Published in final edited form as:

Exp Mol Pathol. 2013 December ; 95(3): 357–363. doi:10.1016/j.yexmp.2013.10.004.

Deep Sequencing and Circos Analyses of Antibody Libraries Reveal Antigen-driven Selection of Ig V_H Genes during HIV-1 Infection

Madelyne Xiao^{a,d}, Prabakaran Ponraj^{b,c,*}, Weizao Chen^b, Bailey Kessing^{d,*}, and Dimitar S. Dimitrov^b

^aUrbana High School, Ijamsville, Maryland 21754, USA

^bProtein Interactions Group, Center for Cancer Research Cancer and Inflammation Program, Science Applications International Corporation-Frederick, Inc., NCI-Frederick, Frederick, MD 21702, USA

^cBasic Research Program, Science Applications International Corporation-Frederick, Inc., NCI-Frederick, Frederick, MD 21702, USA

^dBSP CCR Genetics Core, Science Applications International Corporation-Frederick, Inc., NCI-Frederick, Frederick, MD 21702, USA

^eWerner H. Kirsten Student Intern Program, National Cancer Institute (NCI)-Frederick, National Institutes of Health (NIH), Frederick, MD 21702-1201, USA

Abstract

The vast diversity of antibody repertoires is largely attributed to heavy chain (V_H) recombination of variable (V), diversity (D) and joining (J) gene segments. We used 454 sequencing information of the variable domains of the antibody heavy chain repertoires from neonates, normal adults and an HIV-1-infected individual, to analyze, with Circos software, the VDJ pairing patterns at birth, adulthood and a time-dependent response to HIV-1 infection. Our comparative analyses of the Ig VDJ repertoires from these libraries indicated that, from birth to adulthood, VDJ recombination patterns remain the same with some slight changes, whereas some V_H families are selected and preferentially expressed after long-term infection with HIV-1. We also demonstrated that the immune system responds to HIV-1 chronic infection by selectively expanding certain HV families in an attempt to combat infection. Our findings may have implications for understanding immune responses in pathology as well as for development of new therapeutics and vaccines.

Keywords

454 sequencing; antibodyome; antibody repertoire; HIV-1; human monoclonal antibody; VDJ analysis

*Corresponding authors. Mailing address Prabakaran Ponraj 1050 Boyles St, NCI-Frederick, Bldg 469, Rm 140 Frederick, MD 21702-1201 Phone: (301) 846-5628 Fax: (301) 846-5598 prabakaran.ponraj@nih.gov.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of Interest statement

The authors declare that there are no conflicts of interest.

INTRODUCTION

The adaptive immune system has the genetic basis for encoding an almost unlimited number of distinct antibodies. These antibodies are generated in response to a large number of antigens encountered in a lifetime. Intriguingly, such a vast diversity of antibody repertoire is generated from only a limited number of subfamily genes. Particularly, 38-44 Ig heavy chain variable (HV) functional genes, 27 diversity (D) genes, and 9 junction (J) genes have been identified in humans (Lefranc et al., 2009). These V, D, and J gene segments are in separate regions of chromosomes, and brought together by recombination to code functional immunoglobulin genes (Agrawal et al., 1998). The diversity of antibody heavy chain repertoire is generated from a limited number of V, D, J genes through two mechanisms. On the central level and in naive mature B cells, VDJ recombination of antibody gene segments and junctional diversity together provide approximately 10^{11} possible distinct antibodies (Lloyd et al., 2009). At the peripheral level, post-exposure to antigens, antibody genes undergo somatic hypermutations and class switches and, which add further diversity to the already large antibody repertoire. However, at any given time, there are approximately 10^7 distinct B cell clones in an individual (Chen et al., 2010). Understanding of how B cell clones develop in response to infection and/or progression of a disease may yield valuable insight into our understanding of the maturation process of the immune system.

Previously, high-throughput pyrosequencing of antibodies from humans (Arnaout et al., 2011) (Arnaout et al., 2011; Boyd et al., 2009b; Glanville et al., 2009; Ravn et al., 2010) and zebra fish (Weinstein et al., 2009) have been used to analyze the B cell diversity including germline gene usage, somatic mutation in germline-encoded complementarity determining regions (CDRs), CDR3 diversity, and VDJ usage. One of the major goals of high-throughput sequencing has been to explore the antibodyome (Dimitrov, 2010), the complete sets of antibodies expressed by different subsets of B cells, e.g. those elicited in response to antigens such as HIV-1 (Wu et al., 2011) or non-HLA antigens generated after renal transplantation (Li et al., 2009). Particularly, deep sequencing of antibodies using the 454 pyrosequencing method is opted for sequencing the entire antibody variable domain from a large library, and allows studies of the VDJ usage in normal individuals (Arnaout et al., 2011; Prabakaran et al., 2012b), human lymphocyte clonality in pathological conditions (Boyd et al., 2009a), and haplotyping (Kidd et al., 2012) as well as in antibody repertoire maturation process in zebrafish (Jiang et al., 2011). All these studies showed that the first step in generating antibody diversity through VDJ recombination comes with enormous variability and with unequal frequencies with different Vs, Ds and Js.

In this study, we used Circos (www.circos.ca), software that excels at displaying relational data, to analyze the deep sequencing results of several antibody combinatorial libraries established in our laboratory. These libraries included naïve Fab IgM libraries constructed from 2 cord blood samples (Prabakaran et al., 2012a) and peripheral blood mononuclear cells (PBMCs) from 69 healthy donors (Chen et al., 2010; Prabakaran et al., 2012b). These two libraries were compared in their Ig HV gene composition and VDJ combination to find any changes associated with growth and maturation of the adaptive immune system. We previously reported Fab IgG libraries which were established from the PBMCs of an acutely HIV-1 infected patient at post infection 40 days and 8 months (Chen et al., 2012). In the present analysis, we included IgM libraries constructed from PBMCs, and additionally both IgG and IgM libraries constructed from bone marrows of the same patient at these two time points. Our analyses indicated that, while there are subtle differences between the cord blood and adult IgM library, the IgG and IgM libraries established from the HIV-1 infected patient samples display striking differences in Ig HV gene expression usage, supporting the notion that selection pressure by antigens during infection and the development of the immune system plays an important role in shaping the antibody repertoire in humans.

MATERIALS & METHODS

Libraries

The cord blood IgM library was constructed from the mononuclear cells of two cord blood samples (Prabakaran et al., 2012a). Adult naïve Fab IgM libraries were constructed from peripheral blood collected from 69 healthy adult donors (Chen et al., 2010; Prabakaran et al., 2012b). Details of primers, amplification and sequencing methods are described in a reference (Zhu and Dimitrov, 2009). HIV-1 immune libraries covering both IgM and IgG were derived from the frozen peripheral blood mononuclear cells (PBMCs) and bone marrow (BM) samples of an acutely-infected HIV-1 patient. Further details regarding the IgG libraries from the PBMCs were given in a previous publication (Chen et al., 2012). All libraries were deep sequenced following the standard protocols described for the 454 sequencing method (Prabakaran et al., 2012b). Resulted sequences were trimmed to exclude short sequences and sequences with frame shifts or stop codons. Only sequences with at least 300 nucleotides and encoded complete heavy chain Ig V_H genes were used in the study. IMGT/High V Quest was used to identify V, D and J germline subfamilies and only productive heavy chain Ig V_H domain sequences were used for the Circos analysis (www.circos.ca). The total numbers of antibody sequences for each library are listed in table 1.

Circos Analysis

Circos software was selected for this study for its high data-to-ink ratio and ability to clearly display relational data. Circos open-source software was obtained from www.circos.ca. The VDJ region recombination data were reformatted using the R statistical programming language to comply with Circos data file requirements. Library sizes were normalized with Circos ideogram (circumference segments) scaling and sizing, permitting comparison of individual subgroups within libraries as well as across disparate libraries. Links, drawn from a V region to its observed J region recombinant partner, were used to demonstrate frequency of recombination, with thicker links indicative of higher frequencies of recombination. A heatmap track was implemented to further illustrate frequencies of VDJ recombination and delineate the abundance of each V-region subgroup; the ideogram space allotted to the V-region subgroup corresponds to the frequency of its observance relative to other subgroups. The darker the color of the heatmap, the higher the frequency of observance of that particular V-region subgroup. Links for subgroups with high frequencies of observance across all libraries were highlighted, while subgroup links with <1% of observance were grayscaled for purposes of clarity.

Analysis of the VDJ recombination was conducted with an additional stacked histogram track on each Circos diagram. This track illustrates the relative proportion of each VDJ recombination as a fraction of the total number of D-region sequences observed.

RESULTS AND DISCUSSION

From 454 sequencing results of libraries, only sequences with at least 300 nucleotides with no stop codon or frame shift were included in the final analyses. The total number of antibody sequences from each library is listed in Table 1. The libraries constructed from the HIV-1 infected patient (with $\sim 2 \times 10^8$ independent clones) had small pools of sequences because many of the sequences have short fragments, which were removed from the pools. With the remaining sequences in each library, subfamilies from HV, KL (kappa light chain) and LV (lamda light chain) are still well represented (Chen et al., 2012).

Analysis of D and J regions in all libraries

As the heavy chain V-DJ recombination follows the D-J recombination, we analyzed D and J gene combination frequencies first. D-J combination was analyzed in the cord blood, adult naïve IgM libraries and the 40 day PMBC IgM library (Fig. 1A, 1B, and 1C). The D-J correlations in these three libraries are almost identical to one another. One detectable discrepancy is a slight decrease in frequency of observance of the D7 subfamily in the adult IgM library and 40 day PMBC IgM library. In the latter, the D7 is exclusively joined with the J4 subfamily, but not J3. Data on DJ combination for all other libraries are not shown. DJ combination data are not integrated into the final VJ graph for purposes of clarity.

J4 is the most frequently used J region in all libraries analyzed

We observed that the J4 family was the most frequently recombined J region for all ten libraries analyzed in this study (Fig. 2 and Fig. 3). J4 is the most heavily used J region for joining with the total HV family as well as with any dominant HV subfamilies in a given library. Interestingly, J4 has previously been reported to be a dominant J region in naïve IgM and memory IgG and IgM libraries from healthy donors (Briney et al., 2012; Chapal et al., 2001). Although the J4 gene segment has a high sequence homology to J2, the preference for J4 over other subfamilies is still poorly understood.

Comparison of IgM libraries from cord blood and blood from healthy adults

An antibody library established from cord blood PBMCs would represent the status of the immune system before exposure to antigens from outside world. Studying the repertoire of such a library provides a starting point for monitoring complex changes in the immune system after coming into contact with antigens/pathogens. Similarly, it is important to understand whether the overall antibody repertoire undergoes significant change in response to aging.

To answer these questions, we sought to compare two such antibody IgM libraries: an IgM library constructed from cord blood samples and a naïve IgM library constructed from the blood of 69 healthy adult donors. On the Circos graphs (Fig. 2A and 2B) of the two libraries, the overall profile of heavy chain V and J gene recombination patterns is similar. However, there is an increase in HV6 (by 7.14% %) and HV7 (by 0.82%) families in the adult IgM library in comparison with the cord blood IgM library. A slight increase in J5 and J6 in the adult IgM library is detected. While, in both libraries, the HV1 family is the dominant V gene, the dominance of HV1-69 and HV1-2 is reversed in the two libraries (Table 2). The HV1-2 subgroup was favored in the cord blood library, but the adult naïve IgM library showed a slightly diminished frequency of HV1-2 and an increased frequency of HV1-69.

Overall, the adult IgM library contains a more balanced representation of families from HV1 to HV7 and from J1 to J7, which in turn provides the basis for more antibody diversity.

Analyses of HIV-1 libraries

Libraries from samples collected at 40 days and 8 months post-infection from an acutely HIV-1 infected patient were constructed in our laboratory. At both time points, PBMCs and bone marrow were used as the source for the construction of IgG and IgM libraries. All IgM libraries were compared with the adult naïve IgM library for VDJ recombination and HV family usage. IgG libraries from the HIV-1 infected patient were compared among themselves.

Analyses of IgM libraries

In the 40-day IgM PBMC library (Fig. 3A), the HV1 family is most dominant, with HV3 a close second, as observed in the adult naïve IgM library. Frequency of observance of the HV4 family decreases significantly from 26.15% in the adult IgM library to 8.66% in the 40-day PBMC IgM library. In addition, in the healthy donor IgM library, HV4 is joined to all J families (J1~J6), whereas, in the 40-day PBMC IgM library, HV4 mainly joins to J4 and J5. This VH4 to J4 and J5 pattern remains in the 8-month PBMC IgM library (Fig. 4A), although HV4 percentage now is closer to that in the healthy donor IgM library. In the 8-month IgM libraries, both PBMC and bone marrow (BM) libraries have HV3 as the predominant HV family; this differs from the healthy adult IgM library. Further examination of the VDJ recombination in these 3 libraries reveals a change in favored VDJ combinations. For example, in the healthy adult IgM library, the most frequently observed combination V(1-69)D3J4 encompasses ~1.47% of all HV analyzed. In the 8-month PBMC IgM library, the most frequent VDJ combination is V(3-23)D3J4, at ~4.73%. V(3-23)D3J4 is also the most frequently-observed (~5.77%) combination in the 8 month BM IgM library. Therefore, in the 8-month IgM libraries, in addition to a change in favored VDJ pairing, the selected V(3-23)D3J4 pairing has a much higher frequency, implying the selective expansion of particular B cell clones.

BM IgM (Fig. 3C and 4C) libraries have distinct HV-J joining patterns. Here, HV3 is by far the most dominant HV family in both 40-day and 8-month libraries, and HV3 joins almost exclusively with J4, J5, J6. HV1 is a distant second in the HV family in libraries of both visits.

Analyses of IgG libraries

In the 40-day PMBC IgG library (Fig. 3B), HV3 is the dominant family with HV4 a close second. However, after an 8-month infection (Fig. 4B), HV1 took over and HV3 became a distant second, indicating that the HV1 family was preferentially selected by exposure of B cells to antigens, mostly likely HIV-1 related antigens. In the 8-month IgM libraries, HV3, instead of HV1, was the predominant HV family, with V(3-23)D3J4 being the favored VDJ pairing. Differing HV patterns in IgG and IgM libraries indicate that the IgG and IgM repertoires are genetically distinct and are probably the result of exposure of subsets of B cells to different antigen stimuli.

The regained dominance of HV1 family in the 8-month IgG libraries is fascinating. Many well-studied anti-HIV-1 broadly neutralizing antibodies such as VRC01, b12, X5, 4E10 and m396 are derived from the HV1 germline family (HV1-69, 1-2 and 1-3) (Prabakaran et al., 2012b). These broadly neutralizing antibodies only arise after long-term infection with HIV-1, usually a period of 2 years (Wu et al., 2011). In addition, a recent study has reported that the majority of HIV vaccine-elicited antibodies capable of ADCC (antibody-dependent cell-mediated cytotoxicity) are also derived from the HV1 family (Bonsignori et al., 2011). Therefore, although the HV1 family is the dominant HV gene at birth and in adult IgM library, regaining HV1 dominance in the 8-month IgG libraries after losing dominance at 40 days is still significant. This particular patient did not develop broadly neutralizing antibody at an 8-month visit; nevertheless, the immune system is taking a right turn towards developing neutralizing antibodies. In terms of HIV vaccine design, our result would suggest benefits of boosting at a late stage.

The 40-day BM IgG library (Fig. 3D) bears a closer resemblance to the healthy adult IgM library than does the 8-month BM IgG library (Fig. 4D), where HV4 is significantly decreased (by 16.29%). Frequency of HV5-51 observance is decreased in the 40-day BM IgG library, but returns to a healthy donor's HV5-51 levels by the 8-month mark.

A striking feature in the four 8-month libraries is the high percentage of dominant families, whether HV1 or HV3. These families appeared at much higher frequency than did the second most dominant family in the libraries (see the last column of table 2), again indicating the preferred selection of these families after prolonged infection with HIV-1. These observations are in agreement with the oligoclonal tendency in the chronic viral infections that have been reported in literature (Breden et al., 2011). The oligoclonal expansion of the antibody subset is also observed in autoimmune diseases. It is reported that CD138+ plasma blasts isolated from cerebral spinal fluid (CeSF) of Multiple Sclerosis (MS) patients predominantly express (~70%) the HV4 gene family (Owens et al., 2007). The HV4 bias in MS CeSF B cells is, however, not found in the peripheral blood B cells of these patients, indicating that selective pressure is shaping the accumulation of B cells expressing the HV4 genes, rather than the availability of a B cell pool. In the thyroid tissue of Graves' disease, 33 out of 41 isolated anti-thyroid peroxidase antibodies from thyroid infiltrated B cells are derived from VH1-3 subfamily (Chapal et al., 2001). It is possible that antigenic stimulation through preferential binding to subsets of B cell receptors produces the antigen-driven survival/selection of these B cells.

In conclusion, by analyzing the deep sequencing results of ten antibody libraries with IMGT/V Quest and Circos, we have found that the immune system undergoes subtle changes in the antibody repertoire during growth, and HIV-1 infection selects certain antibody gene families and VDJ recombinations over others. Together, these mechanisms help shape the immune system in individuals. This novel information should help to guide the development of HIV vaccine and immunotherapy. The software used for these analyses are accessible through IMGT/V Quest (www.imgt.org/IMGT_vquest) and Circos (www.circos.ca). This method could be applied to mining large data sets such as, but not limited to, libraries/amplicons of antibodies and T cell receptors.

Acknowledgments

We thank Drs. B. Haynes, H. Liao, C. Broder, and T. Fouts for reagents. We thank the Laboratory of Molecular Technology and Advanced Biomedical Computing Center of SAIC-Frederick Inc. for sequencing service. We thank Ms. Maria G. Singarayan for constructing the PostgreSQL database and developing JAVA applications. This research was supported by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research, and by Federal funds from the NIH, National Cancer Institute, under Contract No. NO1-CO-12400. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does the mention of trade names, commercial products, or organizations imply endorsement by the U. S. Government.

References

- Agrawal A, et al. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature*. 1998; 394:744–51. [PubMed: 9723614]
- Arnaut R, et al. High-resolution description of antibody heavy-chain repertoires in humans. *PLoS ONE*. 2011; 6:e22365. [PubMed: 21829618]
- Bonsignori M, et al. Analysis of a clonal lineage of HIV-1 envelope V2/V3 conformational epitope-specific broadly neutralizing antibodies and their inferred unmutated common ancestors. *J Virol*. 2011; 85:9998–10009. [PubMed: 21795340]
- Boyd SD, et al. Measurement and Clinical Monitoring of Human Lymphocyte Clonality by Massively Parallel V-D-J Pyrosequencing. *Science Translational Medicine*. 2009a; 1
- Boyd SD, et al. Measurement and clinical monitoring of human lymphocyte clonality by massively parallel VDJ pyrosequencing. *Sci Transl Med*. 2009b; 1:12ra23.
- Breden F, et al. Comparison of antibody repertoires produced by HIV-1 infection, other chronic and acute infections, and systemic autoimmune disease. *PLoS ONE*. 2011; 6:e16857. [PubMed: 21479208]

- Briney BS, et al. High-throughput antibody sequencing reveals genetic evidence of global regulation of the naive and memory repertoires that extends across individuals. *Genes Immun.* 2012; 13:469–73. [PubMed: 22622198]
- Chapal N, et al. Thyroid peroxidase autoantibodies obtained from random single chain Fv libraries contain the same heavy/light chain combinations as occur in vivo. *Endocrinology.* 2001; 142:4740–4750. [PubMed: 11606439]
- Chen W, et al. Characterization of human IgG repertoires in an acute HIV-1 infection. *Exp Mol Pathol.* 2012; 93:399–407. [PubMed: 23036472]
- Chen W, et al. Cross-Reactive Human IgM-Derived Monoclonal Antibodies that Bind to HIV-1 Envelope Glycoproteins. *Viruses.* 2010; 2:547–565. [PubMed: 21755021]
- Dimitrov DS. Therapeutic antibodies, vaccines and antibodyomes. *Mabs.* 2010; 2:347–56. [PubMed: 20400863]
- Glanville J, et al. Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire. *Proc Natl Acad Sci U S A.* 2009; 106:20216–21. [PubMed: 19875695]
- Jiang N, et al. Determinism and stochasticity during maturation of the zebrafish antibody repertoire. *Proc Natl Acad Sci U S A.* 2011; 108:5348–53. [PubMed: 21393572]
- Kidd MJ, et al. The inference of phased haplotypes for the immunoglobulin H chain V region gene loci by analysis of VDJ gene rearrangements. *J Immunol.* 2012; 188:1333–40. [PubMed: 22205028]
- Lefranc MP, et al. IMGT, the international ImMunoGeneTics information system. *Nucleic Acids Res.* 2009; 37:D1006–12. [PubMed: 18978023]
- Li L, et al. Identifying compartment-specific non-HLA targets after renal transplantation by integrating transcriptome and “antibodyome” measures. *Proc Natl Acad Sci U S A.* 2009; 106:4148–53. [PubMed: 19251643]
- Lloyd C, et al. Modelling the human immune response: performance of a 1011 human antibody repertoire against a broad panel of therapeutically relevant antigens. *Protein Eng Des Sel.* 2009; 22:159–68. [PubMed: 18974080]
- Owens GP, et al. VH4 gene segments dominate the intrathecal humoral immune response in multiple sclerosis. *J Immunol.* 2007; 179:6343–51. [PubMed: 17947712]
- Prabakaran P, et al. Expressed antibody repertoires in human cord blood cells: 454 sequencing and IMGT/HighV-QUEST analysis of germline gene usage, junctional diversity, and somatic mutations. *Immunogenetics.* 2012a; 64:337–50. [PubMed: 22200891]
- Prabakaran P, et al. Origin, diversity, and maturation of human antiviral antibodies analyzed by high-throughput sequencing. *Front Microbiol.* 2012b; 3:277. [PubMed: 22876240]
- Ravn U, et al. By-passing in vitro screening--next generation sequencing technologies applied to antibody display and in silico candidate selection. *Nucleic Acids Res.* 2010; 38:e193. [PubMed: 20846958]
- Weinstein JA, et al. High-throughput sequencing of the zebrafish antibody repertoire. *Science.* 2009; 324:807–10. [PubMed: 19423829]
- Wu X, et al. Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. *Science.* 2011; 333:1593–602. [PubMed: 21835983]
- Zhu Z, Dimitrov DS. Construction of a large naive human phage-displayed Fab library through one-step cloning. *Methods Mol Biol.* 2009; 525:129–42. xv. [PubMed: 19252833]

Fig. 1. Links drawn from a D region to a J region denote an observed DJ recombination. Wider links indicate higher frequencies of observance. DJ recombination frequencies in cord blood (A), adult naïve IgM libraries (B) and the 40 day (2nd visit) PBMC IgM library (B).

Fig. 2.

Comparison of V and J gene associations in cord blood (A) and adult naïve repertoires (B). Outermost tracks (1) mark the boundaries of each V or J region subfamily in the Circos plot. These tracks, called ideograms, are colored to match their respective links, found in track three (3). Heatmaps (2) indicate the relative frequencies of subgroups within HV subfamilies (e.g., HV1-69 and HV3-23); colors were applied in deciles, with each tenth of the data set assigned a different color. Note: the heatmap is an indicator of percentile rank within the data set, and not observance as a percentage of the total number of sequences. Links (3) indicate the relative frequencies of specific V-J combinations—wider links indicate higher frequencies of recombination. Highlighted links indicate combinations that constitute 1% or more of all sequences observed.

Fig. 3.

Comparison of V and J genes associations in the libraries constructed from 40 days (1st visit) PBMC IgM (A), PBMC IgG (B), BM IgM (C), and BM IgG (D). Outermost tracks indicate the prevalence of each V or J region subfamily in the Circos plot. More frequently observed subfamilies are represented by longer outer tracks, called ideograms. All ideograms are colored to match their respective links, found in the ribbon track. Heatmaps (center track) indicate the relative frequencies of subgroups within HV subfamilies (e.g., HV1-69 and HV3-23); colors were applied in deciles, with each tenth of the data set assigned a different color (see legend for Fig. 2 for color details). Links indicate the relative frequencies of specific V-J combinations and wider links indicate higher frequencies of recombination. Highlighted links indicate combinations that constitute 1% or more of all sequences observed.

Fig. 4.

Comparison of V and J gene associations in the libraries constructed from 8 months (2nd visit) PBMC IgM (A), PBMC IgG (B), BM IgM (C), and BM IgG (D) (see legends for Fig. 2, and Fig. 3 for the details of colors, links and tracks).

Table 1

Number of complete Ig V_h genes analyzed for each library.

Library	Total Sequences of IgHV Gene
Cord blood IgM	27,854
Adult Naïve IgM	69,897
40-day Post HIV-1 Infection:	
PBMCs IgM	1,062
PBMCs IgG	943
BM IgM	1,068
BM IgG	985
8-month Post HIV-1 Infection:	
PBMCs IgM	972
PBMCs IgG	858
BM IgM	607
BM IgG	741

*PBMCs: peripheral blood mononuclear cells. BM: bone marrow.

Table 2

Major HV families/subfamilies in the libraries established from cord blood healthy donor and an HIV-1 infected patient.

Library	Dominant VH family (%)	Main subfamilies of the dominant HV	Second dominant VH family (%)	% Difference between 1 st and 2 nd dominant families
Cord blood IgM	HV1 (45.80)	1-2, 1-69, 1-8, 1-46	HV3(23.43)	22.37
Adult Naïve IgM	HV1 (34.40)	1-69, 1-2, 1-8, 1-46	HV4 (26.15)	8.25
Post HIV-1 Infection 40 days:				
PBMC IgM	HV1 (40.96)	1-69, 1-2, 1-8, 1-46	HV3 (38.42)	2.54
PBMC IgG	HV3 (33.83)	3-33, 3-23, 3-13, 3-15	HV4 (27.89)	5.94
BM IgM	HV3 (51.22)	3-23, 3-48, 3-33, 3-30	HV4 (23.69)	27.53*
BM IgG	HV1 (35.13)	1-69, 1-2, 1-46, 1-18	HV4 (28.32)	6.81
Post HIV-1 Infection 8 months:				
PBMC IgM	HV3 (57.10)	3-23, 3-15, 3-74, 3-48	HV4 (17.18)	39.92*
PBMC IgG	HV1 (53.03)	1-69, 1-2, 1-46, 1-18	HV3 (22.96)	30.77*
BM IgM	HV3 (59.31)	3-23, 3-48, 3-7, 3-74	HV1 (15.98)	43.33*
BM IgG	HV1 (46.29)	1-69, 1-2, 1-46, 1-18	HV3 (29.15)	17.14