

# Next-generation sequencing and single-cell RT-PCR reveal a distinct variable gene usage of porcine antibody repertoire following PEDV vaccination

Ren Li, Fang Fu, Li Feng\* & PingHuang Liu\*

State Key Laboratory of Veterinary Biotechnology, Harbin Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Harbin 150001, China

Received March 28, 2019; accepted May 12, 2019; published online July 17, 2019

Porcine epidemic diarrhea virus (PEDV) is the most common diarrhea-causing pathogen in newborn piglets. The clarifications of the overall antibody repertoire and antigen-specific antibody repertoire are essential to provide important insights into the B-cell response and reshape new vaccines. Here, we applied next-generation sequencing (NGS) technology to investigate immunoglobulin (Ig) variable (V) gene segment usage of swine B-cells from peripheral blood lymphocytes (PBL) and mesenteric lymph node (MLN) cells following PEDV vaccination. We identified the transcripts of all functional Ig V-genes in antibody repertoire. IgHV1S2, IgKV1-11, and IgLV3-4 were the most prevalent gene segments for heavy, kappa, and lambda chains, respectively, in PBL and MLN. Unlike previous studies, IgKV1, instead of IgKV2, and IgLV3, instead of IgLV8, were the prevalent Ig V-gene families for kappa and lambda light chains, respectively. We further examined the antibody repertoire of PEDV spike-specific B cells by single-cell RT-PCR. In contrast to the overall antibody repertoire, Ig V-gene segments of PEDV spike-specific B cells preferentially adopted IgHV1-4 and IgHV1-14 for heavy chain, IgKV1-11 for kappa chain, and IgLV3-3 for lambda chain. These results represent a comprehensive analysis to characterize the Ig V-gene segment usage in the overall and PEDV spike-specific antibody repertoire in PBL and MLN.

**porcine epidemic diarrhea virus (PEDV), next-generation sequencing (NGS), single-cell RT-PCR, antibody repertoire**

**Citation:** Li, R., Fu, F., Feng, L., and Liu, P.H. (2020). Next-generation sequencing and single-cell RT-PCR reveal a distinct variable gene usage of porcine antibody repertoire following PEDV vaccination. *Sci China Life Sci* 63, 1240–1250. <https://doi.org/10.1007/s11427-019-9576-2>

## INTRODUCTION

Porcine epidemic diarrhea (PED) is an acute and highly contagious swine disease caused by an *Alphacoronavirus* porcine epidemic diarrhea virus (PEDV) (Jung et al., 2015; Langel et al., 2016; Wang D et al., 2016). PEDV infection causes serious damage to the pork industry and is characterized by clinical symptoms including watery diarrhea, vomiting, anorexia, dehydration, weight loss, and high morbidity and mortality in piglets (Jung et al., 2015; Pensaert

and de Bouck, 1978; Wang D et al., 2016). PEDV is transmitted mainly through the fecal-oral route and primarily infects small intestinal epithelia *in vivo* (Huang et al., 2013; Jung et al., 2015). The PEDV spike (PEDV-S) protein is the main determinant of viral cellular tropism and the main target being recognized by most neutralizing antibodies (Liu et al., 2015; Park et al., 2011). The humoral immunity, especially the gut local mucosal antibody response against the PEDV-S protein, is the major protective mechanism against PEDV infection (Langel et al., 2016; Wang X et al., 2016). How to elicit optimal humoral immunity against PEDV is a critical question.

Potent humoral immunity fundamentally depends on the

\*Corresponding authors (Pinghuang Liu, email: [liupinghuang@caas.cn](mailto:liupinghuang@caas.cn); Li Feng, email: [fengli@caas.cn](mailto:fengli@caas.cn))

generation of a substantially diverse antibody repertoire. An antibody is composed of two heavy (IgH) and two light chains, where the light chain consists of either a kappa (IgK) or a lambda (IgL) chain. Each chain consists of a variable (V) region and a constant (C) region. The antibody V region is the antigen-binding site and determines the antibody specificity. The antibody V region is encoded using the following separate multigene elements: variable (V), diversity (D), and junctional (J) gene segments for heavy chain variable region (IGHV) or V-J for kappa chain variable region (IGKV) and lambda chain variable region (IGLV). Each gene segment contains multiple allele gene segments and is grouped into a different family based on sequence and structural similarity. Porcine IGHV contains one family of immunoglobulin heavy variable gene segment, 4 D gene segment families, and 5 J gene segment families (Eguchi-Ogawa et al., 2010; Sun and Butler, 1996; Sun et al., 1998). Porcine IGKV contains 5 V $\kappa$  families and 5 J $\kappa$  families (Butler et al., 2004; Schwartz et al., 2012a). Porcine IGLV contains 7 V $\lambda$  families and 3 J $\lambda$  families (Schwartz et al., 2012b; Sun et al., 2012). IGHV and IGLV/IGKV domains each contain four regions of relatively stable sequence, termed framework regions (FRs), and three juxtaposed spans of complementarity-determining regions (CDRs), especially CDR3 of the heavy chain (HCDR3), which is the most diverse portion of the antibody molecule and is derived from DNA rearrangement of some V, the entire D, and J gene segments (Wu et al., 2015). The V-gene segment usage of the pig pre-immune antibody repertoire was previously characterized by traditional cloning and Sanger sequencing. Previous studies by Sanger sequencing have shown that the V-gene segment usage of the porcine pre-immune antibody repertoire is highly selective (Butler et al., 2013; Sun et al., 1998). Four IGHV genes account for almost 80% of the antibody repertoire (Sun et al., 1998). This pattern was also observed in IGKV and IGLV (Butler et al., 2013). However, the information about the total antibody repertoire and the antigen-specific antibody repertoire elicited by PEDV remains unclear.

Additionally, because of the high diversity of antibodies, it is impractical to comprehensively analyze such a diverse antibody repertoire using traditional Sanger sequencing. Next-generation sequencing (NGS) has been widely applied to characterize the antibody repertoire in healthy hosts or following vaccination or infection (Sattentau, 2014; Sundling et al., 2014). The information gained by Ig-seq NGS is proving invaluable for understanding antibody responses in disease and vaccination (Sattentau, 2014).

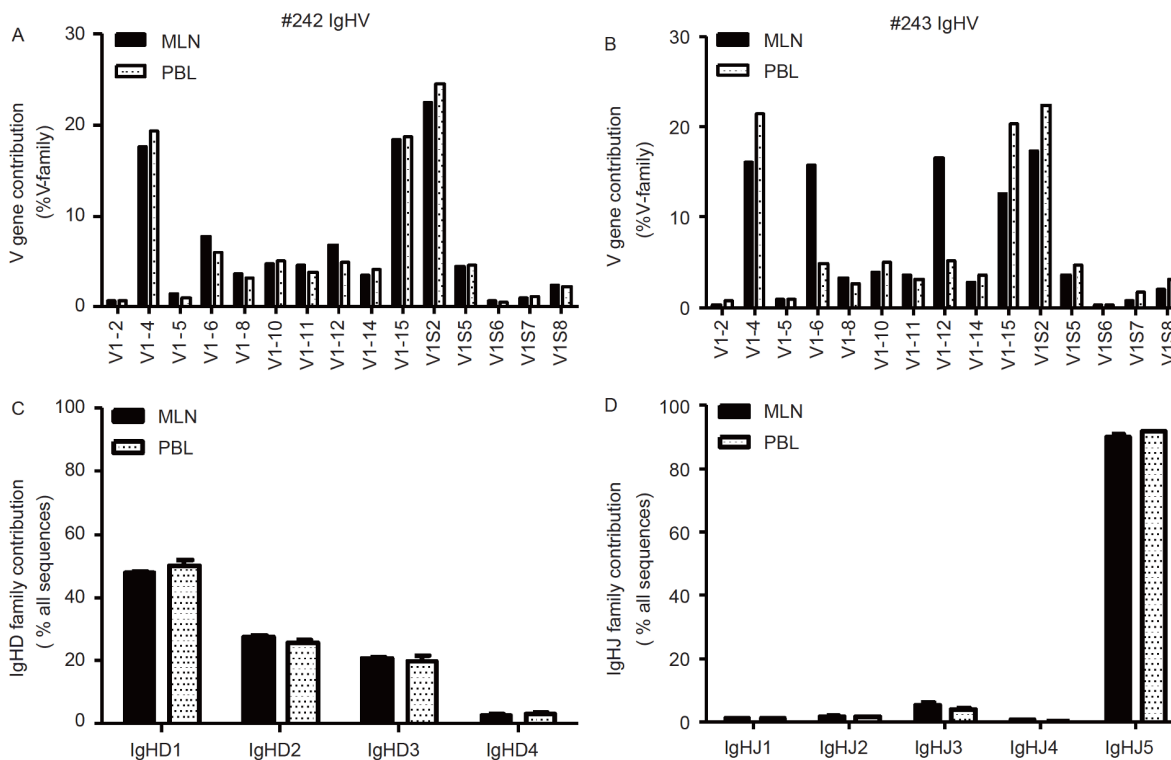
In this study, to characterize the overall Ig V-gene usage in pigs, we investigated the contribution of individual Ig V-gene segments in total B cells from peripheral blood lymphocytes (PBL) or mesenteric lymph node (MLN) by NGS. Moreover, we further characterized the Ig V-gene segments in PEDV S-specific B cells from MLN through sorted single

B cell PCR. The antigen-experienced antibody repertoire characterized by NGS demonstrated that IGHV1S2, IGHV1-4, and IGHV1-15 were the dominant IGHV genes; the highly selective IGKV1-11 accounted for more than 50% IGKV gene usage, and the IGLV3 and IGLV8 families constituted more than 99.99% IGLV V-gene usage. Some degree of compartmentalization in the Ig V-gene usage between PBL and MLN was observed though overall concordance. Interestingly, PEDV S-specific B cells exhibited a different antibody repertoire from the overall antibody repertoire and selectively expanded IGHV1-14, IGLV3-3. These data of the comprehensive analysis of Ig V-gene usage in the swine overall B cells and IgG-switched PEDV S-specific B cells following PEDV immunization are helpful for a comprehensive understanding of antibody immunity and may potentially serve as potential biomarkers for the efficacy of PEDV vaccines.

## RESULTS

### Diverse usage of IGHV genes in the total antibody repertoire

To comprehensively characterize the usage of IGHV genes, an Ig V-gene transcript library was prepared by 5' RACE PCR from cDNA and sequenced separately to reduce PCR amplification bias with NGS. We recovered more than 60,900 IGHV sequences from each sample. All members of 13 functional porcine IGHV genes were detected in both PBL and MLN, and each IGHV gene usage of the antibody repertoire ranged from 0.32% to 24.61%, indicating that all functional IGHV genes are active in the host (Figure 1A and B). Surprisingly, IGHV1S7 and IGHV1S8, annotated as non-functional pseudogenes in the IGMT database, were also transcribed and ranged from 0.75% to 3.1% of IGHV gene transcript. In PBL, the highest frequency of IGHV genes was IGHV1S2 (24.61% for pig #242 and 22.34% for pig #243) followed by IGHV1-4 and IGHV1-15 in both pigs. The usage of IGHV1-6, IGHV1-8, IGHV1-10, IGHV1-11, IGHV1-12, IGHV1-14, IGHV1S5 and IGHV1S8 was more than 2%, respectively (Figure 1A and B). For MLN, just as PBL, IGHV1S2, IGHV1-4, and IGHV1-15 were the top three dominant IGHV genes in pig #242. The frequency of IGHV1-6 and IGHV1-12 usage increased in MLN compared to PBL, especially in pig #243 MLN, in which IGHV1-6 and IGHV1-12 accounted for 15.78% and 16.55%, respectively, indicating that MLN preferentially adopts certain IGHV genes despite the overall concordance between the two compartments (Figure 1B). Meanwhile, we identified the transcripts from 4 IgHD families and 5 IgHJ families in both PBL and MLN, the frequency of each IgHD family and IgHJ family exhibited a highly similar pattern between PBL and MLN (Figure 1C and D). IgHJ5 accounted for 91.10%±0.73% total IgHJ usage, IgHJ3 constituted ranging from 3.88% to 6.45%,



**Figure 1** The usage of IgHV genes in the total antibody repertoire. The frequency of IgHV gene usage (Y-axis) was determined in PBL (white with black dots) or MLN (black) from pig #242 (A) and pig #243 (B). The number of sequences acquired by Ig-seq: A, #242-MLN,  $n=811030$  and #242-PBL,  $n=1057772$ ; B, #243-MLN,  $n=609229$  and #243-PBL,  $n=685151$ . The contribution of IgHD gene family (C) or IgHJ gene family (D) usage in PBL (white with black dots) or MLN (black) was mean $\pm$ SEM from two pigs.

and the proportion of the other 3 IgHJ families were totally about 5% (Figure 1D). All of these results show that certain IgH VDJ genes are selectively adopted, though all swine heavy chain functional V genes are actively transcribed. The undermining mechanisms are potentially consequences of the antigen-selection of B cells and clonal expansion of specific B cells.

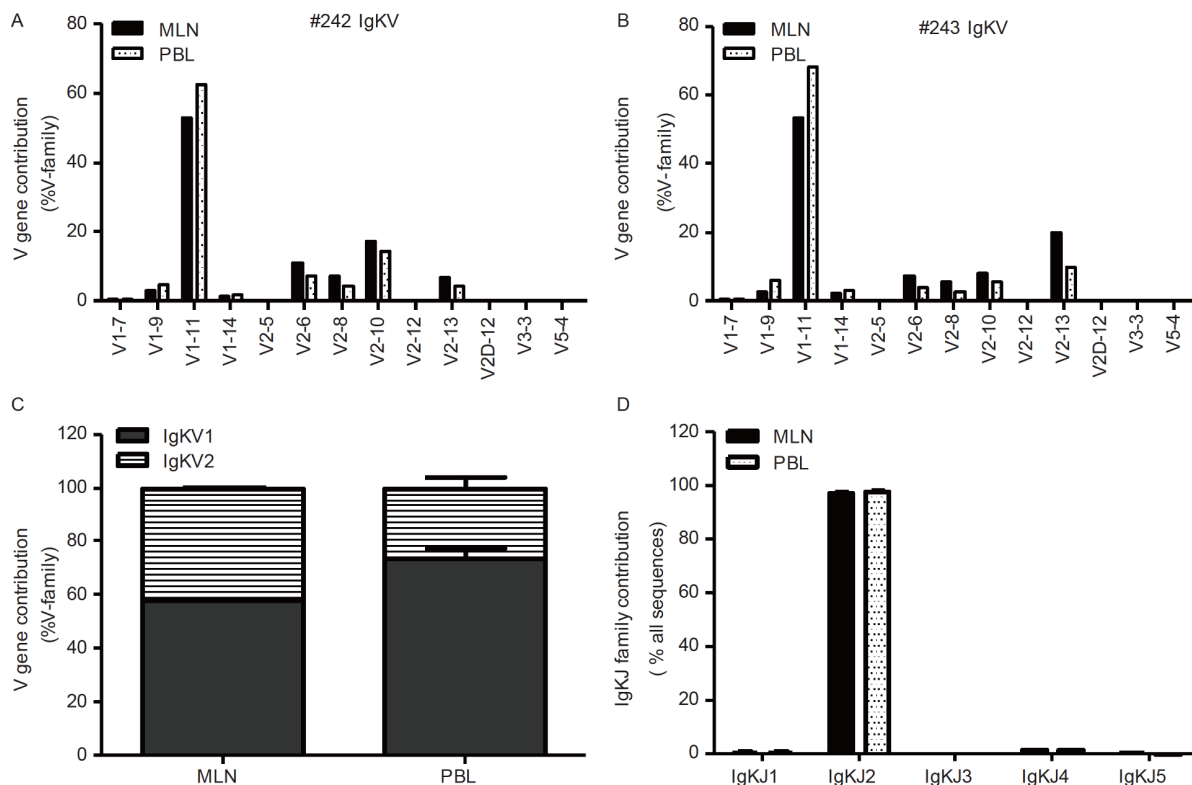
### IgKV gene usage of the antibody repertoire

Based on the IMGT database, swine IgKV genes consist of 5 different IgKV families (IgKV1, IgKV2, IgKV3, IgKV5, and IgKV7), and the only member in the IgKV7 family, IgKV7-2 is a non-functional pseudogene. We obtained ranging from 479,798 to 685,151 sequences of the kappa V-gene transcripts amplified from PBL or MLN by NGS. All the transcripts of 4 functional IgKV family were detectable: IgKV1, IgKV2, IgKV3, and IgKV5 (Figure 2A and B). Among the 4 detectable functional IgKV families, IgKV1 and IgKV2 accounted for more than 99.98% IgKV gene usage in both PBL and MLN (Figure 2C), and the rarely observed IgKV3 (IgKV3-3) and IgKV5 (IgKV5-4) only constituted 0.0011% $\pm$ 0.0002%, and 0.0014% $\pm$ 0.0002% of total IgKV genes usage, respectively (Figure 2A and B). The most prevalent IgKV1 accounted for 58.21% $\pm$ 0.53% in MLN and 73.71% $\pm$ 4.01% in PBL, demonstrating the pre-

ferential usage of IgKV1 in PBL (Figure 2C). In the IgKV1 family, IgKV1-11 accounted for more than 50% total IgKV genes usage in MLN from both pig #242 and #243. The preferential usage of IgKV1-11 in PBL was even more evident and reached 62.70% for pig #242 and 67.82% for pig #243 (Figure 2A and B). The other IgKV1 genes, including IgKV1-9 and IgKV1-14, but not IgKV1D-11, were also detectable and ranged from 1.16% to 6.08%. All the transcripts of 7 functional IgKV2 family genes were presented in our samples. The IgKV2 family genes primarily consisted of IgKV2-6, IgKV2-8, IgKV2-10, and IgKV2-13. The usage of IgKV genes was highly concordant between PBL and MLN (Figure 2A and B). All of these results demonstrate that the usage of IgKV gene is highly selective and is over-represented by IgKV1-11. For IgKJ gene segments, IgKJ2 family dominated in the total IgKJ genes though all the transcripts of 5 IgKJ families were identified. IgKJ2 accounted for more than 97.81% $\pm$ 0.20% of total IgKJ genes usage. IgKJ4 made up 1.29% $\pm$ 0.17%, IgKJ1 and IgKJ5 accounted for only 0.57% $\pm$ 0.15% and 0.30% $\pm$ 0.10% of total IgKJ gene usage respectively (Figure 2D). Thus, the usage of porcine IgKV and IgKJ is highly selective.

### IgLV gene distribution of the antibody repertoire

Among 7 porcine IgLV families (IgLV1, 2, 3, 5, 7, 8, and III),



**Figure 2** High concordance of IgKV gene usage in PBL and MLN. The frequency of IgKV gene contribution (*Y*-axis) was determined in PBL (white with black dots) or MLN (black) from pig #242 (A) and pig #243 (B). The number of sequences acquired by Ig-seq: A, IgKV, #242-MLN,  $n=581612$  and #242-PBL,  $n=590013$ ; B, IgKV, #243-MLN,  $n=479798$  and #243-PBL,  $n=685151$ . C, IgKV1 and IgKV2 two families accounted for more than 99.9% IgKV gene usage. D, The frequency of IgKJ gene family usage in PBL (white with black dots) or MLN (black) was mean $\pm$ SEM from two pigs.

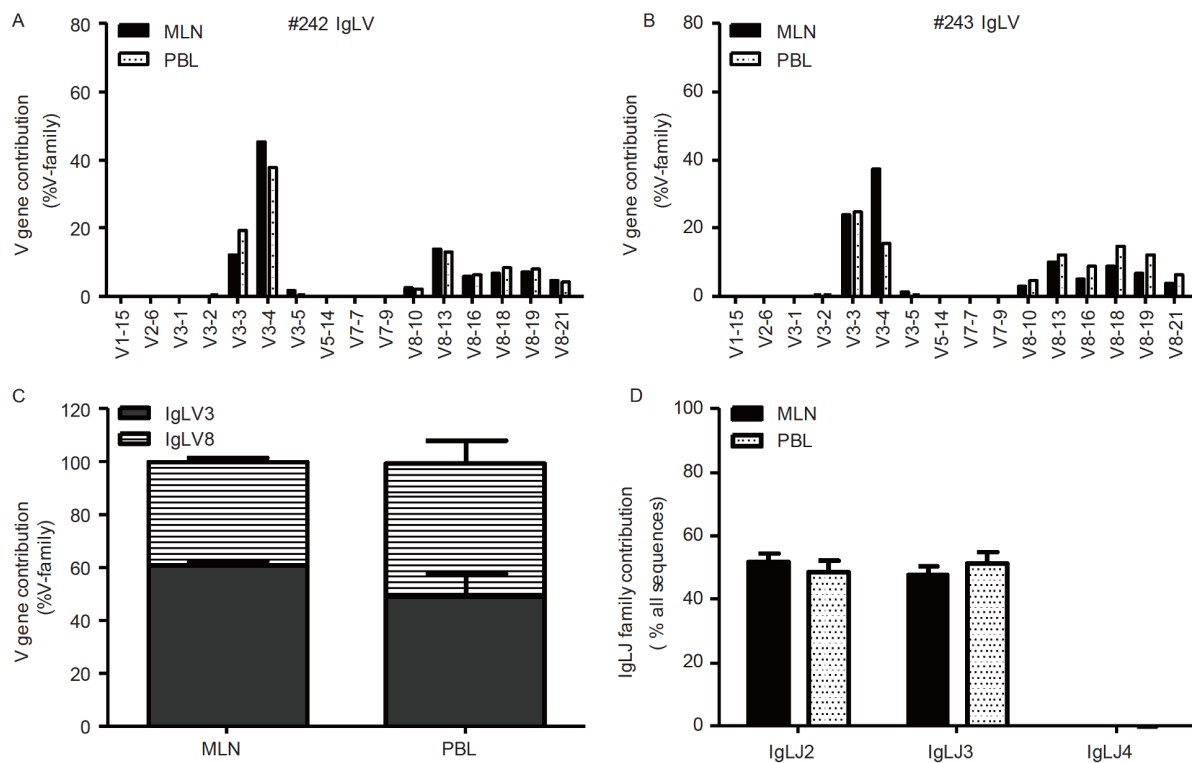
the currently identified members of IgLV1 and IgLV(III) families are annotated as non-functional pseudogenes. With the exception of IgLV(III), the transcripts of the other six IgLV gene families were all identified in PBL and MLN samples. IgLV3 and IgLV8 were the most prevalent two gene families, which totally accounted for more than 99.9% of IgLV gene usage in both PBL and MLN. The rarely observed families of IgLV1, IgLV2, IgLV5, and IgLV7 were only  $0.043\% \pm 0.014\%$ ,  $0.041\% \pm 0.026\%$ ,  $0.048\% \pm 0.017\%$ , and  $0.022\% \pm 0.005\%$  of the total IgLV repertoire, respectively (Figure 3A and B). IgLV genes in PBL were equally composed of IgLV3 and IgLV8; in contrast, MLN B cells showed a skewed usage of the IgLV3 family ( $61.00\% \pm 1.84\%$  for IgLV3 versus  $38.94\% \pm 1.84\%$  for IgLV8) (Figure 3C). The IgLV3 family in PBL and MLN was substantially composed of IgLV3-3 and IgLV3-4 genes. MLN was partial to the usage of IgLV3-4 compared to PBL (Figure 3A and B). All six members of IgLV8 family were present in our samples and there was no significant selection among them (Figure 3A and B). Only the transcripts of three IgLJ families (IgLJ2, 3, 4) were identified in samples. IgLJ2 and IgLJ3 equally accounted for almost all the IgLJ segments usage (Figure 3D).

### IgH CDR3 amino acids length distribution in PBL and MLN

Complementarity-determining region 3 of heavy chain (HCDR3) is the most variable and contributes substantially to the specificity and affinity of antibodies. The diversity of global antibody can be obtained by measuring the length distribution of the CDR3 region by a method called spectratyping (Wu et al., 2015). HCDR3 length in PBL and MLN in both pigs showed a similar Gaussian distribution, although minor individual variation was observed between pig #242 and pig #243 (Figure 4A and B). Comparing the HCDR3 length distribution in PBL and MLN, the HCDR3 14–32 aa in length accounted for  $64.24\% \pm 4.10\%$  in MLN versus  $52.60\% \pm 1.17\%$  in PBL, and the HCDR3 4–13 aa in length accounted for  $35.70\% \pm 4.10\%$  in MLN versus  $47.35\% \pm 1.17\%$  in PBL (Figure 4C). These results indicate that antibodies from MLN exhibit more diversity compared to the antibodies from PBL.

### Distinct Ig V-gene usage in the PEDV S-specific-IgG-switched B cells

We next investigated the contribution of V-gene segments



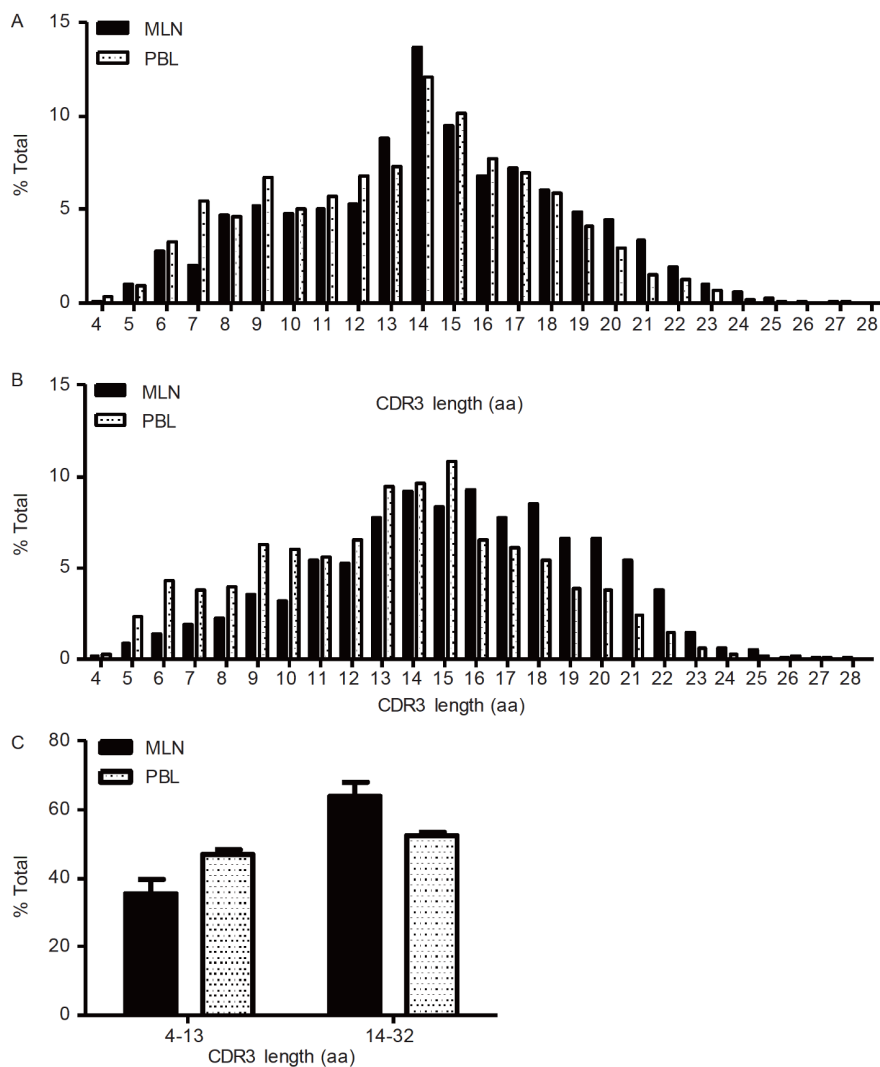
**Figure 3** High concordances of IgLV gene usage in PBL and MLN. The frequency of IgLV gene contribution (Y-axis) was determined in PBL or MLN from pig #242 (A) and pig #243 (B). The number of sequences obtained by Ig-seq: A, IgLV, #242-MLN,  $n=510118$  and #242-PBL,  $n=511364$ . B, IgLV, #243-MLN,  $n=498588$  and #243-PBL,  $n=398715$ . C, IgLV3 and IgLV8 two families accounted for more than 99.9% IgLV gene usage. D, the frequency of IgLJ gene family usage in PBL (white with black dots) or MLN (black) was mean $\pm$ SEM from two pigs.

used in PEDV vaccine-induced responses with FACS-sorted single-cell PCR (Figure 5A). Fifty-eight IgHV genes from 360 single PEDV S-specific IgG<sup>+</sup> cells of pig #242 and Eighty-five IgHV genes from 320 single PEDV S-specific IgG<sup>+</sup> cells of pig #243 were successfully cloned, which included 10 out of 13 functional IgHV1 genes (Figure 5B and E). Unlike the overall IgHV gene contribution, IgHV1-4 was the most dominant IgHV gene (37.93% in pig #242 and 34.12% in pig #243) and was followed by IgHV1-14 and IgHV1S2 (Figure 5B and E). Remarkably, IgHV1-14 was used with greater frequency in PEDV S-specific B cells than that in the total antibody repertoire (Figure 6A), suggesting that PEDV S-specific B cells selectively expand IgHV1-14 B cells. For IgKV, IgKV1-11, IgKV2-10, and IgKV2-13 were detected in a total of 42 IgKV clones from PEDV S-specific IgG<sup>+</sup> cells of pig #242 (Figure 5C). Only IgKV1-11 and IgKV2-13 were detected in a total of 31 IgKV clones of pig #243 (Figure 5F). Consistent with the pattern of the overall IgKV repertoire, IgKV1-11 was the most dominant gene (Figure 6B). For IgLV, only the predominant IgLV3 (IgLV3-3 and IgLV3-4 gene) and IgLV8 families were detected in PEDV S-specific B cells from the both pigs (Figure 5D and G). Like the usage of IgLV8 family in the total B cell repertoire, all six IgLV8 family genes were detected in the PEDV S-specific B cells, indicating that our primers cover

all IgLV alleles. In contrast to the IgLV contribution of the total B cells in MLN, the IgLV8 family outnumbered the IgLV3 family in PEDV S-specific B cells (Figure 6C). Additionally, HCDR3 with 9, 16 and 20 aa of PEDV S-specific antibodies totally accounted for 58.62% in pig #242 and 66.66% in pig #243, whereas HCDR3 of overall antibody repertoire with 9, 16 and 20 aa only made up 16.57% in pig #242 and 19.39% in pig #243, respectively (Figure 6D). In short, the PEDV S-specific B cell repertoire elicited by the PEDV vaccine engaged a similar breadth of V-genes of the total B cells, highlighting the extensive genetic diversity in the B cell response to the PEDV vaccine, but displayed a different pattern of Ig V-gene usage.

## DISCUSSION

Characterization of the overall antibody repertoire or antigen-specific antibody repertoire is critical for the comprehensive understanding of antibody responses to infection or vaccines. No studies have comprehensively investigated the overall or antigen-specific antibody repertoire in pigs with NGS, although several studies have characterized the antibody repertoire in the different growth stages of natural or infected pigs with Sanger sequencing. NGS-based repertoire

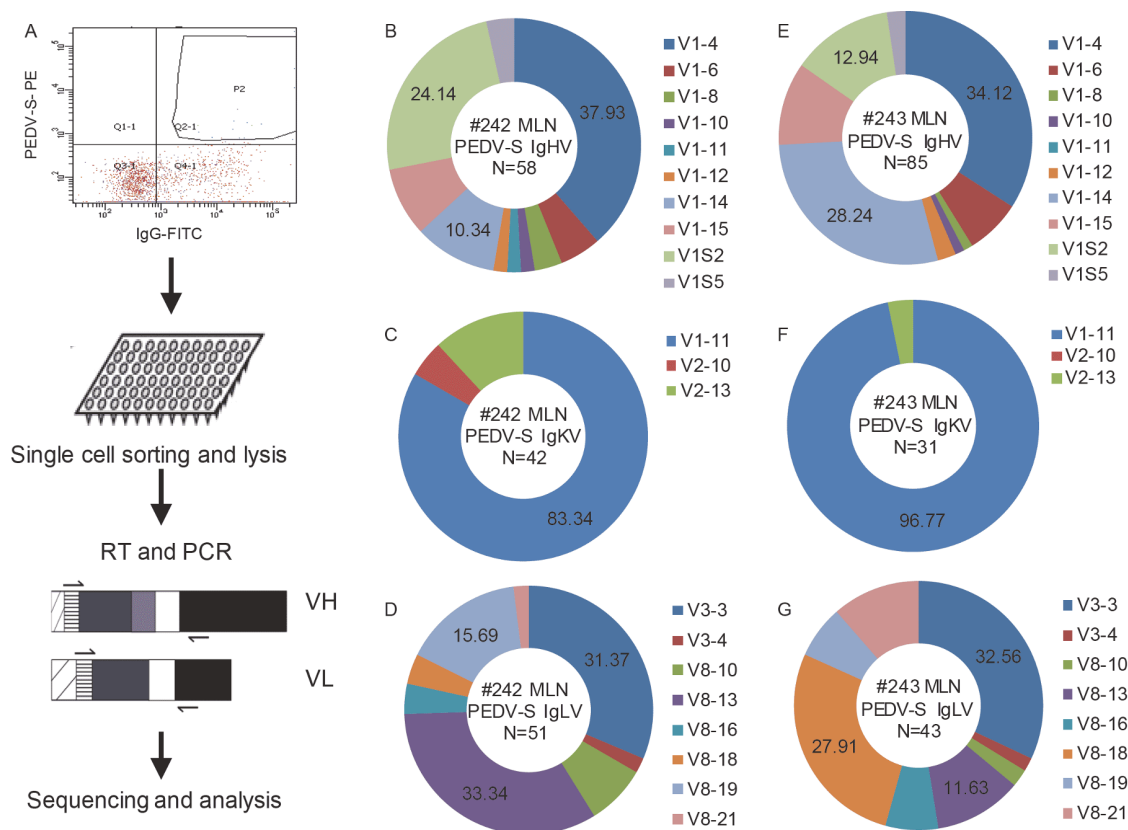


**Figure 4** The distribution of Ig heavy chain CDR3 (HCDR3) length. The HCDR3 length was determined in PBL and MLN from pig #242 (A) and pig #243 (B). C, Comparing the proportion of HCDR3 4–13 aa length and HCDR3 14–32 aa length in MLN and PBL from two pigs. The results presented mean±SEM.

analysis is helpful to comprehensively understand the antibody responses following infection or vaccination. In this study, we comprehensively investigated the porcine Ig V-gene usage in PBL and MLN at high resolution with NGS Ig library generated by 5'-RACE PCR, which offers relatively unbiased view of the antibody repertoire. Furthermore, we dissected the PEDV vaccine-induced antibody repertoire from the total antibody repertoire through sorted single-cell PCR technology combined with NGS.

Previous studies have shown that the usage of 21 IgHV genes in the pre-immune repertoire in pigs is highly restricted and that 4 IgHV genes accounted for 80% of IgHV usage (Sun et al., 1998). IgHV1-4 (denoted IgHVA in Dr. Butler's paper), IgHV1-6 (IgHVB), IgHV1S2 (IgHVB), and IgHV1S5 predominated in DNA and transcripts of IgHV genes in different fetal stages and different tissues (Eguchi-Ogawa et al., 2010; Sun et al., 1998). Here, all 13 identified functional IgHV genes were detected in samples, high-

lighting the sequencing depth of NGS. Consistent with previous studies, porcine IgHV usage is restricted that IgHV1-4 and IgHV1S2 are the two dominant IgHV genes. Unlike previous studies, IgHV1-15 was overrepresented in IgHV usage (up to 20%) in both pigs, regardless of tissue types. Relative to the usage profile of IgHV, the V-gene usage of light chains exhibited greater inconsistency with the pre-immune repertoire in the previous studies. Prior studies demonstrate that only IgKV1 and IgKV2 of the total of 5 IgKV gene families in the pig genome are transcribed and that IgKV2-10 and IgKV2-1 comprise 80% of the kappa repertoire (Butler et al., 2013). In contrast, we identified all 4 functional IgKV family transcripts though IgKV1 and IgKV2 account for more than 99.9% of the kappa repertoire (Figure 2), indicating the high resolution of NGS-based Ig analysis. Furthermore, unlike previous studies that IgKV2 outnumbers IgKV1 (Butler et al., 2013; Schwartz et al., 2012a), IgKV1 was dominantly expressed in both PBL and



**Figure 5** The PEDV spike-specific B cell Ig V-gene usage is highly polyclonal. A, Ig variable genes were amplified from FCAS-sorted PEDV S-specific B cells. Anti-PEDV spike-Dylight550 (PE) and anti-swine IgG-AF488 (FITC) were used as the sorting probes. Differential usage of Ig V-gene segments in PEDV S-specific B cells from pig #242 (B–D) and pig #243 (E–G) was displayed. RT-PCR amplified IgHV, IgKV and IgLV families from sorted single PEDV S-specific B cells were analyzed with IMGT/V-Quest. Families are color coded. The size of the colored area corresponded to the frequency of Ig V-gene usage.

MLN in the present study (Figure 2). Remarkably, the results demonstrated that the IgKV1-11 gene was overrepresented in kappa repertoire of both PBL and MLN; the IgKV1-11 contribution was even greater in PBL and comprised up to 67.82% of the IgKV gene usage (Figure 2). The predominant usage of IgKV1-11 was also observed in PEDV S-specific B cells analyzed by sorted single-cell PCR, consisted 83.34% of IgKV usage in pig #242 and 96.77% in pig #243 (Figure 5C and F). The results indicate that the highly selective usage of IgKV1-11 does occur and does not result from the bias of sequencing methods.

Corresponding with previous data on the IgLV repertoire in fetal-neonatal piglets, IgLV3 and IgLV8 families amounted to more than 99.9% of IgLV gene usage in this study. However, unlike the study described by Wertz et al. (2013), we found that IgLV3-3 and IgLV3-4, not IgLV3-6, were predominant genes in IgLV3 family. Even IgLV3-6 was not identified in more than 398,000 IgLV sequences from our samples. Unlike prior studies, IgLV8-10 had the smallest usage among 6 IgLV8 family genes. Consistent with previous study, only IgLJ2 and IgLJ3 were used in lambda V-genes rearrangements. Unlike previous study that IgLJ3 outnumbers IgLJ2 and accounts for 61% usage (Wertz et al.,

2013), IgLJ2 and IgLJ3 were equally adopted here.

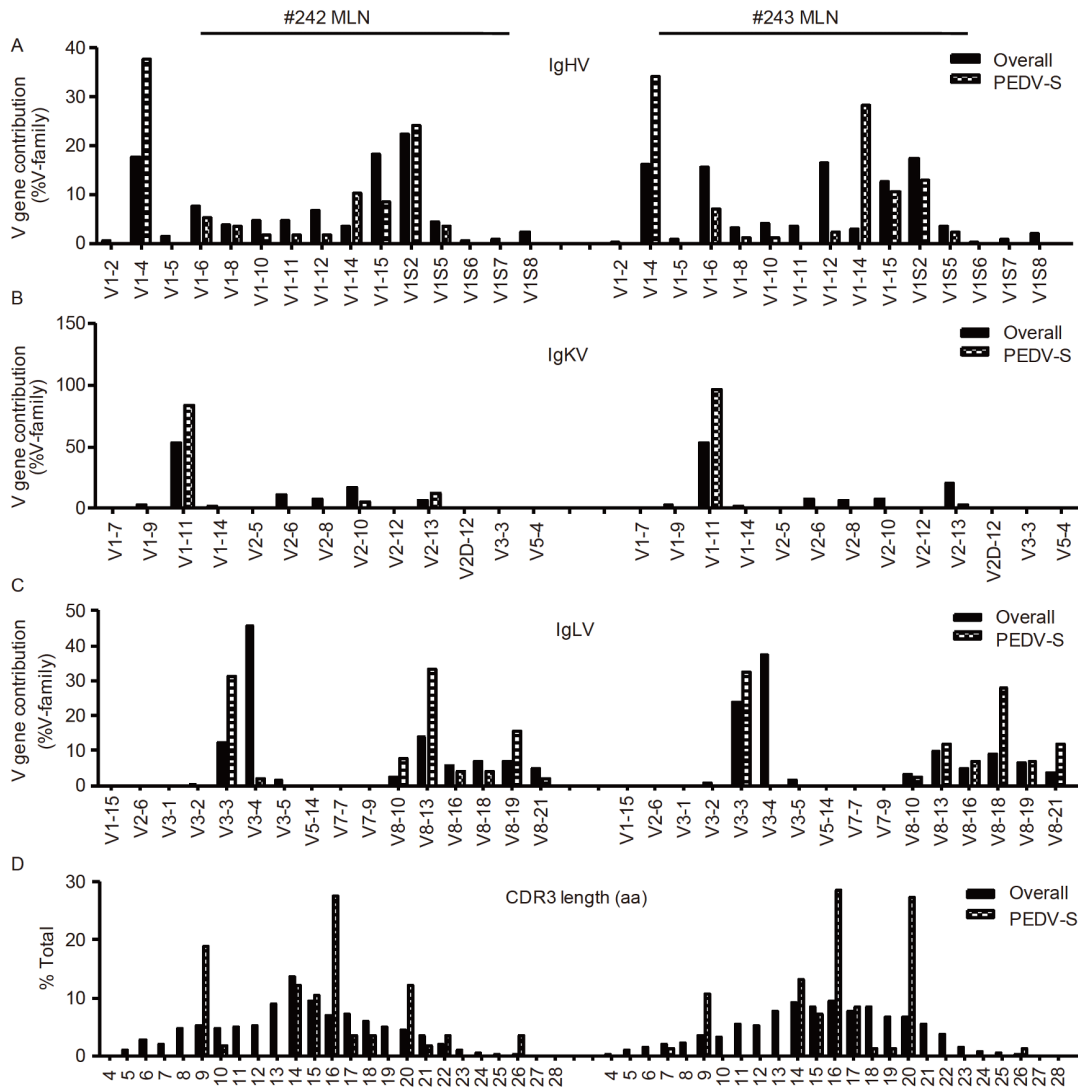
The significant variation of Ig V-gene frequency profiles in this study compared with the previous studies was not caused by genetic differences since a previous study demonstrates that the distribution pattern of IgHV usage is consistent among 35 different outbred pig fetuses (Sun et al., 1998). One of the potential reasons for this variation is the method used to generate the Ig library and sequence. Compared to previous studies using gene-specific primers, we adopted 5' RACE PCR in the template preparation of the NGS Ig library, which eliminated primer bias and provided a more accurate view of the total antibody repertoire. The stark gene usage differences caused by the primer bias of PCR in the analysis of antibody repertoire and T cell receptor repertoire have been previously reported (Carlson et al., 2013; He et al., 2014). We acquired more than 390,000 transcripts of light chains by NGS, which were 1000–10,000 folds sequences of those in previous studies. The high resolution of NGS guaranteed the identification of all active transcripts. Actually, we acquired the transcripts of all the identified functional Ig V-genes in this study. The difference in the ages of the pigs may be another reason for the variable contribution of Ig V-gene usage observed. In previous studies, the anti-

body repertoire in fetuses or neonatal piglets primarily represents the pre-immune repertoire and diversity of “natural antibodies”. In present study, the antibody repertoire represented the post-immune repertoire following the stimulation of multiple rounds of PEDV vaccine or even other non-pathogenic microbiota. These substantial inconsistencies may represent the selective expansion of some V-genes from the pre-immune repertoire following repeated antigen stimulation. The selective expansion by a specific antigen was observed in the analysis of PEDV S-specific B cells. The multiple vaccination may contribute to shape the antibody repertoire.

One unique aspect of the study is that we dissected Ig V-gene distribution of the antigen-specific B cells from total B cells by combining deep sequencing with sorted single-cell PCR. Despite the fact that PEDV S-specific Ig V-gene usage exhibited a similar diversity in the distribution of IgHV and IgK/L V genes compared with total B cells, PEDV S-specific

Ig V-gene usage displayed a highly distinct preferential usage of Ig V-genes, including IgHV1-14, IgKV1-11 and IgLV3-3, and highly restricted the usage of some V-genes such as IgHV1-12 and IgLV3-4 (Figure 6). These results indicate that high IgHV1-14, IgKV1-11 and IgLV3-3 usage by PEDV S-reactive B cells is selectively driven by PEDV S protein antigen. The multiple vaccinations may contribute to shape the PEDV S-specific Ig V-gene selective usage of antibody repertoire because of the repeated antigen stimulation. The specific role of the highly expanded Ig V-genes of PEDV-specific B cells against PEDV infection requires to further produce functional antibodies and characterize their binding profiles.

It is well established that local gut mucosal immunity, not the peripheral system, is critical to control PEDV infection in hosts (Langel et al., 2016). Characterizing the Ig repertoire in the gut like MLN will contribute to a better understanding of how Ig responses are integrated across the periphery and gut



**Figure 6** Comparison of the antibody repertoire in total B cells and PEDV S-specific B cells. The usage frequency of IgHV genes (A), IgKV genes (B), IgLV genes (C), and HCDR3 length distribution (D) in pig #242 and pig #243 was compared in total B cells and PEDV S-specific B cells, respectively.



compartments. Despite the fact that PBL and MLN exhibited a high similarity in Ig V-gene usage and HCDR3 length, there was some compartmentalization between the periphery and gut, such as the skewed usage of IgKV1 in PBL and IgLV3 in MLN (Figure 2C and 3C). We also observed individual variation. Pig #243 showed distinct IgHV usage in PBL and MLN and preferentially used IgHV1-6 and IgHV1-12 in MLN (Figure 1B) as compared to pig #242. It is cautioning when we state the distribution difference of Ig usage between PBL and MLN. The bulk B population in MLN and PBL contains various B-cell types, and the number of Ig RNA transcripts per cell varies greatly between different types of B cells, e.g., the expression of Ig encoding transcripts is much higher in plasma cells than that in naive B cells. Thus, the number of sequences identified in Ig repertoires does not reliably reflect cell numbers.

In summary, we investigated the Ig V-gene usage in the antigen exposed antibody repertoire from periphery and gut mucosa by NGS. We demonstrated the highly selective distributions of Ig V-genes though all functional Ig V-genes were actively transcribed. The antibody repertoire displayed a distinct Ig V-gene usage profile from the pre-immune repertoire previously reported. Moreover, compared to the total antibody repertoire, we found that the Ig V-gene of PEDV S-specific B cells was diverse and exhibited a distinct prevalent usage of Ig V-genes from the overall antibody repertoire. These insights describe a high-resolution picture of swine antigen-experienced antibody repertoire and might hint about the design of effective mucosal vaccination studies in the future.

## MATERIALS AND METHODS

### Isolation of PBL and MLN lymphocytes

We immunized 28-day old male Landrace pigs #242 (*SLA-I\*11:01:01*, *SLA-I\*14:01*, *SLA-2\*10:04*, *SLA-2\*07:01*, *SLA-3\*05:02*, *SLA-3\*03:02*, *SLA-DRB1\*13:01*, *SLA-DRB1\*06:02*, *SLA-DQB1\*09:01*, *SLA-DQB1\*07:01:02*) and #243 (*SLA-I\*15:02*, *SLA-2\*16:01*, *SLA-3\*07:01:02*, *SLA-DRB1\*04:03*, *SLA-DQB1\*03:03*) with an attenuated live PEDV CV777 vaccine and an inactivated PEDV virus (Fu et al., 2017). To enhance and promote the PEDV-specific antibody response, ten consecutive immunizations were performed through intrarectal or intramuscular routes at 2-week intervals. Pigs were euthanized on day 7 following the last immunization. The lymphocytes from MLN or PBL were isolated by using modifications of previously published methods (Yuan et al., 1996). The MLN was collected and placed in RPMI1640 medium (Gibco) containing 10 mmol L<sup>-1</sup> HEPES (Gibco), 200 mg mL<sup>-1</sup> gentamicin and 20 mg mL<sup>-1</sup> ampicillin. The single-cell suspensions of MLN were prepared by gently pressing tissues through stainless

steel 80-mesh screens. Blood was collected in 0.2 mol L<sup>-1</sup> EDTA tubes. PBL or MLN lymphocytes were obtained by Lymphocyte Separation Medium (LSM, MP Biomedical) density gradient centrifugation, according to the product instructions. Cells at the interface were collected and washed twice in Hanks balanced salt solution (HBSS). After extensive washing in HBSS, the purified PBL or MLN lymphocytes were frozen in FBS with 10% DMSO until use. All experiments and procedures were approved by the Local Ethical Committee on Animal Experiments of Harbin Veterinary Research Institute.

### Preparation of antibody library and next-generation sequencing

Total RNA was extracted from 10 million PBL or MLN lymphocytes with RNeasy Mini kit (QIAGEN, Shanghai, China) and quantified with Nanodrop 2000 (ThermoFisher Scientific, USA). Ig repertoire deep sequencing was performed by GENEWIZ Inc. (Suzhou, China). Briefly, 5' RACE was performed with SMARTer® RACE cDNA Amplification Kit (Clontech) with 1 µg total RNA input. Ig library amplification was performed in 50 µL reaction system containing 3 µL cDNA, 20 pmol each of 5' and 3' primers (the Universal Primer Mix (UPM) was used as the 5' primer and Ig constant-specific primers were used as 3' primers (IgG-Fc Rev: 5'-CAGGGGGCCAGAGGGTA-GACC-3'; IgK-Fc Rev: 5'-ATGAAGACGGATGGCTTG-GCATCAG-3'; IgL-Fc Rev: 5'-CGTGGGAGYGGCCTT-GGGCTGAC-3')), 10 µL 5x pfu buffer, 4 µL 2.5 mmol L<sup>-1</sup> dNTPs, and 2.5 U pfu Taq. The PCR program was initiated with a step at 95°C for 4 min, which was followed by a step comprising 25 cycles of 94°C for 30 s; 58°C for 30 s; and 72°C for 10 s. Following cycling, there was a 5-min elongation step at 72°C, followed by a final cooling step at 4°C. After performing quality control on a Bioanalyzer High Sensitivity DNA chip (Agilent, USA), libraries were sequenced on the Illumina MiSeq 2×300 platform as described previously (Zou et al., 2017).

### Single cell sorting by flow cytometry

Frozen lymphocytes from pig #242 and #243 MLN were thawed and resuspended in 10 mL RPMI 1640 media supplemented with 10% FBS. After washing, the cells were then stained using a mixture of antibodies for anti-pig CD3-PE-Cy7 (BD, Biosciences), anti-pig IgG-AF488 (SouthernBiotech), and PEDV S protein labeled with Dylight-550 at 4°C for 60 min. PEDV CV777 S protein was expressed and purified as previously described in detail (Wang X et al., 2016). S protein was labeled with Dylight-550 (ThermoFisher Scientific) according to the product's instructions. Following staining, the cells were washed in 1%BSA/PBS,

resuspended in 500  $\mu$ L, and passed through a 70- $\mu$ m nylon cell mesh. PEDV S-specific memory B cells (CD3<sup>+</sup>IgG<sup>+</sup>PEDV-S<sup>+</sup>) were sorted at single-cell density into 96-well PCR plates containing 10  $\mu$ L lysis buffer (10 mmol L<sup>-1</sup> Tris pH8.0, 1 U RNase Inhibitor) using a BD FACSAria cell sorter.

### Single-cell RT-PCR

The sorted plates were reverse transcribed and the antibody V genes were amplified as our previous study described (Fu et al., 2017). Briefly, the RNA was reverse transcribed to cDNA with a PrimeScript II 1st Strand cDNA Synthesis Kit (TAKARA, Dalian, China). Nested PCR was performed with the PrimeSTAR<sup>®</sup> HS DNA Polymerase Kit (TAKARA, Dalian, China), using primers specific for porcine heavy variable regions, the kappa variable regions, and the lambda variable regions as described in our previous study (Fu et al., 2017). Nested PCR products were evaluated in 1% agarose gels, and positive wells with a specific band of about 400 bp for V<sub>H</sub> and 370 bp for V<sub>k</sub> or V <sub>$\lambda$</sub>  were purified and sequenced by Comate Bioscience Co. (Changchun, China). V, D, and J genes and CDR3s of Ig V-genes were mapped using the International Immunogenetics Information System (IMGT) V-Quest (Brochet et al., 2008).

### Sequence analysis

The analysis of the raw Illumina MiSeq 2 $\times$ 300 data was performed as described in the bioinformatics pipeline developed by GENEWIZ Inc. Briefly, MiSeq reads Fastq data were QC filtered according to standard Illumina criteria. Sequencing primers and adaptors were trimmed along with terminal nucleotides with Q scores below 20. The quality-filtered paired reads were assembled using Pandaseq, and the clean assembled reads were mapped to V, D and J genes and CDR3s extracted using IMGT database (Brochet et al., 2008) and the IgBLAST tool (IMGT, <http://www.imgt.org>; IgBLAST, <http://www.ncbi.nlm.nih.gov/igblast>).

**Compliance and ethics** *The author(s) declare that they have no conflict of interest.*

**Acknowledgements** *This work was supported by the National Natural Science Foundation of China (31772718) and the Open Research Fund of State Key Laboratory of Veterinary Biotechnology (SKLV/BF2018XX). We thank Dr. Caixia Gao (Harbin Veterinary Research Institute, Chinese Academy of Agricultural Sciences, China) for assistance and advice in swine leukocyte antigens (SLA) types identification.*

### References

Brochet, X., Lefranc, M.P., and Giudicelli, V. (2008). 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* 36,

- W503–W508.
- Butler, J.E., Wertz, N., Sun, J., Wang, H., Chardon, P., Piumi, F., Wells, K., 2004. Antibody repertoire development in fetal and neonatal pigs. VII. Characterization of the preimmune kappa light chain repertoire. *J Immunol* 173, 6794–6805.
- Butler, J.E., Wertz, N., and Sun, X.Z. (2013). Antibody repertoire development in fetal and neonatal piglets. XIV. Highly restricted IGKV gene usage parallels the pattern seen with IGLV and IGHV. *Mol Immunol* 55, 329–336.
- Carlson, C.S., Emerson, R.O., Sherwood, A.M., Desmarais, C., Chung, M. W., Parsons, J.M., Steen, M.S., LaMadrid-Herrmannsfeldt, M.A., Williamson, D.W., Livingston, R.J., et al. (2013). Using synthetic templates to design an unbiased multiplex PCR assay. *Nat Commun* 4, 2680.
- Eguchi-Ogawa, T., Wertz, N., Sun, X.Z., Piumi, F., Uenishi, H., Wells, K., Chardon, P., Tobin, G.J., Butler, J.E., 2010. Antibody repertoire development in fetal and neonatal piglets. XI. The relationship of variable heavy chain gene usage and the genomic organization of the variable heavy chain locus. *J Immunol* 184, 3734–3742.
- Fu, F., Li, L., Shan, L., Yang, B., Shi, H., Zhang, J., Wang, H., Feng, L., and Liu, P. (2017). A spike-specific whole-porcine antibody isolated from a porcine B cell that neutralizes both genogroup 1 and 2 PEDV strains. *Vet Microbiol* 205, 99–105.
- He, L., Sok, D., Azadnia, P., Hsueh, J., Landais, E., Simek, M., Koff, W.C., Poignard, P., Burton, D.R., and Zhu, J. (2014). Toward a more accurate view of human B-cell repertoire by next-generation sequencing, unbiased repertoire capture and single-molecule barcoding. *Sci Rep* 4, 6778.
- Huang, Y.W., Dickerman, A.W., Piñeyro, P., Li, L., Fang, L., Kiehne, R., Opriessnig, T., Meng, X.J., and Griffin, D.E. (2013). Origin, evolution, and genotyping of emergent porcine epidemic diarrhea virus strains in the United States. *mBio* 4, e00737.
- Jung, K., Annamalai, T., Lu, Z., and Saif, L.J. (2015). Comparative pathogenesis of US porcine epidemic diarrhea virus (PEDV) strain PC21A in conventional 9-day-old nursing piglets vs. 26-day-old weaned pigs. *Vet Microbiol* 178, 31–40.
- Langel, S.N., Paim, F.C., Lager, K.M., Vlasova, A.N., and Saif, L.J. (2016). Lactogenic immunity and vaccines for porcine epidemic diarrhea virus (PEDV): Historical and current concepts. *Virus Res* 226, 93–107.
- Liu, C., Tang, J., Ma, Y., Liang, X., Yang, Y., Peng, G., Qi, Q., Jiang, S., Li, J., Du, L., et al. (2015). Receptor usage and cell entry of porcine epidemic diarrhea coronavirus. *J Virol* 89, 6121–6125.
- Park, J.E., Cruz, D.J.M., and Shin, H.J. (2011). Receptor-bound porcine epidemic diarrhea virus spike protein cleaved by trypsin induces membrane fusion. *Arch Virol* 156, 1749–1756.
- Pensaert, M.B., and de Bouck, P. (1978). A new coronavirus-like particle associated with diarrhea in swine. *Archiv Virol* 58, 243–247.
- Sattentau, Q.J. (2014). Immunogen design to focus the B-cell repertoire. *Curr Opin HIV AIDS* 9, 217–223.
- Schwartz, J.C., Lefranc, M.P., and Murtaugh, M.P. (2012a). Evolution of the porcine (*Sus scrofa domestica*) immunoglobulin kappa locus through germline gene conversion. *Immunogenetics* 64, 303–311.
- Schwartz, J.C., Lefranc, M.P., and Murtaugh, M.P. (2012b). Organization, complexity and allelic diversity of the porcine (*Sus scrofa domestica*) immunoglobulin lambda locus. *Immunogenetics* 64, 399–407.
- Sun, J., and Butler, J.E. (1996). Molecular characterization of VDJ transcripts from a newborn piglet. *Immunology* 88, 331–339.
- Sun, J., Hayward, C., Shinde, R., Christenson, R., Ford, S.P., Butler, J.E., 1998. Antibody repertoire development in fetal and neonatal piglets. I. Four VH genes account for 80 percent of VH usage during 84 days of fetal life. *J Immunol* 161, 5070–5078.
- Sun, X., Wertz, N., Lager, K., Sinkora, M., Stepanova, K., Tobin, G., and Butler, J.E. (2012). Antibody repertoire development in fetal and neonatal piglets. XXII.  $\lambda$  rearrangement precedes  $\kappa$  rearrangement during B-cell lymphogenesis in swine. *Immunology* 137, 149–159.
- Sundling, C., Zhang, Z., Phad, G.E., Sheng, Z., Wang, Y., Mascola, J.R., Li, Y., Wyatt, R.T., Shapiro, L., and Karlsson Hedestam, G.B. (2014).

- Single-cell and deep sequencing of IgG-switched macaque B cells reveal a diverse Ig repertoire following immunization. *J Immunol* 192, 3637–3644.
- Wang, D., Fang, L., and Xiao, S. (2016). Porcine epidemic diarrhea in China. *Virus Res* 226, 7–13.
- Wang, X., Chen, J., Shi, D., Shi, H., Zhang, X., Yuan, J., Jiang, S., and Feng, L. (2016). Immunogenicity and antigenic relationships among spike proteins of porcine epidemic diarrhea virus subtypes G1 and G2. *Arch Virol* 161, 537–547.
- Wertz, N., Vazquez, J., Wells, K., Sun, J., and Butler, J.E. (2013). Antibody repertoire development in fetal and neonatal piglets. XII. Three IGLV genes comprise 70% of the pre-immune repertoire and there is little junctional diversity. *Mol Immunol* 55, 319–328.
- Wu, Y.C., Kipling, D., Dunn-Walters, D., 2015. Assessment of B cell repertoire in humans. *Methods Mol Biol* 1343, 199–218.
- Yuan, L., Ward, L.A., Rosen, B.I., To, T.L., Saif, L.J., 1996. Systematic and intestinal antibody-secreting cell responses and correlates of protective immunity to human rotavirus in a gnotobiotic pig model of disease. *J Virol* 70, 3075–3083.
- Zou, X., Tang, G., Zhao, X., Huang, Y., Chen, T., Lei, M., Chen, W., Yang, L., Zhu, W., Zhuang, L., et al. (2017). Simultaneous virus identification and characterization of severe unexplained pneumonia cases using a metagenomics sequencing technique. *Sci China Life Sci* 60, 279–286.