

Chapter 3

In-Depth Determination and Analysis of the Human Paired Heavy and Light Chain Antibody Repertoire

3.1 Introduction

The determination of immune receptor repertoires using high throughput (NextGen) DNA sequencing has rapidly become an indispensable tool for the understanding of adaptive immunity, antibody discovery and in clinical practice [1–3]. However, because the variable domains of antibody heavy and light chains (VH and VL, respectively) are encoded by different mRNA transcripts, until recently it was only possible to determine the VH and VL repertoires separately, or else paired VH:VL sequences for small to moderate numbers of B cells (10^4 – 10^5) [4], far smaller than the ~ 0.7 – 4×10^6 B cells contained in a typical 10 ml blood draw. Thus a technology for the facile determination of the paired antibody VH:VL repertoire at great depth (i.e. $>10^6$ cells per analysis) and for a variety of B cell subsets is still needed for clinical research [5], antibody discovery [6, 7], and for addressing a host of important questions related to the shaping of the antibody repertoire [2, 8–13].

Several techniques have been reported for detection or sequencing of genomic DNA or cDNA from single cells; however all are limited by low efficiency or low cell throughput (<200 – 500 cells) and require fabrication and operation of complicated microfluidic devices [14–17]. Chudakov and coworkers recently reported the use of one-pot cell encapsulation within water-in-oil emulsions, cell lysis by heating at 65°C concomitant with TCR α and β reverse transcription and finally, linking by overlap extension PCR to determine TCR α :TCR β pairings, albeit only for TCR β V7 and with a very low efficiency (approximately 700 TCR α :TCR β pairs recovered from 8×10^6 PBMC) [17]. This is likely because one-pot emulsions have a high

The whole content of this chapter was initially published in DeKosky, B.J. et al. In-depth determination and analysis of the human paired heavy- and light-chain antibody repertoire. *Nat. Med.* 21, 89–91 (2014), and it is reproduced here. B.J.D. and G.G. developed the methodology and wrote the manuscript; B.J.D., T.K., G.C.I., A.D.E. and G.G. designed the experiments; B.J.D., T.K., A.R. and W.C. performed the experiments; B.J.D. carried out the bioinformatic analysis; and B.J.D. and T.K. analyzed the data.

degree of droplet size dispersity and since the RT reaction is inhibited in volumes <5 nL [15] only the small fraction of cells encapsulated within larger droplets yields cDNA for further manipulation.

Inspired by methods for the production of highly monodisperse polymeric microspheres for drug delivery purposes [18, 19], we developed a new technology that enables sequencing of the paired VH:VL repertoire from millions of B cells within a few hours of experimental effort and using equipment that can be built inexpensively by any laboratory. For validation we expanded *in vitro* memory B cells isolated from human PBMCs to obtain a sample that contained multiple clones of individual B cells and showed that among aliquots (technical replicates) the accuracy of VH:VL pairing is $>97\%$. We show that ultra-high throughput determination of the paired VH:VL repertoire provides important immunological insights such as: (i) the discovery of human light chains detected in multiple individuals that pair with a wide range of VH genes, (ii) the quantitative analysis of allelic inclusion in humans, i.e. of B cells expressing two different antibodies, and (iii) estimates of the frequencies of antibodies in healthy human repertoires that display known features of broadly neutralizing antibodies to rapidly evolving pathogens.

3.2 Results

3.2.1 Device Construction

For facile high-throughput single-cell manipulation we assembled a simple axisymmetric flow-focusing device comprising three concentric tubes: an inner needle carrying cells suspended in PBS, a middle tube carrying a lysis solution and magnetic poly(dT) beads for mRNA capture from lysed cells, and finally an external tube with a rapidly flowing annular oil phase, all of which passed through a $140\ \mu\text{m}$ glass nozzle (Fig. 3.1a). The rapidly flowing outer annular oil phase focused the slower-moving aqueous phase into a thin, unstable jet that coalesced into droplets with a predictable size distribution; additionally maintaining laminar flow regime within the apparatus prevented mixing of cells and lysis solution prior to droplet formation (Fig. B.1).

To evaluate cell encapsulation and droplet size distribution, MOPC-21 immortalized B cells suspended in PBS were injected through the inner needle at a rate of 250,000 cells/min while a solution of PBS containing the cell viability dye Trypan blue (0.4% v/v) was injected through the middle tubing so that dye mixed with cells at the point of droplet formation. Resulting emulsion droplets were $73 \pm 20\ \mu\text{m}$ in diameter (average \pm SD). Trypan blue exclusion revealed that, as expected, cells remained viable throughout the emulsification process (Fig. B.2). Replacing the Trypan blue stream with cell lysis buffer containing lithium dodecyl sulfate (LiDS) and DTT to inactivate RNases resulted in complete cell lysis as indicated by visual disappearance of cell membranes from emulsion droplets.

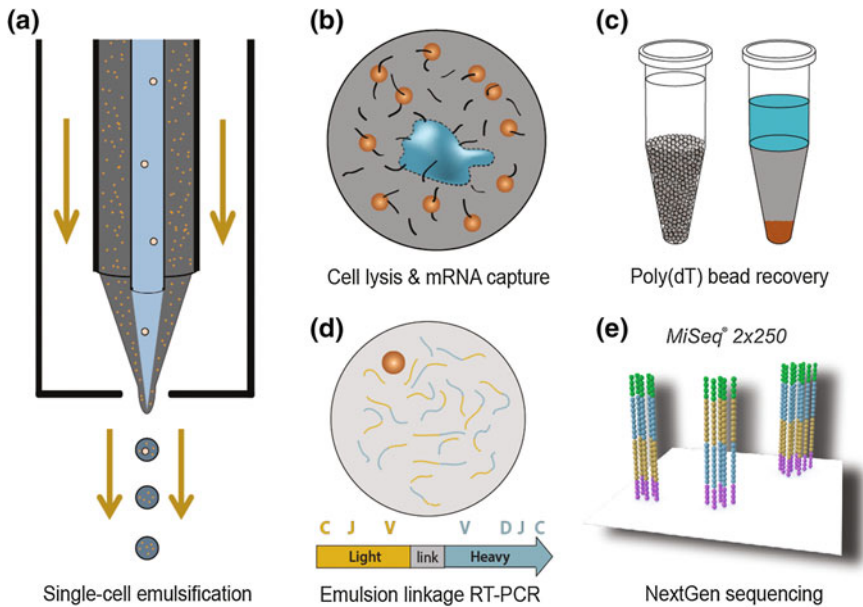


Fig. 3.1 Technical workflow for ultra-high throughput VH:VL sequencing from single B cells. **a** An axisymmetric flow-focusing nozzle isolated single cells and poly(dT) magnetic beads into emulsions of predictable size distributions. An aqueous solution of cells in PBS (center, blue/pink circles) and cell lysis buffer with poly(dT) beads (gray/orange circles) exited an inner and outer needle and were surrounded by a rapidly moving annular oil phase (orange arrows). Aqueous streams focused into a thin jet which coalesced into emulsion droplets of predictable sizes, and cells mixed with lysis buffer only at the point of droplet formation (Fig. B.1). **b** Single cell VH and VL mRNAs annealed to poly(dT) beads within emulsion droplets (blue figure represents a lysed cell, orange circles depict magnetic beads, black lines depict mRNA strands). **c** poly(dT) beads with annealed mRNA were recovered by emulsion centrifugation to concentrate aqueous phase (left) followed by diethyl ether destabilization (right). **d** Recovered beads were emulsified for cDNA synthesis and linkage PCR to generate an ~850-base pair VH:VL cDNA product. **e** Next-generation sequencing of VH:VL amplicons was used to analyze the native heavy and light chain repertoire of input B cells

3.2.2 Single B Cell VH:VL Pairing: Throughput and Pairing Accuracy

Human CD3⁻CD19⁺CD20⁺CD27⁺ memory B cells were isolated from PBMCs from a healthy volunteer and expanded for four days *in vitro* by stimulation with anti-CD40 antibody, IL-4, IL-10, IL-21, and CpG oligodeoxynucleotides [20]. *In vitro* expansion was performed to create a cell population containing a sufficient number of clonal B cells so that the concordance of the VH:VL repertoire in two technical replicates could be assessed. 1,600,000 *in vitro* expanded B cells were divided into two aliquots and passed through the flow-focusing nozzle at a rate of 50,000 cells/min (i.e. 16 min emulsification for each replicate) and processed as

shown in Fig. 3.1. The emulsion of lysed single cells with compartmentalized poly (dT) beads was maintained for three minutes at room temperature to allow specific mRNA hybridization onto poly(dT) magnetic beads (Fig. 3.1b), then the emulsion was broken chemically (Fig. 3.1c), beads were re-emulsified, and overlap extension RT-PCR was performed to generate linked VH:VL amplicons (Fig. 3.1d). The resulting cDNAs were amplified by nested PCR to generate an ~ 850 bp VH:VL product for NextGen sequencing by Illumina MiSeq 2×250 or 2×300 . Due to read length limitations of current NextGen sequencing technologies, the FRH4-(CDR-H3)-FRH3:FRL3-(CDR-L3)-FRL4 was sequenced first to reveal the pairing of the VH and VL hypervariable loops. Each of these VH:VL pairs may also comprise one or more somatic variants containing mutations within the upstream portion of the VH and VL genes. We determine the complete set of somatic variants by separate MiSeq sequencing the VH and VL portions of the paired 850 bp VH:VL amplicon followed by in silico gene assembly [4].

Sequence data were processed by read quality filtering, CDR-H3 clustering, VH:VL pairing, and selection for paired VH:VLs with ≥ 2 reads in the dataset. The clustering step resulted in high-confidence sequence data but with a lower-bound estimate of clonal diversity because clonally expanded or somatically mutated B cells with similar VH sequences collapse into a single CDR-H3 cluster. 129,097 VH:VL clusters were observed after separate analysis and clustering of Replicates 1 and 2. Of these, 37,995 CDR-H3 sequences were observed in both replicates (and hence must have originated from expanded B cells present in both technical replicates) with 36,468 paired with the same CDR-L3 across replicates revealing a VH:VL pairing precision of 98.0% (Fig. 3.2, Table 3.1, Fig. B.3, see Methods). The ratio of VH:VL clusters to input cells observed (typically between 1:10 and 1:15) is a reflection of the clonality of the memory B cell population (i.e. presence of clonally related memory B cells), clustering threshold, RT-PCR efficiency and cell viability. For

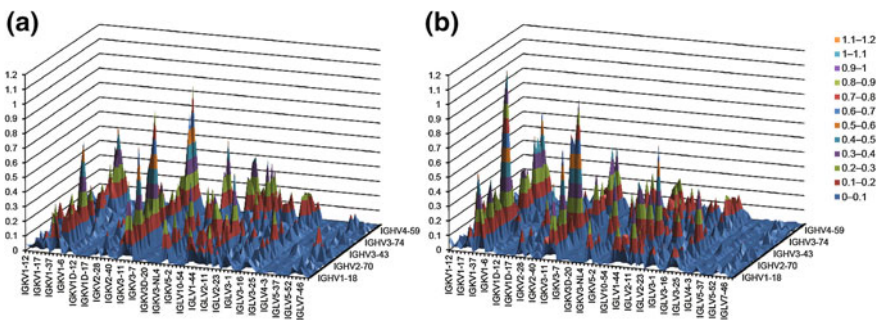


Fig. 3.2 Heavy:light V-gene pairing landscape of $CD3^+CD19^+CD20^+CD27^+$ peripheral memory B cells in two healthy human donors. V genes are plotted in alphanumeric order; height indicates percentage representation among VH:VL clusters. **a** Donor 1 ($n = 129,097$). **b** Donor 2 ($n = 53,679$). VH:VL gene usage was highly correlated between Donors 1 and 2 (Spearman rank correlation coefficient 0.757, $p < 1 \times 10^{-99}$). Additional heat maps are provided in Figs. B.3 and B.4

Table 3.1 High-throughput VH:VL sequence analysis of CD3⁻CD19⁺CD20⁺CD27⁺ in vitro-expanded human B cells

Human donor	V-region primer set	# cells analyzed	Emulsification rate (cells/min)	Observed VH:VL clusters	CDR-H3 detected in both replicates	CDR-H3: CDR-L3 clusters detected in both replicates	VH:VL pairing precision (%)
Donor 1	Framework 1	1,600,000	50,000	129,097	37,995	36,468	98.0
Donor 2	Framework 1	810,000	50,000	53,679	19,096	18,115	97.4
Donor 3	Leader peptide	210,000	33,000	15,372	4267	4170	98.9

comparison, in our hands, sequencing the memory B cell VH repertoire by preparing amplicons directly by standard RT-PCR without pairing and using the same bioinformatic filters (sequences present at ≥ 2 reads, 96% clustering) resulted in a 1:6 ratio of VH clusters:input cells, which compares favorably to the yield of paired VH:VL clusters in Table 3.1. Two additional pairing analyses of somewhat smaller B cell populations from different donors were also performed (Table 3.1, Figs. B.4 and B.5). In a separate experiment designed to verify native VH:VL pairing accuracy, plasmids encoding 11 different known human antibodies were transfected separately into HEK293 cells. Aliquots containing comparable numbers of each of the transfected cells were mixed and processed as described in Fig. 3.1, and native pairings were identified for 11/11 antibodies (Table B.1). In yet another test, approximately 260 ARH-77 immortalized human B cells [4] were mixed with 20,000 CD3⁻CD19⁺CD20⁺CD27⁺ expanded memory B cells (~ 100 -fold excess). ARH-77 heavy and light chains were paired correctly and the ratio of correctly paired ARH-77 VH:VL reads over the top correctVH:incorrectVL, a parameter that we denote signal:topVLnoise, was 96.4:1 (2604 correct ARH-77 VH:VL reads vs. 27 reads for the top ARH-77 VH paired with an incorrect VL, Table B.2).

As discussed above, three sequencing reactions and in silico assembly are needed to determine the sequence of the complete linked VH:VL amplicon with Illumina MiSeq 2×250 or 2×300 . Alternatively, the long-read Pacific Biosciences (PacBio) sequencing platform can be used to obtain the complete ~ 850 bp cDNA encoding linked VH:VL sequences. However, because of its substantially lower throughput and higher cost per read, we find that despite the need for three distinct MiSeq samples compared to only one for PacBio, the former is currently much more cost-effective for deep repertoire analyses. We found PacBio sequencing to be preferable only for certain specialized applications, for example in identifying VH:VL pairs in antibodies with extensive SHM such as broadly neutralizing antibodies that arise following persistent infection with rapidly evolving viruses, most notably HIV-1. For example, we used PacBio to sequence 15,000 VH:VL amplicons from elite controller CAP256 [7] and identified six variants of VRC26-class HIV broadly neutralizing antibodies within the VH:VL repertoire (Figs. B.6 and B.7).

3.2.3 *Promiscuous and Public VL Junctions*

In contrast to the heavy chain, light chain rearrangements do not incorporate a diversity segment and exhibit restricted CDR-L3 lengths with low levels of N-addition. Light chains therefore have a much lower theoretical diversity than heavy chains and the presence of light chain sequences paired with multiple heavy chains within a single donor, referred to as “repeated” or “promiscuous” light chains, is an expected result, especially for VL junctions that are mostly germline encoded and also derive from V- and J-genes with high prevalence in human immune repertoires [21]. However the separate high-throughput sequencing of VH and VL repertoires, as has been practiced until now, cannot provide VH pairing information for a given VL and thus precludes identification and characterization of promiscuous light chains [22, 23]. We observed thousands of heavy chains paired with promiscuous VL nucleotide junctions (34.9, 29.4 and 19.6% of all heavy chains were paired with promiscuous VL junctions in Donors 1, 2, and 3, respectively). We inspected high-frequency promiscuous light chains to see if any promiscuous VL might be shared across individuals (i.e. a “public” VL). We found that highly promiscuous VLs were nearly always public: for example, of the 50 highest-frequency promiscuous VL junctions in Donor 1, 49/50 were also detected in Donors 2 and 3. Promiscuous light chains showed an average of 0.04 non-templated bases in the VL junction compared to an average of 5 non-templated bases in non-promiscuous light chains (i.e. VLs that paired with a single VH in a donor, see Fig. B.8, $p < 10^{-10}$). The lack of non-templated bases in promiscuous VL junctions indicated that promiscuity can be observed mainly in germline-encoded VL genes lacking SHM.

We examined in detail two representative promiscuous and public VL junctions that contained V- and J-genes with high prevalence in steady-state human immune repertoires (*KVI-39:KJ2*, 9 aa CDR-L3, *LVI-44:LJ3*, 11 aa CDR-L3, both observed at a frequency of ~ 1 per 1000 VH:VL clusters) [24, 25] to check for biases in VH pairing of promiscuous VL chains. *KVI-39:KJ2* and *LVI-44:LJ3* both paired with VH genes of diverse germline lineage and CDR-H3 length that reflected the overall VH gene usage in the repertoire (Fig. 3.3, Spearman rank correlation coefficients: *KVI-39:KJ2* $\rho = 0.889$, $p < 10^{-21}$; *LVI-44:LJ3* $\rho = 0.847$, $p < 10^{-17}$), indicating that VL nucleotide-sequence promiscuity arises mostly from distinct VL recombination events rather than B cell activation and subsequent clonal expansion. We note that no two donors shared more than 2 VH nucleotide sequences, and no VH sequence was detected in all three donors, consistent with previous reports which showed that in contrast to VL junctions, the VH nucleotide repertoire is highly private [2, 25].

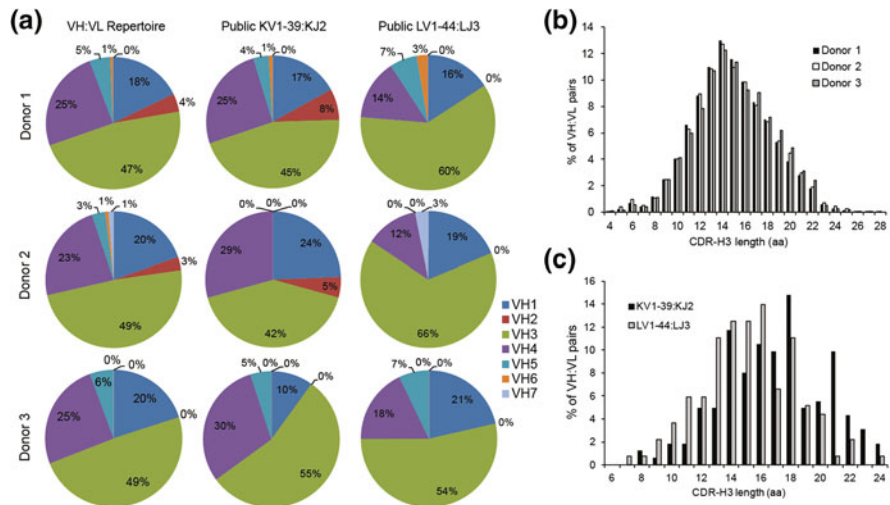


Fig. 3.3 **a** VH gene family utilization in: *left* total paired VH:VL repertoires (Donor 1 $n = 129,097$, Donor 2 $n = 53,679$, Donor 3 $n = 15,372$), *center* heavy chains paired with a representative highly-ranked public and promiscuous VL observed in all three donors (*KV1-39:KJ2* 9 aa CDR-L3, *tgtcaacagatgtacagtaccctgtacactttt*; Donor 1 $n = 106$, Donor 2 $n = 41$, Donor 3 $n = 20$), *right* heavy chains paired with a different highly-ranked public VL in all three donors (*LV1-44:LJ3* 11 aa CDR-L3, *tgtgcagcatgggatgacagcctgaatggttgggtgttc*; $n = 76$, $n = 32$ and $n = 28$, respectively). **b** CDR-H3 length distribution in VH:VL repertoires (Donor 1 $n = 129,097$, Donor 2 $n = 53,679$, Donor 3 $n = 15,372$). **c** CDR-H3 length distribution for all antibodies containing the two representative public VL chains from part (**a**)

3.2.4 Quantifying Allelic Inclusion in Human Memory B Cells

Clonal selection theory postulates that each lymphocyte expresses one antibody. However, studies in mice have confirmed that this is not always the case. Allelic inclusion, the phenomenon whereby one B cell expresses two BCRs, overwhelmingly one VH gene with two different VLs, has been well-documented in mice and has been proposed to be particularly important in autoimmunity because the expression of a second BCR can dilute a pre-existing auto-reactive BCR and limit the expansion of autoreactive B cells. Similarly allelic inclusion can also provide a mechanism for autoreactive antibodies to evade central tolerance [26–30]. Almost 20 years ago A. Lanzavecchia and coworkers used FACS sorting of cells expressing both κ and λ immunoglobulin proteins on their cell surface ($sIg\kappa^+/sIg\lambda^+$, denoting surface-expression of both $Ig\kappa$ and $Ig\lambda$) followed by EBV immortalization to show that $sIg\kappa^+/sIg\lambda^+$ allelic inclusion occurs in 0.2–0.5% of human memory B cells [31]. However, the inability to sort dual $sIg\kappa^+$ and dual $sIg\lambda^+$ human B cells and the absence of methods for the determination of the VH:VL repertoire at sufficient depth (since the frequency of allelic inclusion is low) have precluded

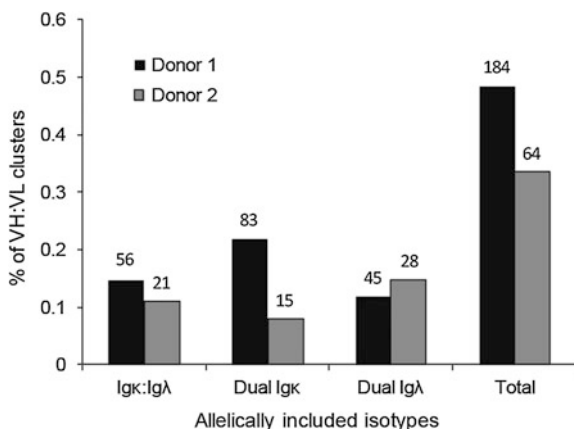


Fig. 3.4 Frequency of VL transcript allelic inclusion in two donors ($n = 184$ and $n = 64$ allelically included antibodies from $n = 37,995$ and $n = 19,096$ VH:VL clusters detected across replicates in Donor 1 and Donor 2, respectively). 14 allelically included antibodies were detected in Donor 3 (8 dual κ/λ , 2 dual κ/κ , 2 dual λ/λ , $n = 4,267$ VH:VL clusters detected across replicates). Numbers above each category indicate the absolute number of observed allelically included antibodies

more comprehensive determination of allelic inclusion in humans. We detected VL allelic inclusion at a rate of approximately 0.4% of VH clusters for Donor 1 and Donor 2, with dual κ/λ -transcribing B cells in approximately equal proportions to dual κ/κ - and λ/λ -transcribing B-cell clones (Fig. 3.4). These heavy chains paired only with their two allelically included light chains (exact nucleotide match) in two technical replicates, and we observed that approximately 80% of these antibodies displayed somatic mutations. The somatic mutation frequency detected in allelically included VH:VL pairs was comparable to previous reports by Lanzavecchia et al. for allelically included $sIg\kappa^+/sIg\lambda^+$ cells (3/5 EBV-immortalized clones [31]). Also consistent with the earlier study, we observed stop codons resulting from somatic mutation that inactivated a subset of allelically included VL transcripts [31]. For the $\sim 20\%$ of allelically included VH that do not display SHM, we cannot rule out the possibility that these clones were derived from pre-B expansion.

3.2.5 Antibodies with Gene Signatures of Known Anti-Viral BNAbs

High-resolution sequence descriptions of the immune repertoire can inform on B cell trajectories for the emergence of broadly neutralizing antibodies (bNAbs) to rapidly evolving pathogens [7, 10, 32, 33]. Many bNAbs display highly unusual features including very long CDR-H3 and short CDR-L3 sequences [7, 32, 34, 35],

and these properties have raised the question as to whether antibodies with similar features are normally found in the repertoire of healthy donors and thus could evolve following stimulation by infection or vaccination to yield neutralizing antibodies. We found approximately 1:6000 VH:VL clusters exhibited general characteristics of known VRC01-class anti-HIV antibodies (22, 9, and 0 for Donors 1, 2, and 3 respectively; germline *VH1-02*, a very short ≤ 5 aa CDR-L3, and CDR-H3 length between 11 and 18 aa [35]), while antibodies with genetic characteristics of anti-influenza FI6 occurred in approximately $2-5 \times 10^4$ memory B cells (6 and 1 antibodies detected in Donors 1 and 2, respectively; *VH3-30*, *KV4-1*, 22aa CDR-H3, 9aa CDR-L3 [32]).

3.3 Discussion

We have developed an easy to implement, ultra high-throughput technology for sequencing the VH:VL repertoire at relatively low cost and with high pairing accuracy. The workflow presented here permits sequence analysis of the entire population of human B cells contained in a 10 mL blood draw, or if needed, even in a unit of blood (450 ml) in a single-day experiment, an improvement orders of magnitude relative to what is feasible using robotic single-cell RT-PCR [36]. As many as 6 million B cells (or alternatively, as few as 1000 B cells) can be analyzed per operator in a single day. Of note, the number of antibody sequences reported here ($\sim 200,000$) dwarfs the entire set of $<19,000$ human VH:VL sequences that had been deposited in the International Nucleotide Sequence Database Collaboration (INSDC) over the past 25 years (in addition to the ~ 5000 human VH:VL pairs we reported previously [4]).

The determination of the paired antibody repertoire at great depth can provide unprecedented insights on a number of medically and immunologically important issues. For example, we used HT single-cell VH:VL sequencing to detect highly-utilized promiscuous and germline-encoded VL junctions that are observed in multiple donors, to identify antibodies with bNAb-like features in HIV-1 patients [7] (Figs. B.6 and B.7) and to quantify the frequency of bNAb-like V gene rearrangements in healthy donors, as described above. The latter is an important factor in determining whether a vaccine immunogen might be able to elicit protective immunity [34, 35]. High-throughput VH:VL sequencing can also be used to search for public antibody VH:VL clonotypes [37, 38] and to identify antibodies having specific features determined by computational or structural biology analyses or with relevance to pathogen neutralization [7, 35, 39–41]. In autoimmunity, high-throughput VH:VL sequencing can reveal an individual's repertoire of allelically included B cells (Fig. 3.4) and the presence of B cell clones expressing antibodies containing hallmark autoimmune signatures (with respect to paratope net charge, CDR-H3 and CDR-L3 lengths, etc.) as well as other attributes of potential diagnostic and therapeutic utility [27, 29, 30].

3.4 Methods

Methods and associated references are reported in a published version of this thesis chapter [42].

References

1. Warren EH, Matsen FA, Chou J (2013) High-throughput sequencing of B- and T-lymphocyte antigen receptors in hematology. *Blood* 122:19–22
2. Georgiou G et al (2014) The promise and challenge of high-throughput sequencing of the antibody repertoire. *Nat Biotechnol* 32:158–168
3. Logan AC et al (2011) High-throughput VDJ sequencing for quantification of minimal residual disease in chronic lymphocytic leukemia and immune reconstitution assessment. *Proc Natl Acad Sci USA* 108:21194–21199
4. DeKosky BJ et al (2013) High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. *Nat Biotech* 31:166–169
5. Sasaki S et al (2011) Limited efficacy of inactivated influenza vaccine in elderly individuals is associated with decreased production of vaccine-specific antibodies. *J Clin Invest* 121:3109–3119
6. Smith K et al (2009) Rapid generation of fully human monoclonal antibodies specific to a vaccinating antigen. *Nat Protoc* 4:372–384
7. Doria-Rose NA et al (2014) Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. *Nat* 509:55–62
8. Wilson PC, Andrews SF (2012) Tools to therapeutically harness the human antibody response. *Nat Rev Immunol* 12:709–719
9. Fischer N (2011) Sequencing antibody repertoires: the next generation. *MAbs* 3:17–20
10. Wu X et al (2011) Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. *Sci* 333:1593–1602
11. Finn JA, Crowe JE Jr (2013) Impact of new sequencing technologies on studies of the human B cell repertoire. *Curr Opin Immunol* 25:613–618
12. Finco O, Rappuoli R (2014) Designing vaccines for the Twenty-First Century society. *Front Immunol* 5
13. Newell EW, Davis MM (2014) Beyond model antigens: high-dimensional methods for the analysis of antigen-specific T cells. *Nat Biotechnol* 32:149–157
14. Marcus JS, Anderson WF, Quake SR (2006) Microfluidic Single-Cell mRNA isolation and analysis. *Anal Chem* 78:3084–3089
15. White AK et al (2011) High-throughput microfluidic single-cell RT-qPCR. *Proc Natl Acad Sci USA* 108:13999–14004
16. Furutani S, Nagai H, Takamura Y, Aoyama Y, Kubo I (2012) Detection of expressed gene in isolated single cells in microchambers by a novel hot cell-direct RT-PCR method. *Analyst* 137:2951–2957
17. Turchaninova MA et al (2013) Pairing of T-cell receptor chains via emulsion PCR. *Eur J Immunol* 43:2507–2515
18. Berkland C, Kim KK, Pack DW (2001) Fabrication of PLG microspheres with precisely controlled and monodisperse size distributions. *J Controlled Release* 73:59–74
19. Berkland C, Pollauf E, Pack DW, Kim K (2004) Uniform double-walled polymer microspheres of controllable shell thickness. *J Controlled Release* 96:101–111
20. Recher M et al (2011) IL-21 is the primary common γ chain-binding cytokine required for human B-cell differentiation in vivo. *Blood* 118:6824–6835

21. Jackson KJL, Kidd MJ, Wang Y, Collins AM (2013) The shape of the lymphocyte receptor repertoire: lessons from the B cell receptor. *Front Immunol* 4:1–12
22. Jackson KJL et al (2012) Divergent human populations show extensive shared IGK rearrangements in peripheral blood B cells. *Immunogenet* 64:3–14
23. Hoi KH, Ippolito GC (2013) Intrinsic bias and public rearrangements in the human immunoglobulin V[lambda] light chain repertoire. *Genes Immun* 14:271–276
24. Ippolito GC et al (2012) Antibody repertoires in humanized NOD-scid-IL2R gamma(null) mice and human B cells reveals human-like diversification and tolerance checkpoints in the mouse. *PLoS ONE* 7:e35497
25. Glanville J et al (2011) Naive antibody gene-segment frequencies are heritable and unaltered by chronic lymphocyte ablation. *Proc Natl Acad Sci USA* 108:20066–20071
26. Pelanda R (2014) Dual immunoglobulin light chain B cells: trojan horses of autoimmunity? *Curr Opin Immunol* 27:53–59
27. Liu S et al (2005) Receptor editing can lead to allelic inclusion and development of B cells that retain antibodies reacting with high avidity autoantigens. *J Immunol* 175:5067–5076
28. Rezanka LJ, Kenny JJ, Longo DL (2005) Dual isotype expressing B cells [$\kappa(+)/\lambda(+)$] arise during the ontogeny of B cells in the bone marrow of normal nontransgenic mice. *Cell Immunol* 238:38–48
29. Casellas R et al (2007) Igk allelic inclusion is a consequence of receptor editing. *J Exp Med* 204:153–160
30. Andrews SF et al (2013) Global analysis of B cell selection using an immunoglobulin light chain-mediated model of autoreactivity. *J Exp Med* 210:125–142
31. Giachino C, Padovan E, Lanzavecchia A (1995) kappa+ lambda+ dual receptor B cells are present in the human peripheral repertoire. *J Exp Med* 181:1245–1250
32. Corti D et al (2011) A neutralizing antibody selected from plasma cells that binds to group 1 and group 2 influenza A hemagglutinins. *Sci* 333:850–856
33. Wrammert J et al (2011) Broadly cross-reactive antibodies dominate the human B cell response against 2009 pandemic H1N1 influenza virus infection. *J Exp Med* 208:181–193
34. Jardine J et al (2013) Rational HIV immunogen design to target specific germline B cell receptors. *Sci* 340:711–716
35. Zhou T et al (2013) Multidonor analysis reveals structural elements, genetic determinants, and maturation pathway for HIV-1 neutralization by VRC01-Class antibodies. *Immunol* 39:245–258
36. Busse CE, Czogiel I, Braun P, Arndt PF, Wardemann H (2014) Single-cell based high-throughput sequencing of full-length immunoglobulin heavy and light chain genes. *Eur J Immunol* 44:597–603
37. Parameswaran P et al (2013) Convergent antibody signatures in human dengue. *Cell Host Microbe* 13:691–700
38. Jackson KJL et al (2014) Human responses to influenza vaccination show seroconversion signatures and convergent antibody rearrangements. *Cell Host Microbe* 16:105–114
39. Wine Y et al (2013) Molecular deconvolution of the monoclonal antibodies that comprise the polyclonal serum response. *Proc Natl Acad Sci* 110:2993–2998
40. Lavinder JJ et al (2014) Identification and characterization of the constituent human serum antibodies elicited by vaccination. *Proc Natl Acad Sci* 111:2259–2264
41. Boutz DR et al (2014) Proteomic identification of monoclonal antibodies from serum. *Anal Chem* 86:4758–4766
42. DeKosky BJ et al (2015) In-depth determination and analysis of the human paired heavy- and light-chain antibody repertoire. *Nat Med* 21:86–91