vaccines

MDPI

Article

A Customizable Suite of Methods to Sequence and Annotate Cattle Antibodies

Kristel Ramirez Valdez [1], Benjamin Nzau [1,2], Daniel Dorey-Robinson [1], Michael Jarman [1], James Nyagwange [1,3], John C. Schwartz [1], Graham Freimanis [1], Angela W. Steyn [1], George M. Warimwe [3], Liam J. Morrison [2], William Mwangi [1], Bryan Charleston [1], Marie Bonnet-Di Placido [1,†] and John A. Hammond [1,*,†]

1 The Pirbright Institute, Pirbright GU24 0NF, UK
2 Roslin Institute, Royal (Dick) School of Veterinary Studies, University of Edinburgh, Midlothian EH25 9RG, UK
3 KEMRI-Wellcome Trust Research Programme CGMRC, Kilifi P.O. Box 230-80108, Kenya
* Correspondence: john.hammond@pirbright.ac.uk
† These authors contributed equally to this work.

Abstract: Studying the antibody response to infection or vaccination is essential for developing more effective vaccines and therapeutics. Advances in high-throughput antibody sequencing technologies and immunoinformatic tools now allow the fast and comprehensive analysis of antibody repertoires at high resolution in any species. Here, we detail a flexible and customizable suite of methods from flow cytometry, single cell sorting, heavy and light chain amplification to antibody sequencing in cattle. These methods were used successfully, including adaptation to the 10x Genomics platform, to isolate native heavy–light chain pairs. When combined with the Ig-Sequence Multi-Species Annotation Tool, this suite represents a powerful toolkit for studying the cattle antibody response with high resolution and precision. Using three workflows, we processed 84, 96, and 8313 cattle B cells from which we sequenced 24, 31, and 4756 antibody heavy–light chain pairs, respectively. Each method has strengths and limitations in terms of the throughput, timeline, specialist equipment, and cost that are each discussed. Moreover, the principles outlined here can be applied to study antibody responses in other mammalian species.

Keywords: antibody discovery; IgMAT; 10x Genomics; antibody sequencing; antibody repertoire

check for updates

Citation: Ramirez Valdez, K.; Nzau, B.; Dorey-Robinson, D.; Jarman, M.; Nyagwange, J.; Schwartz, J.C.; Freimanis, G.; Steyn, A.W.; Warimwe, G.M.; Morrison, L.J.; et al. A Customizable Suite of Methods to Sequence and Annotate Cattle Antibodies. *Vaccines* **2023**, *11*, 1099. https://doi.org/10.3390/vaccines11061099

Academic Editor: Daniel Dory

Received: 12 May 2023
Revised: 7 June 2023
Accepted: 9 June 2023
Published: 14 June 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The specificity and magnitude of the B cell response to vaccination and infection dictates the efficacy of the mammalian humoral response to infectious diseases [1–4]. Cattle are a global protein source that remain burdened by a range of pathogens that reduce health and productivity, such as foot-and-mouth disease, trypanosome, or bovine respiratory disease [5–8]. A detailed high-resolution characterization of the antibody response in cattle is essential to help address the need to develop new or improve existing vaccines for often complex pathogens.

Immunoglobulins (Ig) are B cell antigen receptors that can be cell-surface expressed or secreted as circulating antibodies by plasma cells, which produce large quantities of antibodies tailored to specific antigens [9]. Their structure comprises two disulfide-bond-linked heterodimers of a heavy (H) and light (L) chain. Each chain encompasses a constant region encoding the isotype, which can be IgM, or generate IgD, IgG, IgA, or IgE for the H chains and IgL (Lambda) or IgK (Kappa) for the L chains. The antigen binding variable region (V_H and V_L for the heavy and light chains, respectively) is encoded by V, D, J (heavy chains) and V, J (light chains) gene segments. Both the H and L chains have three hypervariable domains (the complementary determining regions, CDRs) that have been shown to be in contact with antigenic epitopes [10–12], of which the CDR3 of the H chain (CDRH3) exhibits the highest levels of variability.

Cattle antibodies have several unusual characteristics compared to most other species. They have relatively few functional V_H gene segments (12) that are over 90% identical to one another [13]. The diversity of the circulating repertoire is driven through the combinatorial and imprecise assembly of V(D)J gene segments followed by extensive antigen-independent somatic hypermutation of the variable region to generate a virtually infinite diversity of the cattle antibody repertoire [14]. Furthermore, cattle CDRH3 regions average 26 amino acids in length, substantially longer than in human (average 14 amino acids), mouse (average 11 amino acids), and all other species studied [15]. A subgroup of ultralong CDRH3-containing antibodies (average 61–62 amino acids; [16–18]) are unique to cattle, with distinct tertiary structures and potentially distinctive paratopes [19]. Cattle light chain usage for all antibodies is dominated by IgL (95%) over IgK (5%) [20]. These characteristics of cattle V gene segment similarity, extensive somatic mutation, and the wide range of CDRH3 lengths requires some bespoke solutions in order to accurately analyze cattle antibody repertoires and responses [15,21].

Monoclonal antibody discovery has two major applications: first, to understand the immune response in the context of disease or vaccination; second, to generate new tools such as therapeutics or diagnostic reagents. The first two techniques for producing monoclonal antibodies with a predetermined specificity were described in 1975 and 1977 based on hybridoma generation [22] and Epstein–Barr virus infection, respectively [23]. These technologies have been extensively used for many years but can suffer from low efficiency (1–3%) and poor species-specific reagent availability [24–27]. As a higher throughput alternative, antibodies produced by phage display libraries cannot account for native H and L chain pairing [28]. The next generation of methods for antibody discovery combined single B cell sorting followed by single-cell RT-PCR for natural heavy–light chain cloning and sequencing [29–31]. Although robust and practical, this method is low throughput (relative to the whole B cell repertoire), time-consuming, and expensive [32]. Recently, the progress in applying next generation sequencing (NGS) to immunoglobulin repertoires [33] and the introduction of droplet and emulsion nucleic acid amplification technologies have advanced the single B cell techniques into a high-throughput method [34]. Microfluidic platforms can isolate single cells in droplets to either PCR-link H and L chains [35–37] or to barcode both chains with the same index [34,38], followed by the bulk PCR amplification and NGS of H and L chains [39–44]. Similarly, whole antibody repertoire studies have allowed the analysis of bulk antibody chain dynamics but lack natural H and L chain pairing information.

Recent studies showing that pairing affects the antibody structure, specificity, and avidity have confirmed the importance of natural H and L chain pairing during expression for accurate antigen specificity and avidity analyses [45–47]. Additionally, various immunoinformatic tools have been developed to analyze antigen receptor NGS data, mainly in humans and mice [48]. Altogether, the development of these new tools allowed a rapid understanding of the humoral response and antigen-specific antibody discovery in the context of the COVID-19 pandemic, impacting vaccine and therapeutic development [49,50]. These technological advances, combined with the development of new analysis pipelines, reagents, and assays adapted to veterinary immunology research, now allow antibody discovery in veterinary species at a comparable resolution and throughput to that in humans and mice [32,51–54].

Here, we describe a range of validated methods for cattle single-cell antibody sequencing by combining flow cytometry single B cell sorting and (1) limited-cycle semi-nested PCR, cloning, and Sanger sequencing (low-throughput); (2) limited-cycle semi-nested PCR amplification and indexing, followed by NGS (medium-throughput); or (3) flow cytometry bulk sorting and 10x Chromium Next Gel Bead-in-Emulsion (GEM) processing adapted to cattle antibody sequencing (high-throughput). A cattle antibody sequence analysis using the Ig-Sequence Multi-Species Annotation Tool (IgMAT; [51]), showed no bias between methods. We discuss the flexibility of the workflow, emphasized by the interchangeability of the methods within the flow cytometry, PCR amplification, and sequencing modules.

2. Materials and Methods

2.1. Cattle Peripheral Mononuclear Cell (PBMC) Preparation

PBMCs from three Holstein–Friesian cattle (one per strategy below, all aged between 1 and 2.5 years) were used. The three animals were vaccinated using a standard regime for dairy cattle in the UK to protect them against leptospirosis, respiratory, and clostridial diseases. The PBMCs were isolated from heparinized blood samples as previously described [55]. Briefly, blood samples were diluted 1:2 in Dulbecco's phosphate-buffered saline without calcium and magnesium (DPBS, Gibco, Thermo Fisher Scientific, Waltham, MA, USA) and layered on SepMate PBMC isolation tubes (Stemcell Technologies, Vancouver, BC, Canada) containing Histopaque 1.083 g/mL (Sigma-Aldrich, Burlington, MA, USA). The tubes were centrifuged for 30 min at 1200× *g*. The mononuclear cell layer was recovered and treated with ACK lysing buffer (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) to lyse the remaining red blood cells. The recovered mononuclear cells were washed twice with cold DPBS and counted using a TC10 Automated Cell Counter (Bio-Rad, Hercules, CA, USA) as recommended by the manufacturer. The cells were resuspended in horse serum (Sigma-Aldrich, Burlington, MA, USA) containing 10% DMSO (Sigma-Aldrich, Burlington, MA, USA) and cryopreserved until further use.

2.2. Cell Staining and Sorting

The antibodies were obtained from the Immunological Toolbox (The Pirbright Institute, Woking, UK, unless stated otherwise (Table S1) and were labeled with fluorochromes using Lightning coupling kits as recommended by the manufacturer (Expedeon, Abcam, Cambridge, UK). All flow cytometry staining steps were performed in wells from V-bottom 96-well plate for 15 min at room temperature and protected from light followed by a single wash with 200 µL of flow buffer containing 2 mM of EDTA (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) and 0.5% horse serum (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) and a final wash with 200 µL of DPBS.

For staining strategy 1, the PBMCs were stained with LIVE/DEAD Fixable Aqua (L/D-Aqua, Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA), anti-cattle CD3-PE (Clone MM1A, Bio-Rad, Hercules, CA, USA), anti-cattle CD8a-PE (Clone CC63), anti-cattle light-chain IgL-APC (Clone IL-A58), and ER Tracker Green (Thermo Fisher Scientific, Waltham, MA, USA). Viable $CD3^-$ $CD8a^-$ IgL^{lo} ER^{hi} cells were single-cell-sorted into Hard-Shell 96-Well PCR plates (Bio-Rad, Hercules, CA, USA) containing 10 µL of buffer per well (10 mM TRIS pH 8.0 with 1000 units/mL of RNasin Plus RNase Inhibitor (Promega, Fitchburg, WI, USA); Figure 1A). After sorting, the plates were sealed with adhesive foil, spun at 1000 rpm for 1 min at 4 °C, and stored at −80 °C until further use. For staining strategy 2, the cells were stained with LIVE/DEAD Fixable Near-IR (L/D-NIR, Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA), anti-cattle CD8a-FITC (Clone CC63), and anti-cattle IgL-PerCP/Cy5.5 (Clone IL-A58). Viable $CD8a^-$ IgL^+ cells were single-cell sorted as described above (Figure 1B). For staining strategy 3, the cells were stained with LIVE/DEAD Fixable Near-IR (L/D-NIR, Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA), anti-cattle CD40-PerCP/Cy5.5 (clone IL-A158), and anti-cattle IgL-DyLight 405 (Clone IL-A58). Live $CD40^+$ IgL^+ cells were sorted into a 1.5 mL microtube (Eppendorf, Hamburg, Germany) containing DPBS with 0.04% BSA (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) and immediately used for downstream steps (Figure 1C).

All sorting methods were performed on a BD FACS Aria III Cell Sorter (BD Biosciences, Franklin Lakes, NJ, USA) as described previously [56].

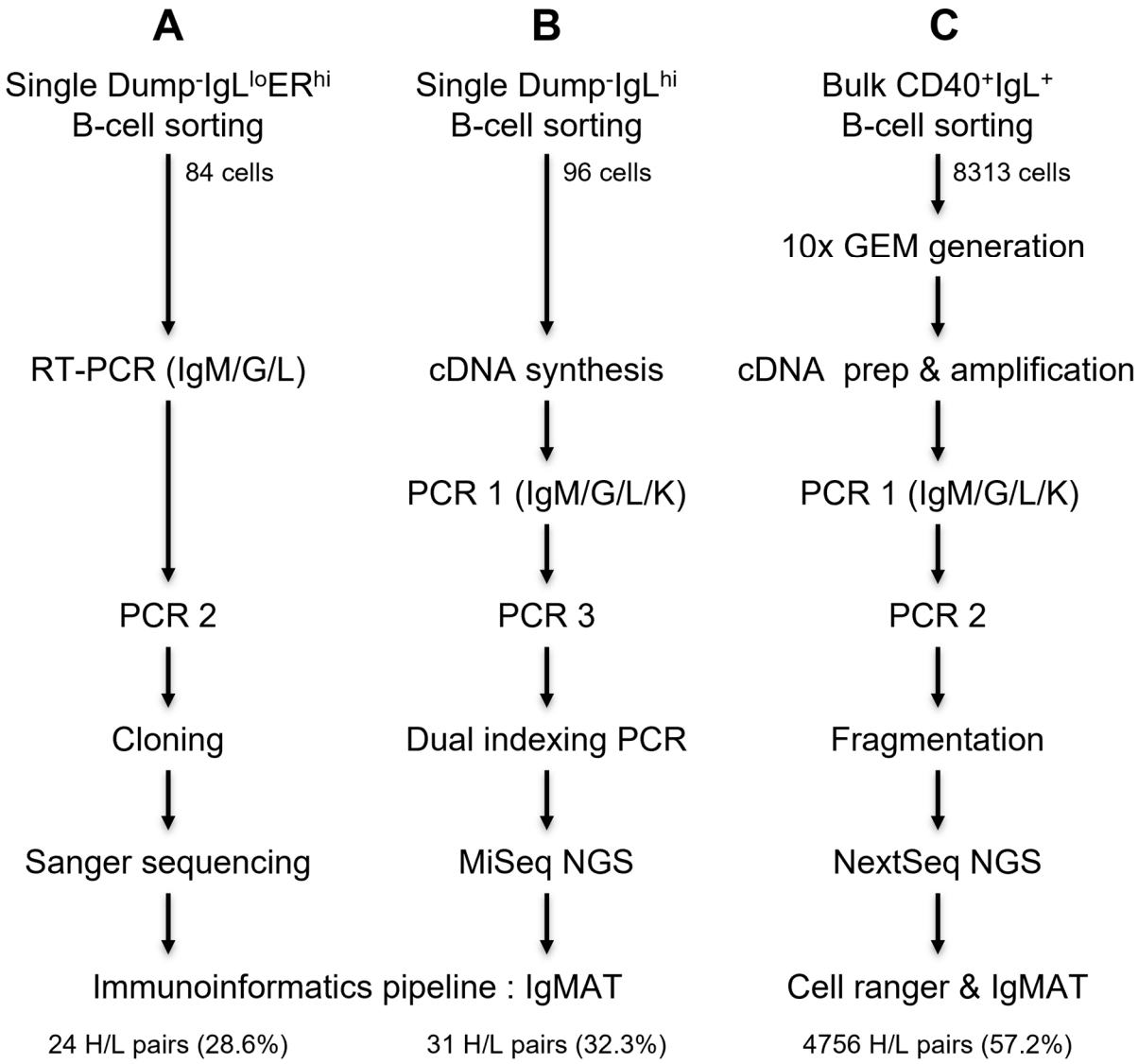

Figure 1. Three strategies for antibody heavy- and light-chain natural pair sequencing. B cells are single-cell-sorted in plates, and a semi-nested PCR is followed by cloning and Sanger sequencing ((**A**) strategy 1—low-throughput workflow) or Illumina adaptors and indexes are added by PCR amplification followed by NGS ((**B**) strategy 2—medium-throughput workflow). B cells are bulk-sorted and single cells are partitioned in Gel Beads-in-Emulsion (GEM) using 10x Chromium Controller followed by semi-nested PCR, fragmentation, and NGS ((**C**) strategy 3—high-throughput workflow). Analyses of sequenced heavy and light chains are performed using the Ig-Sequence Multi-Species Annotation Tool (IgMAT, [51]).

2.3. Primer Design

All primers reported here were designed as part of this study, except for the Igλ inner-reverse primer [31] and Read 1/P5+Read 1 primers [57]. To design antibody heavy and light chain constant domain-specific primers, mRNA, genomic, and predicted gene sequences were mined and aligned from public databases and supported by the available scaffolds in the UMD_3.1 and Btau_4.0 genome assemblies (accessions GCF_000003055.4 and GCF_000003205.2, respectively) and publicly available coding sequences (GenBank accessions AY221099, AY221098, AF097213, S82409, S82407.1, AF109167, AF005274, AF515672, AF411240, BC147881, BC146168, BC114850, BC114809, DQ497627, DQ497628, DQ452014, X62916, AY158087, X16702, and X16701) [58,59]. Several primer sites that did not contain any polymorphisms were identified to amplify cattle IgG (3), IgM (2), IgM and IgD (2),

IgL (3), and IgK (3). IgK primers were manually designed using a sequence from the cattle reference genome assembly, ARS-UCD1.2 (accession GCF_002263795.1) to cover all functional *IGKV* gene segments. The primer sequences were then verified using the NCBI primer design tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/ (accessed on 4 May 2020).

Each candidate IgH and IgL constant region primer was used as the gene-specific primer for the 5′ RACE amplification (SMARTer RACE, Takara Bio Inc., Shiga, Japan) of Ig transcripts from four unrelated Holstein–Friesian cattle cDNA derived from PBMCs, as described previously [60]. All successful amplicons were sequenced at the Centre for Genomic Research (The University of Liverpool, Liverpool, UK) using 454 Titanium technology. Contigs were generated using the MIRA assembler [61] with accurate 454 settings. The contigs were viewed and edited using the gap4 and gap5 programs from the Staden packages [62] to generate the final alignment of unique sequences and their overall frequencies. The reverse primers were further validated using more recent genome assemblies to confirm their specificity across at least *Bos* sp. The genomic regions containing the *IGH*, *IGL*, and *IGK* gene loci within the Hereford reference genome (ARS-UCD1.2, accession GCF_002263795.1) and the long-read assemblies for a yak cow (*Bos grunniens*; ARS_UNL_BGru_maternal_1.0_p; accession GCA_009493645.1), Highland bull (*Bos taurus*; ARS_UNL_Btau-highland_paternal_1.0_alt; accession GCA_009493655.1), and water buffalo (*Bubalus bubalis*; UOA_WB_1; accession GCA_003121395.1) were queried for primer sequence conservation using both Artemis [63] and BLAST [64].

All forward V gene primers matched each of the assemblies to varying extents; however, as the V regions are highly repetitive, they are often both poorly annotated and prone to mis-assembly. Combined with this potential for mis-assembly and haplotype variation in the gene content, there is uncertainty in the correct number of V genes present in any given assembly or individual. Therefore, it is difficult to ascertain whether the forward primers enable efficient amplification across bovine species. For this reason, multiple degenerate forward primers based on the available public sequences were used to maximize the number of amplified V genes. Comprehensive heavy and light chain V gene primer sets were designed manually, including some redundancy, to amplify all recovered transcripts. Furthermore, the coverage of each constant region primer was compared to limit bias, with the final primers selected based on complete representation across pools. These V and constant gene primer combinations were then used on the same cDNA samples for the PCR (95 °C for 15 min, 35 cycles as follows: 94 °C for 30 s, 58 °C for 30 s, 72 °C for 45 s, and a final extension at 72 °C for 10 min, using HotStart Taq DNA Polymerase; Qiagen, Venlo, The Netherlands) and sequenced using the same 454 process to validate the comprehensiveness of the V gene amplification. These primers (see the list of primers in Table 1) were then taken forward into the subsequent amplification pipelines.

Table 1. Cattle heavy- and light-chain targeted primer sequences used for the different amplification strategies *.

Primer Name	Primer Sequence	Orientation	Workflow 1		Workflow 2		Workflow 3	
			RT-PCR	PCR 2	PCR 1	PCR 3	PCR 1	PCR 2
IgG_C3 #	GGCACCCGAGTTCCAGGTCA	RV	✓		✓		✓	
IgG_C2	CACCGGCTCGGGCATGTAGCTGG	RV						✓
IgG_C1	GGGTAGACTTTCGGGGCTGTGGTGG	RV		✓				
IgG_C1_OH	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGTA GACTTTCGGGGCTGTGGT	RV				✓		
IgM_C3	AGGGCCACCGTGCTCTCATC	RV			✓		✓	
IgM_C2	CGAGCTCACGCAGGACACCA	RV	✓					✓
IgM_C1	CTTTCGGGTGTGATTCACCT	RV		✓				
IgM_C2_OH	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGAC TCTCRGGTGTGATTCAC	RV				✓		
IgM1_C1_OH	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGAC TCTCGGGAGCGATTCAC	RV				✓		

Table 1. *Cont.*

Primer Name	Primer Sequence	Orientation	Workflow 1		Workflow 2		Workflow 3	
			RT-PCR	PCR 2	PCR 1	PCR 3	PCR 1	PCR 2
IgH_V1	ACTGTGGACCCTCCTCYTKGTGY	FW	✓	✓	✓			
IgH_V2	ACTGTGGACCCTSSTCTTTGTGC	FW	✓	✓	✓			
IgH_V3	ACTGTGGRCYCTCCTCTTTGTGC	FW	✓	✓	✓			
IgH_V4	VYYGTGGACCCTCCTCTTTGTGC	FW	✓	✓	✓			
IgH_V5	ACTGTGGACCCTCCTCTTTVTVC	FW	✓	✓	✓			
IgH_V1_OH	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACTGT GGACCCTCCTCYTKGTGY	FW				✓		
IgH_V2_OH	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACTGT GGACCCTSSTCTTTGTGC	FW				✓		
IgH_V3_OH	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACTGT GGRCYCTCCTCTTTGTGC	FW				✓		
IgH_V4_OH	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGVYYGT GGACCCTCCTCTTTGTGC	FW				✓		
IgH_V5_OH	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACTGT GGACCCTCCTCTTTVTVC	FW				✓		
IgL_C4	CGGGTAGAAGTCGCTGATGA	RV			✓		✓	
IgL_C2	CCGTTGAGCTCCTCCGTGGAG	RV	✓					
IgL_C1	CGAGGGTGSGGACTTGGGCTGAC	RV		✓				
Igλ chain Inner-Reverse #	GCGGGAACAGGGTGACCGAG	RV						✓
IgL_C3_OH	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCGGG TAGAAGTCGCTGATGA	RV				✓		
IgL_V1	CAGGCTSYACTGACTCAGCCR	FW	✓	✓				
IgL_V2	CAGVCTGKSCTGACTCAGCCK	FW	✓	✓				
IgL_V3	CAGGMTRTGCTGACKCAGCCG	FW	✓	✓				
IgL_V4	CAGGMTVTRCTGACTCAGCCG	FW	✓	✓				
IgL_V5	CAGGMTRTGCTGACKCAGCCG	FW	✓	✓				
IgL_V6	CAGGCTGGYCTGACTCAGCCG	FW	✓	✓				
IgL_V7	CAGGCTGTVCTRACBCAGCCG	FW	✓	✓				
IgL_V8	CAGGCTGTGCTKRCTCARCCG	FW	✓	✓				
IgL_V9	CAGGCTGTKYTGACTCAGCCR	FW	✓	✓				
IgL_V10	CAGGGTGTGCTGACTCAGCCR	FW	✓	✓				
IgL_V11	TCSTATGAACTGACMCAGYYG	FW	✓	✓				
IgL_V12	TCTTCTCARCTGACTCAGCCG	FW	✓	✓				
IgL_L1	ATGGCCYGGTCCCCTCTG	FW			✓			
IgL_L2	ATGGCCTTGGCCCCTCTG	FW			✓			
IgL_L3	ATGGCCTGGATGCTGCTT	FW			✓			
IgL_L4	ATGGYCTGGGCTCTGCTY	FW			✓			
IgL_L5	ATGGCCTGGACCCCTCTC	FW			✓			
IgL_L1_OH	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATGGC CYGGTCCCCTCTG	FW				✓		
IgL_L2_OH	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATGGC CTTGGCCCCTCTG	FW				✓		
IgL_L3_OH	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATGGC CTGGATGCTGCTT	FW				✓		
IgL_L4_OH	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATGGY CTGGGCTCTGCTY	FW				✓		
IgL_L5_OH	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATGGC CTGGACCCCTCTC	FW				✓		
IgK_C3	TTCACCAAGCACACGACAGA	RV			✓		✓	
IgK_C2	TTCAGCTGCTCATCAGATGGTT	RV						✓
IgK_C1_OH	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGAAG ACGGATGGC TCAGCATCAGACC	RV				✓		
IgK_L1	CCTTGGTCTCCTGCTGCTC	FW						
IgK_L2	CCCACTCAGCTCCTCAGTCT	FW						
IgK_L3	TGAGATTCYCTGCTCAGYTCC	FW			✓			
IgK_L4	ATGAGGTTCCCTGTCAGCTC	FW			✓			
IgK_L5	ARATTCCCTGCTCAGCTCCT	FW			✓			
IgK_L6	CTGTTCTTCTGGCTCCCAGC	FW			✓			

Table 1. *Cont.*

Primer Name	Primer Sequence	Orientation	Workflow 1		Workflow 2		Workflow 3	
			RT-PCR	PCR 2	PCR 1	PCR 3	PCR 1	PCR 2
IgK_L3_OH	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGAGA TTCYCTGCTCAGYTCC	FW				✓		
IgK_L4_OH	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATGAG GTTCCCTGTCAGCTC	FW				✓		
IgK_L5_OH	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGARATT CCCTGCTCAGCTCCT	FW				✓		
IgK_L6_OH	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTGTT CTTCTGGCTCCCAGC	FW				✓		
Read1 primer #	TCTACACTCTTTCCCTACACGACG	FW					✓	
P5+ Read1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCT ACACGAC	FW						✓

* Gene-specific primers designed for amplifying the heavy and light chains of cattle antibodies by using three different throughput approaches; Ticks identify primers used in each strategy. # Primer sequences reported elsewhere: IgG_C3 [15], Igλ chain inner-reverse [31], Read1 primer [57]; other primers were designed in- house.

IgG, IgM, IgL, and IgK primer pairs were validated by the amplification of cattle cDNA using the above-mentioned PCR conditions (see the combination of primers in the Results section). For the next-generation sequencing (medium-throughput strategy), the constant region inner primers used in PCR 3 ensured similar sizes for all isotype amplicons to limit sequencing bias against longer amplicons. For the next-generation sequencing, Illumina overhang adapter sequences were added to the gene-specific primers based on the 16S Metagenomic Sequencing Library Preparation Protocol [65] as an anchor for the indexing primers [66].

2.4. Immunoglobulin Gene Amplification and Library Preparation

2.4.1. Strategy 1: PCR Amplification and Cloning Method Used in the Low-Throughput Workflow

A semi-nested PCR was used to amplify the antibody H and L chain variable regions (V_H and V_L, respectively) from sorted single cells. All reactions measured 25 µL, and all primers are described in Table 1. First, a multiplex RT-PCR reaction was performed using the OneStep RT-PCR Kit (Qiagen, Venlo, The Netherlands) as recommended by the manufacturer with 400 µM of dNTPs, 100 µM of forward primers annealed to the 5′ of the variable region, and 100 µM of reverse primers at the 5′ of the constant region of IgG, IgM, or IgL. The RT-PCR reaction mix was incubated at 50 °C for 30 min for cDNA synthesis, followed by an activation step at 95 °C for 15 min and 15 PCR amplification cycles (94 °C for 30 s, 58 °C for 30 s, 72 °C for 1 min), with a final extension step at 72 °C for 10 min (Table S2A). The amplified products were diluted 1:20 and 3 µL was used as a template for further PCR reactions (PCR 2). The PCR 2 reactions were performed using 0.025 U/µL (0.625 U) of HotStart Taq DNA Polymerase (Qiagen, Venlo, The Netherlands) with 200 µM of dNTPs and 100 µM of each primer. The amplification procedure was as follows: 95 °C for 15 min, followed by 49 cycles (94 °C for 30 s, 58 °C for 30 s, 72 °C for 1 min) and a final extension step of 10 min at 72 °C (Table S2A). The PCR products were resolved on a 2% low melting temperature agarose gel (Sigma-Aldrich, Burlington, MA, USA) at 100 volts for 45 min and the amplicons (of approximately 500 bp for V_H and 400 bp for V_L) were excised and purified using the Monarch DNA Gel Extraction Kit (New England Biolab, NEB, Ipswich, MA, USA), cloned into pGEM-T Easy Vector (Promega, Fitchburg, WI, USA), and propagated in *E. coli* JM109 competent cells (Promega, Fitchburg, WI, USA). The selection was performed using LB agar media plates containing 100 µg/mL of ampicillin (Sigma-Aldrich, Burlington, MA, USA) and supplemented post-solidification with 100 µL of 100 mM IPTG (Sigma-Aldrich, Burlington, MA, USA) and 20 µL of 50 mg/mL X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside, Promega, Fitchburg, WI, USA). White colonies were picked and the correct insert was confirmed

by PCR using M13 Forward and Reverse Primers [67] and DreamTaq DNA Polymerase (Thermo Fisher Scientific, Waltham, MA, USA).

2.4.2. Strategy 2: PCR Amplification and Indexing Method Used in the Medium-Throughput Workflow

Here, cDNA synthesis and single-plex semi-nested PCR reactions were used to amplify the antibody H and L chains. First, the cDNA was synthesized from sorted single cells as previously described [56] in a 20 µL reaction containing 1 µL of Sensiscript reverse transcriptase (Qiagen, Venlo, The Netherlands), 2 µM dT23VN oligonucleotides (NEB, Ipswich, MA, USA), 4 µM Random Primer 9 (NEB, Ipswich, MA, USA), 10 units of RNaseOUT (Thermo Fisher Scientific, Waltham, MA, USA), and 500 µM of dNTPs (Qiagen, Venlo, The Netherlands), and incubated at 37 °C for one hour. The amplification of the V_H and V_L was achieved in 25 µL reactions using a semi-nested PCR approach, using 4 µL of the cDNA product and 0.016 U/µL (0.4 Units) of Q5 Hot Start High-Fidelity DNA Polymerase (NEB, Ipswich, MA, USA), with 200 µM of dNTPs, 50 µM of each primer, Q5 high-GC enhancer as recommended by the manufacturer, and 1 M of betaine (Sigma-Aldrich, Burlington, MA, USA). All primers are described in Table 1. Touch-down PCR 1 was composed of 30 s of activation at 98 °C, followed by five cycles at 63 °C (98 °C for 10 s, 63 °C for 15 s, and 72 °C for 20 s), then 20 cycles at 58 °C (98 °C for 10 s, 58 °C for 15 s, and 72 °C for 20 s) and a final extension step of 2 min at 72 °C (Table S2B). The PCR 1 products were used as the template for a second round of PCR reactions (named PCR 3) in which overhang (-OH) Illumina adapter sequences were added to primers as anchors for the indexing PCR [65,68]. The PCR 3 reaction mix was as above and the amplification was performed in two steps starting with five cycles at a lower annealing temperature (58 °C) followed by either 30, 25, 27, and 28 cycles at 63 °C annealing for IgG, IgM, IgL, and IgK, respectively (Table S2B). The expected amplicon sizes were approximately 600 bp. The PCR products were purified with AMPure XP beads (Beckman Coulter Life Sciences; Indianapolis, IN, USA) using a 0.8 bead/amplicon volume ratio following the manufacturer's protocol. MiSeq Illumina indexes [66] were added using an 8 cycles amplification PCR (PCR 4: 98 °C for 10 s, 58 °C for 15 s, 72 °C for 20 s; [68]). The final products were purified using the AMPure XP beads.

2.4.3. Strategy 3: PCR Amplification Method Used in the High-Throughput Workflow

Sorted B cells were processed on a 10x Chromium Controller device using Next GEM Single Cell V(D)J v1.1 Reagent Kits (10x Genomics, Pleasanton, CA, USA) following the manufacturer's instructions. Briefly, the cells were prepared for loading in 1× PBS with 0.04% w/v BSA and counted on a TC10 Automated Cell Counter (Bio-Rad, Hercules, CA, USA). The cells were split into three samples containing approximately 1500, 3000, and 7700 cells for recovery of approximately 900, 1800, and 4500 antibodies (designated samples 1, 2, and 3, respectively) and dispensed into three channels of a G-chip to generate single-cell Gel Bead-in-Emulsions (GEMs). The formed GEMS were collected into 0.2 mL microtubes (Eppendorf, Hamburg, Germany) and a GEM reverse transcription reaction to generate cDNA was performed at 53 °C for 45 min, followed by inactivation for 5 min at 85 °C. A GEM clean-up was performed using Dynabeads MyOne SILANE (10x Genomics, Pleasanton, CA, USA) followed by amplifying barcoded full-length cDNA from polyadenylated mRNA with 16, 14, and 13 cycles for 900, 1800, and 4500 target cell recovery, respectively. Sample 3 cDNA was used to test the compatibility of three constant reverse primers of each isotype (IgM_C1-3, IgG_C1-3, IgL_C2, IgL_C4, Ig λ chain inner [31], IgK_C1-3) with 10x forward primer (Read 1) [57] using Kapa HiFi Taq polymerase (Roche, Basel, Switzerland) for 3 min at 95 °C, followed by 32 amplification cycles (98° C for 20 s, 62 °C for 15 s, and 72 °C for 30 s) and a final extension step of 2 min at 72 °C. Two reverse primers were selected for each isotype to process samples for NGS (see Table 1). The first semi-nested V(D)J target enrichment was performed using Read1 and constant region-specific outer reverse primers. In contrast, the second semi-nested PCR used the P5+Read1 forward primer with constant region inner primers (Table 1). Two trials with different PCR

conditions were tested. Trial 1 consisted of 6 cycles in the first PCR reactions and 8 cycles in the second PCR reactions as recommended by 10x Genomics protocols. Trial 2 consisted of two additional cycles for the first PCR (8 cycles in the first PCR) followed by four extra cycles for the second PCR (12 cycles; Table S2C). The PCR amplicons were profiled using a Bioanalyzer High-Sensitivity DNA Kit (Agilent Technologies, Inc., Santa Clara, CA, USA) followed by library preparation and indexing for multiplexing according to the manufacturer's instructions. The libraries were bead-purified and profiled again by the bioanalyzer before sequencing.

2.5. Sequencing

For the low-throughput strategy, two clones bearing inserts at expected sizes (500 bp for H and 400 bp for L) were Sanger sequenced for each sample and chain (ABI 3730XL, Eurofins Scientific, Luxembourg, Luxembourg). Below 6 nucleotides differences between clones (calculated using Taq polymerase error rate as 1.8×10^{-4} errors/base/doubling and the number of PCR cycles used; [69]), sequences were considered to arise from the same antibody transcript.

For the medium-throughput method, amplicons derived from individual samples were visualized on a Tapestation 4200 using a DNA D1000 ScreenTape (Agilent Technologies, Inc., Santa Clara, CA, USA). Samples exhibiting a band at approximately 600 bp were pooled at a 5 nM final concentration. The library pool was quantitated using the NEBNext Library Quantitation Kit for Illumina (NEB, Ipswich, MA, USA) and confirmed using the Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). The sequencing pool was loaded on a MiSeq v3 reagent kit (Illumina Inc., San Diego, CA, USA) for 2×300 cycle paired-end sequencing according to the manufacturer's specifications with a 40% PhiX spike-in to introduce diversity.

For the high-throughput strategy, pooling was performed following the 10x Genomics pooling worksheet to produce a target library pool concentration of 5 nM, aiming to achieve a target of 5000 reads per cell (5000 reads/cell). The final library pool was quantitated using the NEBNext Library Quantitation Kit for Illumina (NEB, Ipswich, MA, USA) and confirmed using the Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). The sequencing pool was loaded on a NextSeq 150 Cycle High-Output Cartridge (Illumina Inc., San Diego, CA, USA) for paired-end sequencing on a NextSeq 550 sequencer according to the manufacturer's specifications. Samples from trial 1 were sequenced once, while samples from trial 2 were sequenced twice.

2.6. Bioinformatic Analysis

Alignments of sequences obtained from the low-throughput method were created using MEGA version 7 software [70], and gene usage data were generated using IgBLASTn (v1.0.0, https://pubmed.ncbi.nlm.nih.gov/23671333/ (accessed on 22 November 2021)) with V, D, and J sequences from IMGT [71]. The region annotation was performed by translation into six reading frames using BBMAP translate6frames (V38.71, https://jgi.doe.gov/data-and-tools/bbtools/bb-tools-user-guide/bbmap-guide/ (accessed on 22 November 2021)). Sequences that did not contain stop codons were sequence-annotated with IgMAT [51] using hidden Markov models built from IMGT cattle data. For the medium-throughput method, where a dual combinatorial indexing strategy was used, demultiplexing was carried out on the Illumina sequencer. Subsequently, each sample was quality trimmed with Trim Galore (v0.5.0, https://github.com/FelixKrueger/TrimGalore (accessed on 24 November 2021)) using a quality cut-off of 30 before the reads were merged using FLASH (v1.2.11; [72]). The gene usage and region annotation was performed with IgBLASTn, BBMAP translate6frames, and IgMAT as above. Full amino acid sequences were further clustered together using an amino acid identity rate of 96% using UCLUST (v11.0.667; [73]) and the top 10 clusters were extracted for further processing. Single chains per cell were determined by dominant clusters encompassing over 90% of the UCLUST

input sequences. Where multiple dominant clusters were present, each dominant clusters' centroid was assigned as that sample's amino acid sequence.

For the high-throughput method, samples from independent channels were demultiplexed using cellranger mkfastq (10x Genomics Cell Ranger v6.1.1, https://support.10xgenomics.com/single cell-vdj/software/downloads/latest (accessed on 1 December 2021)). Cellranger vdj was ran using inner primers from the semi-nested PCR and cattle reference created from the IMGT cattle antibody sequences using the 10x script fetch-imgt following the developers' instructions (https://support.10xgenomics.com/single cell-vdj/software/pipelines/latest/advanced/references (accessed on 1 December 2021); Table S3). The gene usage and region annotation was performed with IgBLASTn, BBMAP translate6frames, and IgMAT as above.

3. Results and Discussion

3.1. Flow Cytometry

We have developed three broad strategies for cattle B cell selection by flow cytometry (Figure 1). In the first approach, we aimed to enrich antibody-secreting plasma cells, which are known to have low levels of antibodies on their cell surfaces [74] and high endoplasmic reticulum activity [75,76]. To target the plasma cells, we stained and excluded dead cells, gated on singlets using FSC-A and FSC-H parameters, and used anti-cattle CD3 and CD8a surface markers to exclude T cells. An antibody for surface IgL combined with ER staining selected the population of IgL-low (IgL^{lo}) cells with a high ER expression signal (ER^{hi}). IgL^{lo} ER^{hi} cells comprised 0.76% of the live $CD8a^{-}$ $CD3^{-}$ singlet cells (Figure 2A) and 1.78×10^{-3}% of live singlet PBMCs, which is broadly equivalent to what has been described from plasma cells in humans and mice [76].

Vaccine efficacy prediction may also require the analysis of the overall B cell population (i.e., including memory cells) rather than plasma cells alone [77]. For the second staining strategy, we gated on the 27.2% of IgL^{+} cells from live $CD8a^{-}$ singlet PBMCs (Figure 2B) and sorted IgL^{+} single cells. Comparing the above staining strategies, plotting IgL against CD3 or CD8a expectedly showed no overlap and we concluded that staining for CD3 and CD8a was superfluous, in line with cattle B cell phenotyping strategies [53]. However, to provide more confidence in cell phenotyping, we added CD40 to the panel (a marker for sustained antibody production in humans and mice [78]). In this third strategy, CD40 enabled gating on B cells before selecting the 16.8% of IgL^{+} cells (Figure 2C). Approximately 12,000 sorted $CD40^{+}IgL^{+}$ cells were immediately processed on the 10x Chromium Controller for antibody V_H and V_L region amplification and sequencing.

These three simple flow cytometry strategies based on IgL surface staining isolate overlapping B cell phenotypes that can be preferentially selected depending on the research question and available reagents, independent of any downstream processing. Additional markers such as CD20, CD21, CD71, or CD138 can easily be included to select other and more specific B cell subsets [53]. As vaccine efficacy can be correlated to the antigen-specific antibody response [79–82], fluorescently labeled antigens could also be added to the flow cytometry panel for sorting antigen-specific B cells [21,83] for downstream antigen-specific antibody discovery [31].

3.2. Primer Design and Validation

Cattle exist as many breeds and subspecies due to their history of domestication. Over the last decade, the number, and importantly the quality, of cattle genomes has increased, allowing significantly more confidence in the design of PCR primers to conserved regions or within immunogenetic regions. However, the success of bulk and single-cell PCR strategies for antibody discovery across multiple gene segments and Ig subclasses relies on primer design, ensuring that they cover all gene segments, all SNPs are considered, and a good level of efficiency is achieved.

V_H and V_L were amplified using isotype-specific reverse constant region primers by 5′ RACE followed by 454 sequencing to enable the design of V gene segment primers. Between

24,000 and 36,000 reads per amplicon were recovered. As predicted, the amplification with IgK primers did not yield enough product for sequencing and was not pursued further using this strategy. Forward degenerate primers designed to anneal to all functional heavy and light chain V gene segments were used to amplify the cDNA and sequenced using the same 454 sequencing process to assess the comprehensiveness of the V gene segment amplification. The sequence recovery in the most abundant 500 transcripts was remarkably similar, with over 98% of sequences being in both the 5′ RACE and PCR amplicon pools, and the most abundant 20 sequences from one all ranked in with the top 22 from the other.

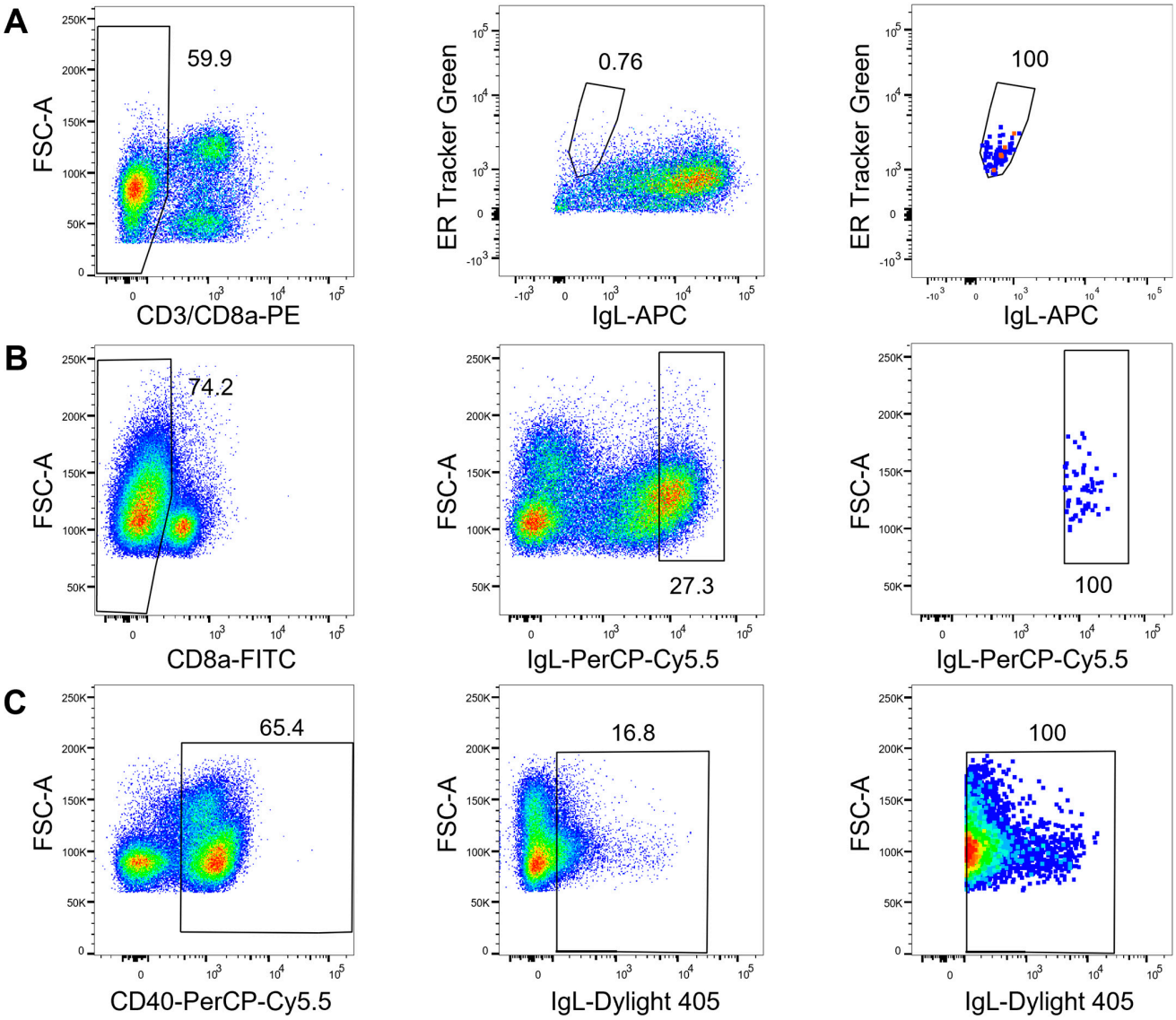

Figure 2. Three strategies of flow cytometry B cell gating. Plots in the left column represent enrichment gating strategies, the middle column represents the sort gating strategy, and the right column shows the sorted cells: (**A**) $CD3^{-}CD8a^{-}IgL^{low}ER^{hi}$ cells were single-cell-sorted for strategy 1; (**B**) $CD8a^{-}IgL^{hi}$ single cells were sorted for strategy 2; (**C**) $CD40^{+}IgL^{+}$ cells were sorted in bulk for strategy 3. All gatings were performed on live singlets lymphocytes. Percentages within each gate are indicated on the plots.

The reverse isotype-specific constant region primers were examined for their sequence conservation using a sampling of diverse long-read bovine genome assemblies (i.e., water buffalo, yak, and Highland cattle). All constant-region specific primers were overall highly conserved across bovines, with the exception of IgM_C1 and IgL_C2 (see details in Table

S4). For all antibody chains, the designed constant region reverse primers were tested independently with one V primer and showed amplification at the expected size (Figure 3A, first row). All V primers were tested independently with a constant region primer for each isotype to confirm its efficiency (Figure 3A). Band intensity variations may be explained by the number of gene segments covered by each degenerate primer or the frequency of usage of the associated (set of) gene segments. We were able to design primers to amplify all known functional V gene segments at expected sizes (400–600 bp; Figure 3) for IgG, IgM, and IgL. For the singleplex amplification used in strategy 2, we included IgK primers. Of note, the IGKV1.1 primer failed to amplify transcripts (sample 20), which correlates with the classification of the associated V gene segment as an ORF [71]. The designed primers ensured all amplicons used in the second round of PCR amplicons were similar enough in size to avoid NGS sequencing bias against longer amplicons [84]. To establish the 10x Chromium, cDNA generated from GEMs was used to test several constant region reverse primers for each isotype combined with the Read1 10x primer. Figure 3B shows that all constant region forward primers tested were compatible. For each isotype, we selected two efficient reverse primers for the first and the semi-nested PCR of the V(D)J amplification. Altogether, we designed and validated a set of primers for each isotype and strategy for cattle antibody amplification, adaptable to downstream sequencing workflows (Table 1).

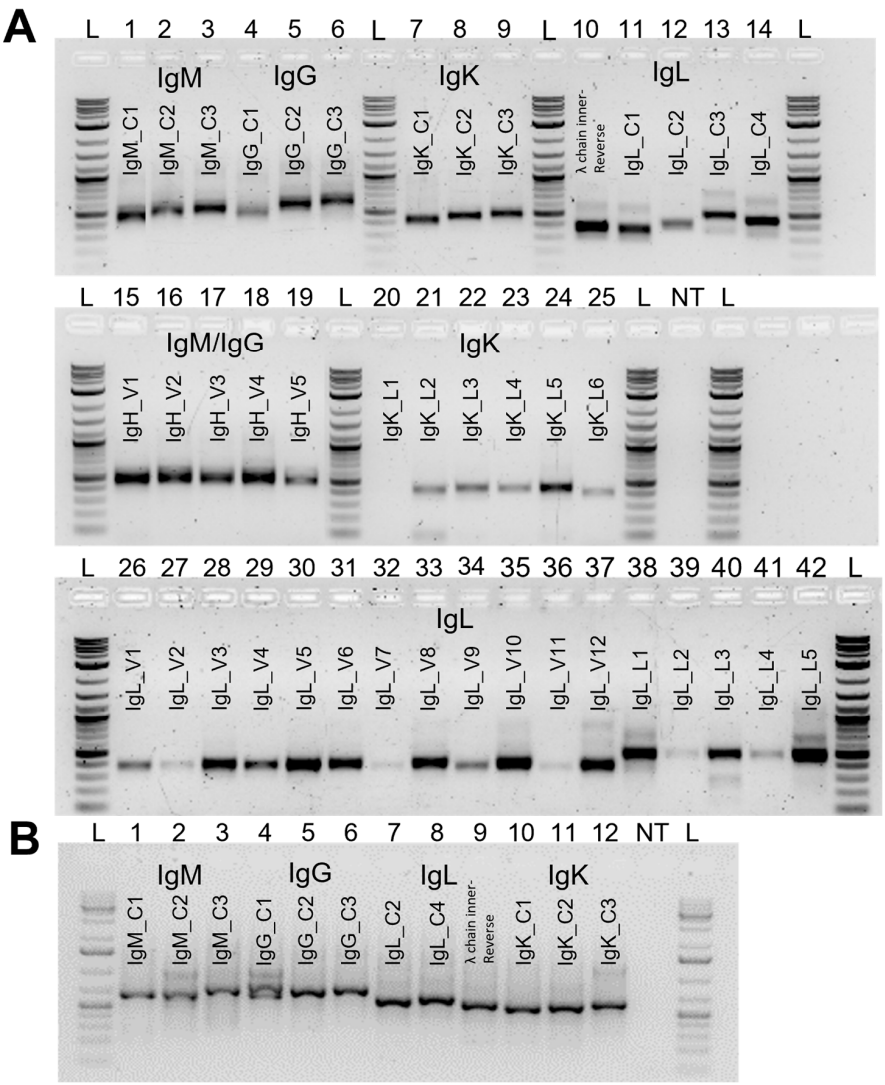

Figure 3. PCR primer validation. (**A**) Cattle cDNA amplification using heavy and light antibody chain primers. The first row shows amplicons generated using the set of constant region primers from the

heavy IgM (1–3: IgM_C1-3) and IgG (4–6: IgG_C1-3), and the light IgK (7–9: IgK_C1-C3) and IgL (10: Ig λ chain inner-reverse [31], 11–14: IgL_C1-4) chains paired with IgH_V4 (1–6), IgK_L3 (7–9), and IgL_V6 (10–12, 14) or IgL_L1 (13), respectively. The second row shows amplicons from individual heavy-chain VHset primers paired with IgM_C2 (15–19), and the kappa chain individual VKset primers paired to IgK_C2 (20–25). The third row shows individual lambda light chains primers from the VL12set (26–37) paired with IgL_C2 and the individual VL5set primers (38–42) paired to IgL_C4. All amplicons are between 400 bp and 600 bp. (**B**) The 10x barcoded cDNA amplification using Read1 forward and cattle constant region reverse primers. The Read1 forward primer [57] was paired with several reverse primers to select suitable pairs for IgM (1–3: IgM_C1-3), IgG (4–6: IgG_C1-3), IgL (7: IgL_C2; 8: IgL_C4; 9: Ig λ chain inner-reverse [31]) and IgK (10–12: IgK_C1-3). NT: no template control; L: NEB 1 Kb Plus ladder.

3.3. PCR Strategy and Sequence Analysis

After the flow cytometric sorting of B cells, we explored three antibody chain amplification strategies (Figure 1). The sequencing strategy determines the ultimate throughput of the antibody sequencing, while upstream isolation is more flexible and adaptable to the project needs (See graphical abstract).

3.3.1. Strategy 1: Antibody Amplification and Sequencing in a Low-Throughput Workflow

The first strategy for antibody VH and VL amplification and sequencing was based on a single-step multiplexed RT-PCR, in which cDNA synthesis was followed by simultaneous amplification of the heavy (IgM or IgG) and light (IgL) chains. IgK, representing only 5% of cattle light chains [20], was not included due to multiplexing difficulties and a general lack of expression. A semi-nested PCR (PCR 2) was subsequently performed with constant region inner primers. Next, gel-purified amplicons of wells exhibiting PCR bands for both H and L chains were cloned into a pGEMTeasy vector, Sanger-sequenced, and analyzed using IgBLASTn and IgMAT [51]. We PCR-amplified H and L chains from 39.3% and 48.8% of the wells (Figure 4A and Figure S1A, Tables S4A and S6A). Out of 84 sorted cells we were able to PCR an amplicon of the expected size for both H and L chains from 26 wells, of which 24 (28.6%) consisted of a single H and L chain after sequencing (Figure 4A and Figure S1A, Table S5A). The few wells that produced more than one H or L chain likely contained more than one cell from the original sorting or potentially a level of biallelism [36,85]. We were only able to amplify an unpaired H or L chain in 7 and 15 wells, respectively. Overall, 42.9% of the wells produced no product, which could be a consequence of many factors, including (i) the sorting parameters and gating strategy resulting in the absence of cells, the sorting of non-B cells, or cells falling on the walls of the wells; (ii) a primer design that did not account for all SNPs (which we consider unlikely due to the primer design strategy); (iii) sample processing that might have degraded the RNA, such as the time between sorting and freezing; or (iv) the number of PCR cycles did not permit the amplification of rare transcripts.

3.3.2. Strategy 2: Antibody Amplification and Sequencing in a Medium-Throughput Workflow

For our second approach, we combined a semi-nested PCR with the Illumina 16S Metagenomics library preparation protocol that ensures unique combinatorial dual indexing of amplicons for subsequent NGS sequencing and demultiplexing [65]. We also included the IgK chain but excluded the V_{K_L2} primer, which was inhibiting the IgK PCR efficiency (Figure S2). We synthesized cDNA in the first step followed by four independent PCR 1 runs for IgM, IgG, IgL, and IgK chains. Primers in the semi-nested PCR ensured that the V_H and V_L amplicons were of similar sizes to avoid NGS sequencing bias for smaller amplicons. They encompassed Illumina adapters used as anchors for the indexing PCR. H and L amplicons from the same well were indexed with identical index primers to facilitate demultiplexing. The final amplicons were loaded on a Tapestation (Agilent Technologies, Inc., Santa Clara, CA, USA) and wells yielding a band at 600 bp were pooled at equimolar concentrations for Illumina sequencing. Due to similarities between ampli-

cons, 40% of PhiX was spiked in to increase the library diversity. Sequences were then analyzed using a custom pipeline and IgMAT. Out of the 96 sorted cells, we were able to amplify and sequence 31 H-L antibody pairs (32.3%); 27 wells (28.1%) did not produce any amplicon, while we obtained more than one heavy, light, or both chains in 8, 7, and 3 of the wells, respectively (18.8%). In total, we amplified 64.6% of the heavy chains and 58.3% of the light chains (Figure 4A, Tables S5B and S6A, and Figure S1B). Altogether, these ranges of H-L chain recovery are comparable for both low- (28.6%) and medium- (32.3%) throughput methods.

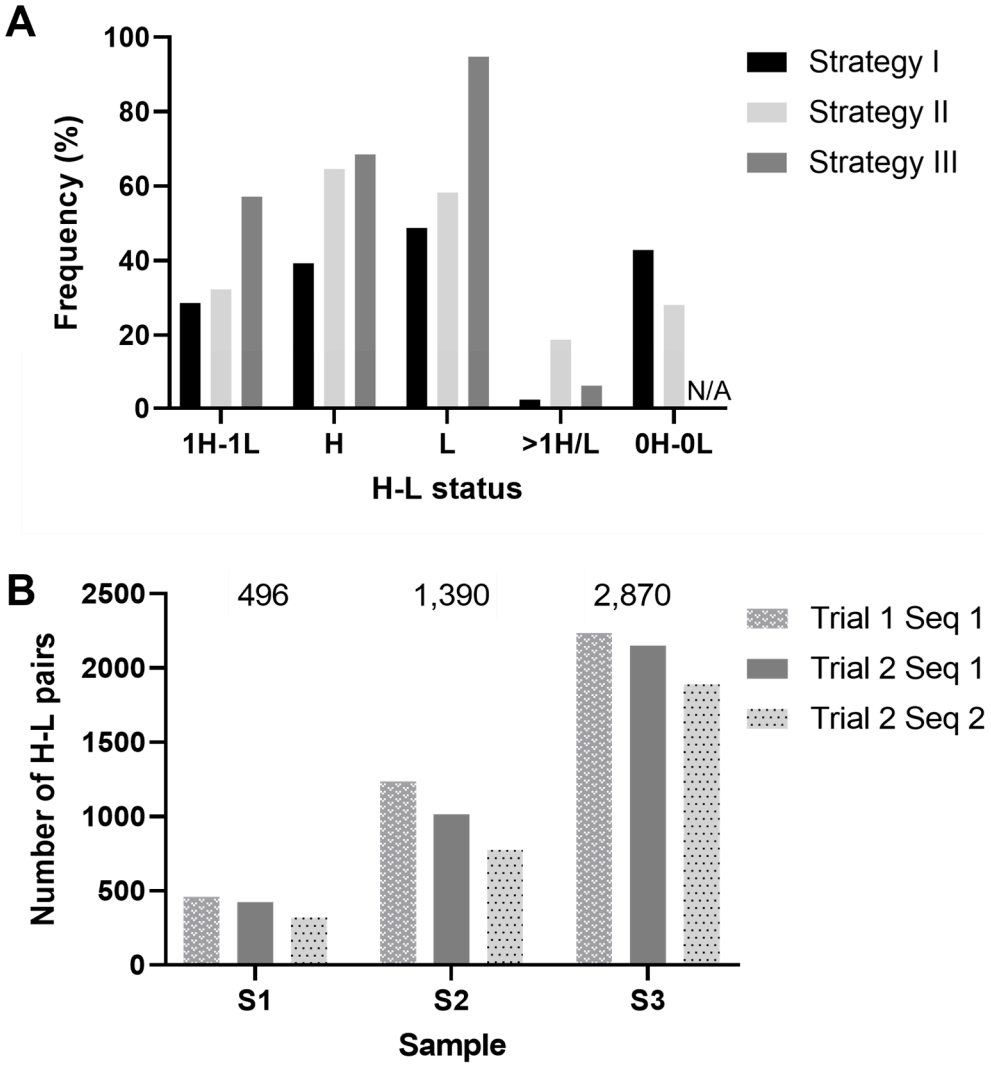

Figure 4. Proportions of sequenced heavy (H) and light (L) chains. (**A**) Frequencies of H/L chain statuses for the three strategies. Percentages of each status are indicated as frequencies of wells processed (strategies 1 and 2) or estimated numbers of cells (strategy 3): 1H-1L: single H and single L pairs; H: overall H chains amplified (including multiplets); L: overall L chains amplified (inc. multiplets); >1H/L: wells or 10x GEMs with more than one H or L chain; 0H-0L: no amplification of either chain. (**B**) Within strategy 3, the numbers of antibody pairs recovered for sample 1 (S1), sample 2 (S2), and sample 3 (S3) are shown. The pairs of single H and single L chains obtained for the two trials (trial 1 and 2) and two sequencing runs (Seq 1 and Seq 2) are shown. Total numbers of antibodies recovered for each sample are indicated above each sample's set of bars.

A similar medium-throughput approach using a dual indexing strategy can also be used to bulk sequence whole repertoires of circulating B cells by indexing samples from distinct animals, time points, and tissues (as opposed to indexing single cells). This would

allow the tracking of antibody sequences from samples taken at different points during infection or vaccination studies [80] from several individuals and in multiple tissues, all in one sequencing run. With the possibility to pool up to 384 samples (considering the number of currently available Illumina indexes), the limiting factor is the sequencing depth.

3.3.3. Strategy 3: Antibody Amplification and Sequencing in a High-Throughput Workflow

To further increase the throughput of the paired H-L chain sequencing, we adapted the 10x Genomics workflow [57] for antibody discovery to cattle by using droplet-based partitioning with the 10x Chromium Controller combined with the amplification of V_H and V_L. We aimed to recover 900, 1800, and 4500 cells in three samples (1, 2, and 3, respectively). The PCR target enrichment conditions recommended by 10x Genomics were used in trial 1, while trial 2 had additional cycles in both PCRs to evaluate the effect on H-L pair recovery, in line with multiple studies demonstrating that increasing the cycle number improves the overall performance [86,87]. The data were processed using Cell Ranger with default settings. All V(D)J processed single-cell samples exhibited Q30 scores above 90% and largely exceeded the recommended number of reads/samples, with an average of 17,000 reads/cell. The Cell Ranger estimates of sequenced cell numbers were 596 and 663 cells for sample 1, 1639 and 1683 cells for sample 2, and 2966 and 4581 cells for sample 3 of trial 1 and trial 2, respectively. The differences between trial 1 and trial 2's estimated numbers of cells may have been related to the increased amplification of transcripts that pass filtering thresholds or the sequencing depth. As we obtained similar numbers of H-L pairs for the two trials of each sample (Figure 4B and Table S6A–E), we concluded that increasing the PCR cycle number did not increase the recovery of paired antibody H-L chains.

The average ratios of GEMs exhibiting more than one H or L chain ranged from 5.21% for sample 1 to 5.93% for sample 2 and 6.25% for sample 3, and the benefit of increasing the cell numbers per channel was supported by the higher number of antibody sequences recovered (Table S6B). With both trials and sequencing runs combined, we obtained 496 (sample 1), 1390 (sample 2), and 2870 (sample 3) single H-L pairs, for a total of 4756 antibody H-L sequences (Table S5C), dramatically increasing the throughput compared to techniques based on single-cell sorting in 96-well plates, as expected. To our knowledge this is the first experiment using 10x technology to sequence cattle antibodies. The 5′ RACE-based 10x protocol ensures no bias against V gene segment usage. We aligned all V_L reads to the V_L12set of primers used for the low-throughput strategy, confirming that 98% of the lambda chain reads were covered by V_L12set exact sequences, with 99.6% allowing 1 mismatch. This validates the completeness of the V_L12set primers. A similar analysis was not possible for the V_H, V_L5, and V_K sets of primers, as the primer sequence is located on the leader sequence trimmed by Cell Ranger.

3.4. Cattle Antibody Characterization

The IgMAT based annotation of antibody sequences isolated in this study confirmed the expected structure of the framework and complementary determining regions, from which we could clearly identify the CDR3. The overall distribution of CDRH3 lengths was in line with previous studies, with average lengths of 20–24 amino acids for the three methods ([51,88]; Figure 5A, Table S4A–C). Overall, 53 ultralong heavy chain sequences (CDRH3 length $\geq$ 50 amino acids) paired with light chains were amplified at 4.2% (1 antibody), 3.2% (1 antibody), and 1.1% (51 antibodies) for the low-, medium-, and high-throughput approaches, respectively (Tables S7A and S8). These percentages were lower than the previously described 5–10% [89,90], which could be attributed to (i) PCR bias against ultralong antibodies, (ii) NGS sequencing bias against longer amplicons, or (iii) variation in ultralong heavy-chain proportions depending on individuals.

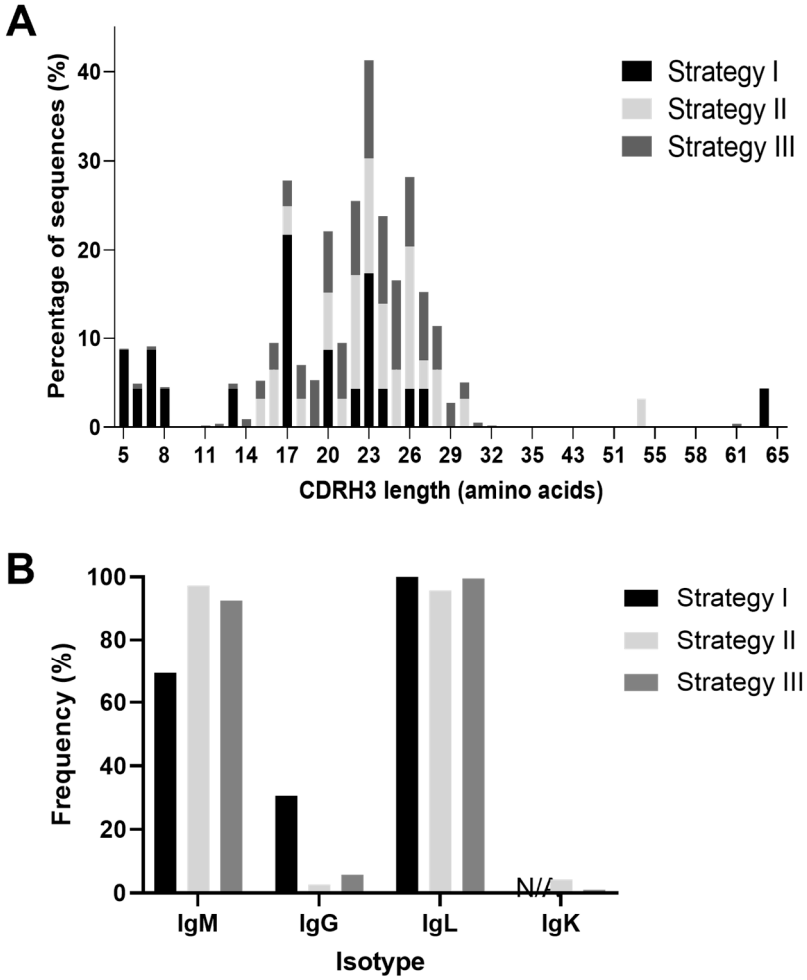

Figure 5. Antibody characteristics. (**A**) Heavy chain complementarity determining region 3 (CDRH3) length distribution from antibodies recovered using the three strategies. Percentages of sequences for each CDRH3 amino acid length are shown. (**B**) H- and L-chain isotypes from all productive chains sequenced by the 3 strategies. Frequencies of heavy-chain IgM and IgG and light-chain lambda (IgL) and kappa (IgK) are shown. IgK frequency was not assessed using strategy 1 (N/A).

Given that cattle VH gene segments share an average of 90% identity [13], the genomic origin of a single antibody sequence is often ambiguous depending on the level of somatic mutation. Therefore, the most probable V gene segment used was assigned to each sequence, and when ambiguous alternative V gene segments were added (Table S4A–C and Figure S3). All VH gene segments belonged to the IGHV1 gene family, and all of the ultralong heavy chains were identified as IGHV1-7 and IGHD8-2, as previously described [16], except one that aligns to the IGHV1-20 gene segment. We also observed restricted pairing of 86.8% (46) of the ultralong heavy chains to light chains using the VL1-47 gene segment, out of which 35 were the so-called invariant light chain CDRL3 previously described (ASAEDSSSNAV; [18]), with another seven paired light chains exhibiting a single amino acid difference to the conserved CDRL3 (Table S7B). Overall, we did not observe significant differences in CDRH3 length distribution (Figure 5A) or gene usage (Figure S3) between the three approaches, implying no obvious bias between the three amplification strategies.

IgM-expressing cells dominated, with IgG comprising only 30.6% of cells processed in strategy 1, 2.7% for strategy 2, and 5.8% for strategy 3 (Figure 5B, Table S8). These data indicate a potentially higher frequency of IgG-secreting cells within the IgL^{lo} ER^{hi} population compared to the overall B cell population, which are likely to be plasma cells and plasmablasts. From the mid-throughput experiment, three light chains were confirmed to be IgK, corresponding to 4.5% of all amplified light chains (including from multiplets),

in line with the described 5% of the repertoire ([20]; Figure 5B and Table S8). However, none were associated with a single H-L pair (2 had multiple heavy, 1 had multiple light chains), which may indicate a lack of function or dysfunctional development of IgK positive B cells. From the 10x experiment, IgK accounted for only 0.5% of the light chains. This low IgK frequency may be due to a suboptimal constant region reverse primer being used (IgKC2 used in the nested PCR of the 10x experiment versus IgKC1 used in mid-throughput strategy) or a difference between individual animals.

The combined analysis of antibody light chains from the three animals revealed that 39.1% (nine chains, out of which seven were distinct), 64.5% (20; 17 distinct), and 15.1% (720; 21 distinct) of the light chains from strategies 1, 2, and 3 were shared between at least two individuals (public chains, Table S9A,B). Out of the distinct chains, three were shared between all three animals. No public heavy chains were observed, consistent with previous reports [36]. Altogether, increasing the throughput of antibody discovery in cattle could shed light on the conflicting issue of heavy and light chain pairing and sharing. In particular, identifying public antibodies among vaccinated or infected animals and convergent CDRH3 could allow the characterization of the antigen-specific antibody response [80].

4. Conclusions

All three workflows successfully isolated natural H-L antibody pairs from single cattle B cells, with no obvious bias between the different approaches, despite using three different animals. This is an advance to enable the study of the antibody response that in turn may inform the optimization or development of new vaccines. The three throughput methods described here have pros and cons when selecting the most suitable strategy for a given aim (Table 2). However, both the single-cell amplification methods and the 10x workflow are very sensitive to cell quality, with the former requiring intact heavy and light chain transcripts and the latter being sensitive to ambient RNA contamination from non-intact cells.

Table 2. Features of the three workflows used in the study.

		Workflow 1	Workflow 2	Workflow 3
Flow cytometry	Staining strategy	Stain PBMCs with fluorescently labeled B cell markers		
	Sorting	Single-cell in plates	Single-cell in plates	Bulk sorting in tubes
	Cost differences	Cost of staining and sorting is equivalent for the 3 strategies		
	Duration	1 day	1 day	1 day
Amplification	cDNA preparation	RT-PCR	cDNA preparation	10x + cDNA preparation + amplification
	PCR amplification	semi-nested PCR	semi-nested PCR	semi-nested PCR
	Purification	Gel purification	AMPure bead purification	AMPure bead purification
	Sample preparation for sequencing	pGEM-T cloning + test PCR	PCR-indexing	Fragmentation + indexing
	Duration	10–12 days	3–4 days	2–3 days
	Cost of reagents *	1750 USD/plate	875 USD/plate	1250 USD/sample
Sequencing	Platform	Sanger	Illumina	Illumina
	Analysis	V-Quest/IgMat	IgMat	Cell Ranger + IgMat
Price/antibody * (excluding flow cytometry and FTE)		80 USD	80 USD	2 USD

* Prices calculated on reagents and sequencing costs only (no FTE) for the experiment detailed in the current study.

When considering the most appropriate method, it is essential to determine the number of H-L pairs required. To equal the throughput of strategy 3, hundreds of plates using either strategy 1 or 2 would need to be processed, with correspondingly high expenditure. If the B cell isolation includes antigen specificity, a few cells may be enough instead of analyzing thousands of cells with feature barcodes to identify the rare population of antigen-specific cells in the high-throughput strategy. The time frame should also be considered since the

cloning or Sanger sequencing approach, although robust, is time-consuming and PCR or NGS is more time-efficient. Another obvious decision-making factor is the availability of specialist equipment such as a 10x device, although equivalent instrument-free approaches are likely to emerge [91]. Alternatively, the intermediate-throughput approach enables several hundred antibodies to be sequenced in one sequencing run. Importantly, the flow cytometry approaches are independent of downstream processing, and any PCR strategy (i.e., multiplexed RT-PCR *versus* singleplex PCR including IgK chains) can interchangeably be Sanger-sequenced or indexed for NGS.

The other key consideration linked to the throughput is the cost. A guide to reagent and consumable costing for each method amplification or sequencing module is provided based on the experiments in this study (Table 2). In brief, the low and medium strategies cost approximately $80 USD per sequenced antibody, while the high-throughput approach costs $2 USD per antibody (Table 2). Although the costings for the low-throughput strategy are fixed, pooling antibodies from multiple plates to sequence 384 antibodies simultaneously can considerably decrease the price/antibody for the medium-throughput approach. We also tested desalt indexing primers and found no difference in antibody sequences compared to Illumina HPLC primers. However, the efficiency was slightly reduced, which was compensated for by two additional cycles in the indexing PCR. Using desalt primers for the medium-throughput approach will have an initial cost but would be cost-effective in the long term.

Overall, the three strategies described here enable successful cattle natural antibody H-L pair sequencing. In particular, 10x protocols were adapted to cattle antibody sequencing for the first time. Implementing these strategies with antigen-specific flow cytometry sorting or feature barcoding before 10x GEM generation would enable the antigen-specific component of the antibody response to vaccination or infection to be identified. Finally, this workflow can be applied to sequence antibodies in other veterinary species.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/vaccines11061099/s1. Figure S1A: Graphic representation of H- and L-chain recovery by amplification and sequencing strategy 1. Figure S1B: Graphic representation of H- and L-chain recovery by amplification and sequencing strategy 2. Figure S2: Testing IgK multiplexing. Figure S3: V_H gene segment usage. Table S1: Markers used for staining bovine B cells in the three strategies. Table S2A: PCR cycling conditions for amplification of bovine antibody amplicons used in strategy 1. Table S2B: PCR cycling conditions for amplification of bovine antibody amplicons used in strategy 2. Table S2C: Conditions for cDNA amplification and PCR cycling used in strategy 3. Table S3: Bioinformatic tools used for the analysis of antibody NGS data. Table S4: Comparative primer specificity in bovine assemblies shown as pairwise identity. Table S5A: Isotype, gene segment usage, CDR3 sequence, and lengths of heavy and light antibody chains obtained from bovine B cells using strategy 1. Table S5B: Isotype, gene segment usage, CDR3 sequence, and lengths of heavy and light chains obtained from bovine B cells using strategy 2. Table S5C: Barcode, gene segment usage, and mismatches to germline, CDR3 sequence and length, and isotype of heavy and light chains obtained from bovine B cells using strategy 3. Table S6A: Recovered antibody heavy and light chains using three different strategies. Table S6B–E: GEM statuses of recovered H and L chains obtained from strategy 3, trials 1 and 2 and sequencing runs 1 and 2. Table S7A: Ultralong antibody heavy chain CDR3 and paired CDRL3 from pairs recovered by the three strategies. Table S7B: CDRL3 sequences associated with ultralong antibodies and their frequencies. Table S8: Isotypes and ultralong chain distribution from all sequenced chains versus H-L pairs obtained by the three strategies. Table S9A: Percentages of public light chains. Table S9B Frequency of CDRL3 sequences shared between the animals used in the three strategies.

Author Contributions: Conceptualization, J.A.H., M.B.-D.P., L.J.M., B.C. and G.M.W.; methodology, K.R.V., B.N., J.N., J.C.S., G.F., W.M., M.B.-D.P. and J.A.H.; software, D.D.-R. and M.J.; validation, K.R.V., B.N. and A.W.S.; formal Analysis, D.D.-R., M.B.-D.P., M.J., W.M., K.R.V., J.C.S. and B.N.; investigation, K.R.V., B.N., J.N. and M.B.-D.P.; resources, W.M. and G.F.; data curation, D.D.-R. and M.J.; writing—original draft preparation, K.R.V., B.N., D.D.-R. and M.B.-D.P.; writing—review and editing, K.R.V., B.N., B.C., L.J.M., M.B.-D.P., W.M. and J.A.H.; visualization, K.R.V., D.D.-R., M.J. and M.B.-D.P.; supervision, M.B.-D.P., L.J.M. and J.A.H.; project administration, M.B.-D.P.; funding Acquisition, J.A.H. and L.J.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by funding from the Bill and Melinda Gates Foundation (grants OPP1192002 and OPP1215550) and the UK Research and Innovation—Biotechnology and Biological Sciences Research Council (UKRI-BBSRC) (grants BBS/E/I/00001410, BBS/E/I/00007030, BBS/E/I/00007031, BBS/E/I/00007037, BBS/E/I/00007038, and BBS/E/I/00007039). The Roslin Institute was supported through core funding from a UKRI-BBSRC award (BBS/E/D/20002173). The PhD studentship of Benjamin Nzau was jointly funded by The Pirbright Institute and the Royal (Dick) School of Veterinary Studies.

Institutional Review Board Statement: Animal blood sampling was approved by the ethical review process at the Pirbright Institute and the Centre for Diary Research of the University of Reading (CEDAR) or the Animal and Plant Agency (APHA) and was conducted following the UK Government Animal (Scientific Procedures) Act 1986 under project license (PPL 7998958), conforming to the ARRIVE Guidelines.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available as BioProject accession PRJNA962094.

Acknowledgments: We acknowledge the Pirbright Science Technology Platforms' support through the Immunological Toolbox for providing some of the antibodies used in this study, the Flow Cytometry unit for their assistance in the sorting experiments, and the Bioinformatics, Sequencing, and Proteomics unit for performing the next-generation sequencing.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ahmed, S.F.; Quadeer, A.A.; McKay, M.R. Preliminary Identification of Potential Vaccine Targets for the COVID-19 Coronavirus (SARS-CoV-2) Based on SARS-CoV Immunological Studies. *Viruses* **2020**, *12*, 254. [CrossRef]
2. Bekker, L.-G.; Tatoud, R.; Dabis, F.; Feinberg, M.; Kaleebu, P.; Marovich, M.; Ndung'u, T.; Russell, N.; Johnson, J.; Luba, M.; et al. The complex challenges of HIV vaccine development require renewed and expanded global commitment. *Lancet* **2020**, *395*, 5. [CrossRef]
3. Hatherill, M.; White, R.G.; Hawn, T.R. Clinical Development of New TB Vaccines: Recent Advances and Next Steps. *Front. Microbiol.* **2019**, *10*, 12. [CrossRef]
4. Matz, K.M.; Marzi, A.; Feldmann, H. Ebola vaccine trials: Progress in vaccine safety and immunogenicity. *Expert. Rev. Vaccines* **2019**, *18*, 14. [CrossRef] [PubMed]
5. Antos, A.; Miroslaw, P.; Rola, J.; Polak, M.P. Vaccination Failure in Eradication and Control Programs for Bovine Viral Diarrhea Infection. *Front. Vet. Sci.* **2021**, *8*, 11. [CrossRef]
6. Belsham, G.J. Towards improvements in foot-and-mouth disease vaccine performance. *Acta Vet. Scand.* **2020**, *62*, 20. [CrossRef] [PubMed]
7. Black, S.J.; Mansfield, J.M. Prospects for vaccination against pathogenic African trypanosomes. *Parasite Immunol.* **2016**, *38*, 9. [CrossRef]
8. Guzman, E.; Taylor, G. Immunology of bovine respiratory syncytial virus in calves. *Mol. Immunol.* **2015**, *66*, 9. [CrossRef] [PubMed]
9. Cancro, M.P.; Tomayko, M.M. Memory B cells and plasma cells: The differentiative continuum of humoral immunity. *Immunol. Rev.* **2021**, *303*, 11. [CrossRef]
10. Stanfield, R.L.; Berndsen, Z.T.; Huang, R.; Sok, D.; Warner, G.; Torres, J.L.; Burton, D.R.; Ward, A.B.; Wilson, I.A.; Smider, V.V. Structural basis of broad HIV neutralization by a vaccine-induced cow antibody. *Sci. Adv.* **2020**, *6*, 12. [CrossRef]
11. Tsuchiya, Y.; Mizuguchi, K. The diversity of H3 loops determines the antigen-binding tendencies of antibody CDR loops. *Protein Sci.* **2016**, *25*, 11. [CrossRef]
12. Xu, J.L.; Davis, M.M. Diversity in the CDR3 Region of VH Is Sufficient for Most Antibody Specificities. *Immunity* **2000**, *13*, 9. [CrossRef]

13. Ma, L.; Qin, T.; Chu, D.; Cheng, X.; Wang, J.; Wang, X.; Wang, P.; Han, H.; Ren, L.; Aitken, R.; et al. Internal Duplications of DH, JH, and C Region Genes Create an Unusual IgH Gene Locus in Cattle. *J. Immunol.* **2016**, *196*, 9. [CrossRef] [PubMed]
14. Zhao, Y.; Jackson, S.M.; Aitken, R. The bovine antibody repertoire. *Dev. Comp. Immunol.* **2006**, *30*, 12. [CrossRef] [PubMed]
15. Oyola, S.O.; Henson, S.P.; Nzau, B.; Kibwana, E.; Nene, V. Access to ultra-long IgG CDRH3 bovine antibody sequences using short read sequencing technology. *Mol. Immunol.* **2021**, *139*, 9. [CrossRef] [PubMed]
16. Deiss, T.C.; Vadnais, M.; Wang, F.; Chen, P.L.; Torkamani, A.; Mwangi, W.; Lefranc, M.-P.; Criscitiello, M.F.; Smider, V.V. Immunogenetic factors driving formation of ultralong VH CDR3 in Bos taurus antibodies. *Cell. Mol. Immunol.* **2019**, *16*, 12. [CrossRef]
17. Smider, B.A.; Smider, V.V. Formation of ultralong DH regions through genomic rearrangement. *BMC Immunol.* **2020**, *21*, 30. [CrossRef]
18. Wang, F.; Ekiert, D.C.; Ahmad, I.; Yu, W.; Zhang, Y.; Bazirgan, O.; Torkamani, A.; Raudsepp, T.; Mwangi, W.; Criscitiello, M.F.; et al. Reshaping antibody diversity. *Cell* **2013**, *153*, 15. [CrossRef]
19. Burke, M.J.; Stockley, P.G.; Boyes, J. Broadly Neutralizing Bovine Antibodies: Highly Effective New Tools against Evasive Pathogens? *Viruses* **2020**, *12*, 473. [CrossRef]
20. Arun, S.S.; Breuer, W.; Hermanns, W. Immunohistochemical Examination of Light-chain Expression (Λ/k Ratio) in Canine, Feline, Equine, Bovine and Porcine Plasma Cells. *J. Vet. Med. Ser. A* **1996**, *43*, 4. [CrossRef]
21. Barroso, R.; Morrison, W.I.; Morrison, L.J. Molecular Dissection of the Antibody Response: Opportunities and Needs for Application in Cattle. *Front. Immunol.* **2020**, *11*, 10. [CrossRef]
22. Köhler, G.; Milstein, C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **1975**, *256*, 3. [CrossRef]
23. Steinitz, M.; Klein, G.; Koskimies, S.; Makel, O. EB virus-induced B lymphocyte cell lines producing specific antibody. *Nature* **1977**, *269*, 3. [CrossRef]
24. Smith, S.A.; Crowe, J.E., Jr. Use of Human Hybridoma Technology To Isolate Human Monoclonal Antibodies. *Microbiol. Spectr.* **2015**, *3*, 12. [CrossRef]
25. Crain, M.J.; Sanders, S.K.; Butler, J.L.; Cooper, M.D. Epstein-Barr virus preferentially induces proliferation of primed B cells. *J. Immunol.* **1989**, *143*, 6. [CrossRef]
26. Borrebaeck, C.A.; Danielsson, L.; Möller, S.A. Human monoclonal antibodies produced by primary in vitro immunization of peripheral blood lymphocytes. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 5. [CrossRef] [PubMed]
27. Groves, D.J.; Tucker, E.M. The production and application of non-rodent monoclonal antibodies in veterinary science. *Vet. Immunol. Immunopathol.* **1989**, *23*, 14. [CrossRef] [PubMed]
28. Clackson, T.; Hoogenboom, H.R.; Griffiths, A.D.; Winter, G. Making antibody fragments using phage display libraries. *Nature* **1991**, *352*, 5. [CrossRef] [PubMed]
29. Tiller, T.; Meffre, E.; Yurasov, S.; Tsuiji, M.; Nussenzweig, M.C.; Wardemann, H. Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. *J. Immunol. Methods* **2008**, *329*, 13. [CrossRef]
30. Neu, K.E.; Guthmiller, J.J.; Huang, M.; La, J.; Vieira, M.C.; Kim, K.; Zheng, N.Y.; Cortese, M.; Tepora, M.E.; Hamel, N.J.; et al. Spec-seq unveils transcriptional subpopulations of antibody-secreting cells following influenza vaccination. *J. Clin. Investig.* **2019**, *129*, 13. [CrossRef]
31. Li, K.; Wang, S.; Cao, Y.; Bao, H.; Li, P.; Sun, P.; Bai, X.; Fu, Y.; Ma, X.; Zhang, J.; et al. Development of Foot-and-Mouth Disease Virus-Neutralizing Monoclonal Antibodies Derived From Plasmablasts of Infected Cattle and Their Germline Gene Usage. *Front. Immunol.* **2020**, *10*, 18. [CrossRef]
32. Pedrioli, A.; Oxenius, A. Single B cell technologies for monoclonal antibody discovery. *Trends Immunol.* **2021**, *42*, 16. [CrossRef] [PubMed]
33. Hou, X.L.; Wang, L.; Ding, Y.L.; Xie, Q.; Diao, H.Y. Current status and recent advances of next generation sequencing techniques in immunological repertoire. *Genes Immun.* **2016**, *17*, 12. [CrossRef] [PubMed]
34. Goldstein, L.D.; Chen, Y.J.; Wu, J.; Chaudhuri, S.; Hsiao, Y.C.; Schneider, K.; Hoi, K.H.; Lin, Z.; Guerrero, S.; Jaiswal, B.S.; et al. Massively parallel single-cell B-cell receptor sequencing enables rapid discovery of diverse antigen-reactive antibodies. *Commun. Biol.* **2019**, *2*, 10. [CrossRef]
35. McDaniel, J.R.; DeKosky, B.J.; Tanno, H.; Ellington, A.D.; Georgiou, G. Ultra-high-throughput sequencing of the immune receptor repertoire from millions of lymphocytes. *Nat. Protoc.* **2016**, *11*, 14. [CrossRef]
36. DeKosky, B.J.; Kojima, T.; Rodin, A.; Charab, W.; Ippolito, G.C.; Ellington, A.D.; Georgiou, G. In-depth determination and analysis of the human paired heavy- and light-chain antibody repertoire. *Nat. Med.* **2015**, *21*, 6. [CrossRef] [PubMed]
37. DeKosky, B.J.; Ippolito, G.C.; Deschner, R.P.; Lavinder, J.J.; Wine, Y.; Rawlings, B.M.; Varadarajan, N.; Giesecke, C.; Dörner, T.; Andrews, S.F.; et al. High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. *Nat. Biotechnol.* **2013**, *31*, 4. [CrossRef]
38. Saikia, M.; Burnham, P.; Keshavjee, S.H.; Wang, M.F.Z.; Heyang, M.; Moral-Lopez, P.; Hinchman, M.M.; Danko, C.G.; Parker, J.S.L.; De Vlaminck, I. Simultaneous multiplexed amplicon sequencing and transcriptome profiling in single cells. *Nat. Methods* **2019**, *16*, 4. [CrossRef]

39. Wang, B.; DeKosky, B.J.; Timm, M.R.; Lee, J.; Normandin, E.; Misasi, J.; Kong, R.; McDaniel, J.R.; Delidakis, G.; Leigh, K.E.; et al. Functional interrogation and mining of natively paired human V(H):V(L) antibody repertoires. *Nat. Biotechnol.* **2018**, *36*, 4. [CrossRef]
40. Rajan, S.; Kierny, M.R.; Mercer, A.; Wu, J.; Tovchigrechko, A.; Wu, H.; Dall Acqua, W.F.; Xiao, X.; Chowdhury, P.S. Recombinant human B cell repertoires enable screening for rare, specific, and natively paired antibodies. *Commun. Biol.* **2018**, *1*, 5. [CrossRef]
41. DeKosky, B.J.; Lungu, O.I.; Park, D.; Johnson, E.L.; Charab, W.; Chrysostomou, C.; Kuroda, D.; Ellington, A.D.; Ippolito, G.C.; Gray, J.J.; et al. Large-scale sequence and structural comparisons of human naive and antigen-experienced antibody repertoires. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 10. [CrossRef]
42. Asensio, M.A.; Lim, Y.W.; Wayham, N.; Stadtmiller, K.; Edgar, R.C.; Leong, J.; Leong, R.; Mizrahi, R.A.; Adams, M.S.; Simons, J.F.; et al. Antibody repertoire analysis of mouse immunization protocols using microfluidics and molecular genomics. *MAbs* **2019**, *11*, 14. [CrossRef] [PubMed]
43. Adler, A.S.; Mizrahi, R.A.; Spindler, M.J.; Adams, M.S.; Asensio, M.A.; Edgar, R.C.; Leong, J.; Leong, R.; Johnson, D.S. Rare, high-affinity mouse anti-PD-1 antibodies that function in checkpoint blockade, discovered using microfluidics and molecular genomics. *MAbs* **2017**, *9*, 12. [CrossRef]
44. Adler, A.S.; Bedinger, D.; Adams, M.S.; Asensio, M.A.; Edgar, R.C.; Leong, R.; Leong, J.; Mizrahi, R.A.; Spindler, M.J.; Bandi, S.R.; et al. A natively paired antibody library yields drug leads with higher sensitivity and specificity than a randomly paired antibody library. *MAbs* **2018**, *10*, 12. [CrossRef]
45. Jayaram, N.; Bhowmick, P.; Martin, A.C. Germline VH/VL pairing in antibodies. *Protein Eng. Des. Sel.* **2012**, *25*, 7. [CrossRef] [PubMed]
46. Ren, J.; Nettleship, J.E.; Harris, G.; Mwangi, W.; Rhaman, N.; Grant, C.; Kotecha, A.; Fry, E.; Charleston, B.; Stuart, D.I.; et al. The role of the light chain in the structure and binding activity of two cattle antibodies that neutralize bovine respiratory syncytial virus. *Mol. Immunol.* **2019**, *112*, 8. [CrossRef] [PubMed]
47. Curtis, N.C.; Lee, J. Beyond bulk single-chain sequencing: Getting at the whole receptor. *Curr. Opin. Syst. Biol.* **2020**, *24*, 7. [CrossRef]
48. Smakaj, E.; Babrak, L.; Ohlin, M.; Shugay, M.; Briney, B.; Tosoni, D.; Galli, C.; Grobelsek, V.; D'Angelo, I.; Olson, B.; et al. Benchmarking immunoinformatic tools for the analysis of antibody repertoire sequences. *Bioinformatics* **2019**, *36*, 9. [CrossRef]
49. Ehling, R.A.; Weber, C.R.; Mason, D.M.; Friedensohn, S.; Wagner, B.; Bieberich, F.; Kapetanovic, E.; Vazquez-Lombardi, R.; Di Roberto, R.l.B.; Hong, K.-L.; et al. SARS-CoV-2 reactive and neutralizing antibodies discovered by single-cell sequencing of plasma cells and mammalian display. *Cell. Rep.* **2022**, *38*, 21. [CrossRef]
50. Qi, H.; Liu, B.; Wang, X.; Zhang, L. The humoral response and antibodies against SARS-CoV-2 infection. *Nat. Immunol.* **2022**, *23*, 13. [CrossRef]
51. Dorey-Robinson, D.; Maccari, G.; Borne, R.; Hammond, J.A. IgMAT: Immunoglobulin sequence multi-species annotation tool for any species including those with incomplete antibody annotation or unusual characteristics. *bioRxiv* **2021**. [CrossRef]
52. Mwangi, W.; Maccari, G.; Hope, J.C.; Entrican, G.; Hammond, J.A. The UK Veterinary Immunological Toolbox Website: Promoting vaccine research by facilitating communication and removing reagent barriers. *Immunology* **2020**, *161*, 3. [CrossRef]
53. Roos, E.O.; Bonnet-Di Placido, M.; Mwangi, W.N.; Moffat, K.; Fry, L.M.; Waters, R.; Hammond, J.A. OMIP-085: Cattle B-cell phenotyping by an 8-color panel. *Cytom. Part. A* **2023**, *103*, 4. [CrossRef]
54. Wang, L.; Madera, R.; Li, Y.; Gladue, D.P.; Borca, M.V.; McIntosh, M.T.; Shi, J. Development of Porcine Monoclonal Antibodies with In Vitro Neutralizing Activity against Classical Swine Fever Virus from C-Strain E2-Specific Single B Cells. *Viruses* **2023**, *15*, 863. [CrossRef] [PubMed]
55. Grant, C.F.; Lefevre, E.A.; Carr, B.V.; Prentice, H.; Gubbins, S.; Pollard, A.J.; Charreyre, C.; Charleston, B. Assessment of T-dependent and T-independent immune responses in cattle using a B cell ELISPOT assay. *Vet. Res.* **2012**, *43*, 68. [CrossRef]
56. Holzer, B.; Rijal, P.; McNee, A.; Paudyal, B.; Martini, V.; Clark, B.; Manjegowda, T.; Salguero, F.J.; Bessell, E.; Schwartz, J.C.; et al. Protective porcine influenza virus-specific monoclonal antibodies recognize similar haemagglutinin epitopes as humans. *PLoS Pathog.* **2021**, *17*, 25. [CrossRef]
57. 10x_Genomics. Chromium Next GEM Single Cell V(D)J Reagent Kits v1.1 User Guide. Available online: https://support.10xgenomics.com/single-cell-vdj/library-prep/doc/user-guide-chromium-single-cell-vdj-reagent-kits-v11-chemistry (accessed on 12 October 2020).
58. Elsik, C.G.; Tellam, R.L.; Worley, K.C.; Gibbs, R.A.; Muzny, D.M.; Weinstock, G.M.; Adelson, D.L.; Eichler, E.E.; Elnitski, L.; Guigó, R.; et al. The genome sequence of taurine cattle: A window to ruminant biology and evolution. *Science* **2009**, *324*, 7. [CrossRef]
59. Zimin, A.V.; Delcher, A.L.; Florea, L.; Kelley, D.R.; Schatz, M.C.; Puiu, D.; Hanrahan, F.; Pertea, G.; Van Tassell, C.P.; Sonstegard, T.S.; et al. A whole-genome assembly of the domestic cow, Bos taurus. *Genome Biol.* **2009**, *10*, 10. [CrossRef]
60. Schwartz, J.C.; Philp, R.L.; Bickhart, D.M.; Smith, T.P.L.; Hammond, J.A. The antibody loci of the domestic goat (*Capra hircus*). *Immunogenetics* **2018**, *70*, 10. [CrossRef]
61. Chevreux, B.; Wetter, T.; Suhai, S. Genome Sequence Assembly Using Trace Signals and Additional Sequence Information. *Comput. Sci. Biol. Proc. Ger. Conf. Bioinform. (GCB)* **1999**, *99*, 12.
62. Bonfield, J.K.; Whitwham, A. Gap5—Editing the billion fragment sequence assembly. *Bioinformatics* **2010**, *26*, 5. [CrossRef]
63. Rutherford, K.; Parkhill, J.; Crook, J.; Horsnell, T.; Rice, P.; Rajandream, M.-A.; Barrell, B. Artemis: Sequence visualization and annotation. *Bioinformatics* **2000**, *16*, 2. [CrossRef]

64. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 8. [CrossRef] [PubMed]
65. Klindworth, A.; Pruesse, E.; Schweer, T.; Peplies, J.; Quast, C.; Horn, M.; Glöckner, F.O. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **2013**, *41*, 11. [CrossRef] [PubMed]
66. Illumina. Illumina Adapter Sequences. Available online: https://support-docs.illumina.com/SHARE/AdapterSeq/illumina-adapter-sequences.pdf (accessed on 19 April 2021).
67. Messing, J. New M13 vectors for cloning. *Methods Enzym.* **1983**, *101*, 59. [CrossRef]
68. Illumina. 16S Metagenomic Sequencing Library Preparation. Preparing 16S Ribosomal RNA Gene Amplicons for the Illumina MiSeq System. Available online: https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf (accessed on 9 March 2020).
69. Potapov, V.; Ong, J.L. Examining Sources of Error in PCR by Single-Molecule Sequencing. *PLoS ONE* **2017**, *12*, 19. [CrossRef] [PubMed]
70. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 5. [CrossRef] [PubMed]
71. Lefranc, M.-P.; Giudicelli, V.; Duroux, P.; Jabado-Michaloud, J.; Folch, G.; Aouinti, S.; Carillon, E.; Duvergey, H.; Houles, A.; Paysan-Lafosse, T.; et al. IMGT®, the international ImMunoGeneTics information system® 25 years on. *Nucleic Acids Res.* **2014**, *43*, 10. [CrossRef]
72. Magoč, T.; Salzberg, S.L. FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **2011**, *27*, 7. [CrossRef]
73. Edgar, R.C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **2010**, *26*, 2. [CrossRef]
74. Jourdan, M.; Caraux, A.; De Vos, J.; Fiol, G.; Larroque, M.; Cognot, C.; Bret, C.; Duperray, C.; Hose, D.; Klein, B. An in vitro model of differentiation of memory B cells into plasmablasts and plasma cells including detailed phenotypic and molecular characterization. *Blood* **2009**, *114*, 9. [CrossRef]
75. Shaffer, A.L.; Shapiro-Shelef, M.; Iwakoshi, N.N.; Lee, A.H.; Qian, S.B.; Zhao, H.; Yu, X.; Yang, L.; Tan, B.K.; Rosenwald, A.; et al. XBP1, downstream of Blimp-1, expands the secretory apparatus and other organelles, and increases protein synthesis in plasma cell differentiation. *Immunity* **2004**, *21*, 13. [CrossRef] [PubMed]
76. Kurosawa, N.; Yoshioka, M.; Fujimoto, R.; Yamagishi, F.; Isobe, M. Rapid production of antigen-specific monoclonal antibodies from a variety of animals. *BMC Biol.* **2012**, *10*, 14. [CrossRef]
77. Diks, A.M.; Overduin, L.A.; van Leenen, L.D.; Slobbe, L.; Jolink, H.; Visser, L.G.; van Dongen, J.J.M.; Berkowska, M.A. B-Cell Immunophenotyping to Predict Vaccination Outcome in the Immunocompromised—A Systematic Review. *Front. Immunol.* **2021**, *12*, 21. [CrossRef]
78. Lee, B.O.; Moyron-Quiroz, J.; Rangel-Moreno, J.; Kusser, K.L.; Hartson, L.; Sprague, F.; Lund, F.E.; Randall, T.D. CD40, but Not CD154, Expression on B Cells Is Necessary for Optimal Primary B Cell Responses. *J. Immunol.* **2003**, *171*, 5707–5717. [CrossRef]
79. Frölich, D.; Giesecke, C.; Mei, H.E.; Reiter, K.; Daridon, C.; Lipsky, P.E.; Dörner, T. Secondary immunization generates clonally related antigen-specific plasma cells and memory B cells. *J. Immunol.* **2010**, *185*, 8. [CrossRef] [PubMed]
80. Galson, J.D.; Pollard, A.J.; Trück, J.; Kelly, D.F. Studying the antibody repertoire after vaccination: Practical applications. *Trends Immunol.* **2014**, *35*, 13. [CrossRef]
81. Jiang, N.; He, J.; Weinstein, J.A.; Penland, L.; Sasaki, S.; He, X.-S.; Dekker, C.L.; Zheng, N.-Y.; Huang, M.; Sullivan, M.; et al. Lineage structure of the human antibody repertoire in response to influenza vaccination. *Sci. Transl. Med.* **2013**, *5*, 17. [CrossRef] [PubMed]
82. Laserson, U.; Vigneault, F.; Gadala-Maria, D.; Yaari, G.; Uduman, M.; Vander Heiden, J.A.; Kelton, W.; Taek Jung, S.; Liu, Y.; Laserson, J.; et al. High-resolution antibody dynamics of vaccine-induced immune responses. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 6. [CrossRef]
83. Newman, J.; Rice, J.S.; Wang, C.; Harris, S.L.; Diamond, B. Identification of an antigen-specific B cell population. *J. Immunol. Methods* **2003**, *272*, 11. [CrossRef]
84. Illumina. How Short Inserts Affect Sequencing Performance. Available online: https://emea.support.illumina.com/bulletins/2020/12/how-short-inserts-affect-sequencing-performance.html (accessed on 12 October 2021).
85. Shi, Z.; Zhang, Q.; Yan, H.; Yang, Y.; Wang, P.; Zhang, Y.; Deng, Z.; Yu, M.; Zhou, W.; Wang, Q.; et al. More than one antibody of individual B cells revealed by single-cell immune profiling. *Cell. Discov.* **2019**, *5*, 13. [CrossRef]
86. Singh, M.; Al-Eryani, G.; Carswell, S.; Ferguson, J.M.; Blackburn, J.; Barton, K.; Roden, D.; Luciani, F.; Giang Phan, T.; Junankar, S.; et al. High-throughput targeted long-read single cell sequencing reveals the clonal and transcriptional landscape of lymphocytes. *Nat. Commun.* **2019**, *10*, 13. [CrossRef]
87. Setliff, I.; Shiakolas, A.R.; Pilewski, K.A.; Murji, A.A.; Mapengo, R.E.; Janowska, K.; Richardson, S.; Oosthuysen, C.; Raju, N.; Ronsard, L.; et al. High-Throughput Mapping of B Cell Receptor Sequences to Antigen Specificity. *Cell* **2019**, *179*, 11. [CrossRef] [PubMed]
88. Berens, S.J.; Wylie, D.E.; Lopez, O.J. Use of a single VH family and long CDR3s in the variable region of cattle Ig heavy chains. *Int. Immunol.* **1997**, *9*, 11. [CrossRef]

89. Dong, J.; Finn, J.A.; Larsen, P.A.; Smith, T.P.L.; Crowe, J.E., Jr. Structural Diversity of Ultralong CDRH3s in Seven Bovine Antibody Heavy Chains. *Front. Immunol.* **2019**, *10*, 9. [CrossRef]
90. Shojaei, F.; Saini, S.S.; Kaushik, A.K. Unusually long germline DH genes contribute to large sized CDR3H in bovine antibodies. *Mol. Immunol.* **2003**, *40*, 7. [CrossRef] [PubMed]
91. Komatsu, J.; Cico, A.; Poncin, R.; Le Bohec, M.; Morf, J.; Lipin, S.; Graindorge, A.; Eckert, H.; Saffarian, A.; Cathaly, L.; et al. RevGel-seq: Instrument-free single-cell RNA sequencing using a reversible hydrogel for cell-specific barcoding. *Sci. Rep.* **2023**, *13*, 11. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.