located at least 55 nt upstream from the terminal intron can trigger NMD [51,52]. However, this rule is not applicable to some genes, such as the TCRβ and Igμ transcripts, which were downregulated even when the PTCs were located downstream of the −55 nt boundary [49,53]. Moreover, a polar effect of the NMD efficiency was also observed in both TCRβ and Igμ genes. The efficiency of NMD increased gradually as the PTC moved further downstream in Igμ gene [49], but the effect is opposite for the TCRβ [53]. Importantly, in our investigations, the novel feature of the CDRH3 length distribution of the nonproductive rearrangements was the regular and distinct reduction of the frequency in the order $3n-2$, $3n-1$ and $3n$. An earlier report indicated that the productively (PTC−) and nonproductively (PTC +) rearranged Igμ heavy chain alleles seemed to be equally well transcribed [54]. Hence, the difference in the frequency of the three kinds of CDRH3 lengths ($3n-2$, $3n-1$ and $3n$) probably reflects the difference in the sensitivity to NMD. Due to the usage of RF3 in D_H segments, PTC(s) in the overwhelming majority of the nonproductive rearrangements with $3n$-length-CDRH3 are located several nt downstream of the V_H-D junction, where they can just trigger strong NMD [49]. Each of the other two length types ($3n-2$ and $3n-1$) cause frameshifts downstream of the D-J junction, resulting in the accumulation of multiple PTCs in Cμ exons. The first PTC appears ∼180 nt or ∼60 nt downstream of the J-C junction, when the length of CDRH3 is $3n$-2 or $3n$-1. Therefore, the polarity of NMD efficiency observed in our study is similar to that reported for TCRβ transcripts: 5′ PTCs triggered more effective NMD than did 3′ PTCs [53].

The presence of excess positively charged amino acids in the CDRH3 loop is an important feature of autoreactive antibodies, especially the dsDNA binding antibodies [10,55,56]. Sequences of CDRH3 containing positively charged amino acids have been reported to be sequentially removed from the population during B cell development [10], and another study also showed that mice forced to express an antibody repertoire enriched with positively charged amino acids showed impaired B cell development and antibody production [57]. Consistent with these previous studies, we observed that the repertoire moved toward less positively charged residues in the CDRH3 loop in the transition from pre-B to S-B cells, accompanied by an increase in the negatively charged amino acids, which might be beneficial for B cell development. Furthermore, a shift in average hydrophobicity of the CDRH3 loops from nearly neutral to mildly hydrophilic has been observed from early-pre B to mature B cells when the $V_H7183DJ_HC\mu$ transcripts were analyzed [39]. A decrease of aliphatic index in CDRH3 in the development from transitional to naive B cells has also been observed in the

human [29]. Here we found a significant reduction in nonpolar amino acids and an increase in polar amino acids during B cell development, which was compatible with the previous findings.

5 Conclusions

The data presented show that the expressed μ chain repertoire, including gene segment usage, V_HDJ_H combination profile, and especially the length and amino acid composition of CDRH3 are fine-tuned during B cell development in order to establish an optimal humoral immune response to antigen.

Supplementary materials   The online version of this article at http://dx. doi.org/10.15302/J-FASE-2014017 contains supplementary materials (Appendix).

References

1. Bassing C H, Swat W, Alt F W. The mechanism and regulation of chromosomalV(D)J recombination. *Cell*, 2002, 109(2, Suppl 1): S45–S55

2. Jung D, Giallourakis C, Mostoslavsky R, Alt F W. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annual Review of Immunology*, 2006, 24(1): 541–570

3. Meffre E, Wardemann H. B-cell tolerance checkpoints in health and autoimmunity. *Current Opinion in Immunology*, 2008, 20(6): 632–638

4. Herzog S, Jumaa H. Self-recognition and clonal selection: autoreactivity drives the generation of B cells. *Current Opinion in Immunology*, 2012, 24(2): 166–172

5. Kline G H, Hartwell L, Beck-Engeser G B, Keyna U, Zaharevitz S, Klinman N R, Jäck H M. Pre-B cell receptor-mediated selection of pre-B cells synthesizing functional mu heavy chains. *The Journal of Immunology*, 1998, 161(4): 1608–1618

6. Nishimoto N, Kubagawa H, Ohno T, Gartland G L, Stankovic A K, Cooper M D. Normal pre-B cells express a receptor complex of mu heavy chains and surrogate light-chain proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 1991, 88(14): 6284–6288

7. Lassoued K, Nuñez C A, Billips L, Kubagawa H, Monteiro R C, LeBlen T W, Cooper M D. Expression of surrogate light chain receptors is restricted to a late stage in pre-B cell differentiation. *Cell*, 1993, 73(1): 73–86

8. Köhler F, Hug E, Eschbach C, Meixlsperger S, Hobeika E, Kofer J,

Wardemann H, Jumaa H. Autoreactive B cell receptors mimic autonomous pre-B cell receptor signaling and induce proliferation of early B cells. *Immunity*, 2008, 29(6): 912–921

9. Eschbach C, Bach M P, Fidler I, Pelanda R, Köhler F, Rajewsky K, Jumaa H. Efficient generation of B lymphocytes by recognition of self-antigens. *European Journal of Immunology*, 2011, 41(8): 2397–2403

10. Wardemann H, Yurasov S, Schaefer A, Young J W, Meffre E, Nussenzweig M C. Predominant autoantibody production by early human B cell precursors. *Science*, 2003, 301(5638): 1374–1377

11. Goodnow C C, Crosbie J, Adelstein S, Lavoie T B, Smith-Gill S J, Brink R A, Pritchard-Briscoe H, Wotherspoon J S, Loblay R H, Raphael K, Trent R J, Basten A. Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice. *Nature*, 1988, 334(6184): 676–682

12. Nemazee D A, Bürki K. Clonal deletion of B lymphocytes in a transgenic mouse bearing anti-MHC class I antibody genes. *Nature*, 1989, 337(6207): 562–566

13. Erikson J, Radic M Z, Camper S A, Hardy R R, Carmack C, Weigert M. Expression of anti-DNA immunoglobulin transgenes in non-autoimmune mice. *Nature*, 1991, 349(6307): 331–334

14. Gay D, Saunders T, Camper S, Weigert M. Receptor editing: an approach by autoreactive B cells to escape tolerance. *The Journal of Experimental Medicine*, 1993, 177(4): 999–1008

15. Tiegs S L, Russell D M, Nemazee D. Receptor editing in self-reactive bone marrow B cells. *The Journal of Experimental Medicine*, 1993, 177(4): 1009–1020

16. Casellas R, Shih T A, Kleinewietfeld M, Rakonjac J, Nemazee D, Rajewsky K, Nussenzweig M C. Contribution of receptor editing to the antibody repertoire. *Science*, 2001, 291(5508): 1541–1544

17. Halverson R, Torres R M, Pelanda R. Receptor editing is the main mechanism of B cell tolerance toward membrane antigens. *Nature Immunology*, 2004, 5(6): 645–650

18. Chung J B, Silverman M, Monroe J G. Transitional B cells: step by step towards immune competence. *Trends in Immunology*, 2003, 24(6): 342–348

19. Loder F, Mutschler B, Ray R J, Paige C J, Sideras P, Torres R, Lamers M C, Carsetti R. B cell development in the spleen takes place in discrete steps and is determined by the quality of B cell receptor-derived signals. *The Journal of Experimental Medicine*, 1999, 190(1): 75–90

20. Allman D, Lindsley R C, DeMuth W, Rudd K, Shinton S A, Hardy R R. Resolution of three nonproliferative immature splenic B cell subsets reveals multiple selection points during peripheral B cell maturation. *The Journal of Immunology*, 2001, 167(12): 6834–6840

21. Chung J B, Sater R A, Fields M L, Erikson J, Monroe J G. CD23 defines two distinct subsets of immature B cells which differ in their responses to T cell help signals. *International Immunology*, 2002, 14(2): 157–166

22. Su T T, Rawlings D J. Transitional B lymphocyte subsets operate as distinct checkpoints in murine splenic B cell development. *The Journal of Immunology*, 2002, 168(5): 2101–2110

23. Petro J B, Gerstein R M, Lowe J, Carter R S, Shinners N, Khan W N. Transitional type 1 and 2 B lymphocyte subsets are differentially responsive to antigen receptor signaling. *The Journal of Biological Chemistry*, 2002, 277(50): 48009–48019

24. Schelonka R L, Tanner J, Zhuang Y, Gartland G L, Zemlin M, Schroeder H W Jr. Categorical selection of the antibody repertoire in splenic B cells. *European Journal of Immunology*, 2007, 37(4): 1010–1021

25. Lange H, Hecht O, Zemlin M, Trad A, Tanasa R I, Schroeder H W Jr, Lemke H. Immunoglobulin class switching appears to be regulated by B-cell antigen receptor-specific T-cell action. *European Journal of Immunology*, 2012, 42(4): 1016–1029

26. Khass M, Buckley K, Kapoor P, Schelonka R L, Watkins L S, Zhuang Y, Schroeder H W Jr. Recirculating bone marrow B cells in C57BL/6 mice are more tolerant of highly hydrophobic and highly charged CDR-H3s than those in BALB/c mice. *European Journal of Immunology*, 2013, 43(3): 629–640

27. Weinstein J A, Jiang N, White R A 3rd, Fisher D S, Quake S R. High-throughput sequencing of the zebrafish antibody repertoire. *Science*, 2009, 324(5928): 807–810

28. Jiang N, Weinstein J A, Penland L, White R A 3rd, Fisher D S, Quake S R. Determinism and stochasticity during maturation of the zebrafish antibody repertoire. *Proceedings of the National Academy of Sciences of the United States of America*, 2011, 108(13): 5348–5353

29. Wu Y C, Kipling D, Leong H S, Martin V, Ademokun A A, Dunn-Walters D K. High-throughput immunoglobulin repertoire analysis distinguishes between human IgM memory and switched memory B-cell populations. *Blood*, 2010, 116(7): 1070–1078

30. Glanville J, Zhai W, Berka J, Telman D, Huerta G, Mehta G R, Ni I, Mei L, Sundar P D, Day G M, Cox D, Rajpal A, Pons J. Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire. *Proceedings of the National Academy of Sciences of the United States of America*, 2009, 106(48): 20216–20221

31. Glanville J, Kuo T C, von Büdingen H C, Guey L, Berka J, Sundar P D, Huerta G, Mehta G R, Oksenberg J R, Hauser S L, Cox D R, Rajpal A, Pons J. Naive antibody gene-segment frequencies are heritable and unaltered by chronic lymphocyte ablation. *Proceedings of the National Academy of Sciences of the United States of America*, 2011, 108(50): 20066–20071

32. Arnaout R, Lee W, Cahill P, Honan T, Sparrow T, Weiand M, Nusbaum C, Rajewsky K, Koralov S B. High-resolution description of antibody heavy-chain repertoires in humans. *PLoS ONE*, 2011, 6(8): e22365

33. Rubelt F, Sievert V, Knaust F, Diener C, Lim T S, Skriner K, Klipp E, Reinhardt R, Lehrach H, Konthur Z. Onset of immune senescence defined by unbiased pyrosequencing of human immunoglobulin mRNA repertoires. *PLoS ONE*, 2012, 7(11): e49774

34. Rohatgi S, Ganju P, Sehgal D. Systematic design and testing of nested (RT-)PCR primers for specific amplification of mouse rearranged/expressed immunoglobulin variable region genes from small number of B cells. *Journal of Immunological Methods*, 2008, 339(2): 205–219

35. Johnston C M, Wood A L, Bolland D J, Corcoran A E. Complete sequence assembly and characterization of the C57BL/6 mouse Ig heavy chain V region. *The Journal of Immunology*, 2006, 176(7): 4221–4234

36. Ye J. The immunoglobulin IGHD gene locus in C57BL/6 mice. *Immunogenetics*, 2004, 56(6): 399–404

37. Solin M L, Kaartinen M. Allelic polymorphism of mouse Igh-J locus, which encodes immunoglobulin heavy chain joining (JH) segments. *Immunogenetics*, 1992, 36(5): 306–313

38. Yousfi Monod M, Giudicelli V, Chaume D, Lefranc M P. IMGT/JunctionAnalysis: the first tool for the analysis of the immunoglobulin and T cell receptor complex V-J and V-D-J JUNCTIONs. *Bioinformatics*, 2004, 20(Suppl 1): i379–i385

39. Ivanov I I, Schelonka R L, Zhuang Y, Gartland G L, Zemlin M, Schroeder H W Jr. Development of the expressed Ig CDR-H3 repertoire is marked by focusing of constraints in length, amino acid use, and charge that are first established in early B cell progenitors. *The Journal of Immunology*, 2005, 174(12): 7773–7780

40. Yancopoulos G D, Desiderio S V, Paskind M, Kearney J F, Baltimore D, Alt F W. Preferential utilization of the most JH-proximal VH gene segments in pre-B-cell lines. *Nature*, 1984, 311(5988): 727–733

41. Wu G E, Paige C J. VH gene family utilization in colonies derived from B and pre-B cells detected by the RNA colony blot assay. *The EMBO Journal*, 1986, 5(13): 3475–3481

42. Malynn B A, Yancopoulos G D, Barth J E, Bona C A, Alt F W. Biased expression of JH-proximal VH genes occurs in the newly generated repertoire of neonatal and adult mice. *The Journal of Experimental Medicine*, 1990, 171(3): 843–859

43. Rolink A, Karasuyama H, Grawunder U, Haasner D, Kudo A, Melchers F. B cell development in mice with a defective lambda 5 gene. *European Journal of Immunology*, 1993, 23(6): 1284–1288

44. ten Boekel E, Melchers F, Rolink A. The status of Ig loci rearrangements in single cells from different stages of B cell development. *International Immunology*, 1995, 7(6): 1013–1019

45. ten Boekel E, Melchers F, Rolink A G. Changes in the V(H) gene repertoire of developing precursor B lymphocytes in mouse bone marrow mediated by the pre-B cell receptor. *Immunity*, 1997, 7(3): 357–368

46. Meng W, Yunk L, Wang L S, Maganty A, Xue E, Cohen P L, Eisenberg R A, Weigert M G, Mancini S J, Prak E T. Selection of individual VH genes occurs at the pro-B to pre-B cell transition. *The Journal of Immunology*, 2011, 187(4): 1835–1844

47. Decker D J, Boyle N E, Klinman N R. Predominance of nonproductive rearrangements of VH81X gene segments evidences a dependence of B cell clonal maturation on the structure of nascent

48. Shiokawa S, Mortari F, Lima J O, Nuñez C, Bertrand F E 3rd, Kirkham P M, Zhu S, Dasanayake A P, Schroeder H W Jr. IgM heavy chain complementarity-determining region 3 diversity is constrained by genetic and somatic mechanisms until two months after birth. *The Journal of Immunology*, 1999, 162(10): 6060–6070

49. Bühler M, Paillusson A, Mühlemann O. Efficient downregulation of immunoglobulin mu mRNA with premature translation-termination codons requires the 5′-half of the VDJ exon. *Nucleic Acids Research*, 2004, 32(11): 3304–3315

50. Gudikote J P, Wilkinson M F. T-cell receptor sequences that elicit strong down-regulation of premature termination codon-bearing transcripts. *The EMBO Journal*, 2002, 21(1–2): 125–134

51. Nagy E, Maquat L E. A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends in Biochemical Sciences*, 1998, 23(6): 198–199

52. Brogna S, Wen J. Nonsense-mediated mRNA decay (NMD) mechanisms. *Nature Structural & Molecular Biology*, 2009, 16(2): 107–113

53. Wang J, Gudikote J P, Olivas O R, Wilkinson M F. Boundary-independent polar nonsense-mediated decay. *EMBO Reports*, 2002, 3(3): 274–279

54. Eberle A B, Herrmann K, Jäck H M, Mühlemann O. Equal transcription rates of productively and nonproductively rearranged immunoglobulin mu heavy chain alleles in a pro-B cell line. *RNA*, 2009, 15(6): 1021–1028

55. Barbas S M, Ditzel H J, Salonen E M, Yang W P, Silverman G J, Burton D R. Human autoantibody recognition of DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 1995, 92(7): 2529–2533

56. Shlomchik M, Mascelli M, Shan H, Radic M Z, Pisetsky D, Marshak-Rothstein A, Weigert M. Anti-DNA antibodies from autoimmune mice arise by clonal expansion and somatic mutation. *The Journal of Experimental Medicine*, 1990, 171(1): 265–292

57. Ippolito G C, Schelonka R L, Zemlin M, Ivanov I I, Kobayashi R, Zemlin C, Gartland G L, Nitschke L, Pelkonen J, Fujihashi K, Rajewsky K, Schroeder H W Jr. Forced usage of positively charged amino acids in immunoglobulin CDR-H3 impairs B cell development and antibody production. *The Journal of Experimental Medicine*, 2006, 203(6): 1567–1578