WHY MAINTAIN LIGHT CHAIN ISOTYPES? THE INFLUENCE OF HEAVY CHAIN ISOTYPE AND COMPLEMENTARY DETERMINING REGION LENGTHS UPON LIGHT CHAIN ISOTYPE IN *XENOPUS LAEVIS*

A Thesis

by

NATALIE JO JACOBS

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Michael F. Criscitiello
Co-Chair of Committee,	C. Jane Welsh
Committee Member,	Scott V. Dindot
Head of Department,	Roger Smith III

December 2015

Major Subject: Biomedical Sciences

ABSTRACT

Different immunoglobulin (Ig) heavy chain (H) isotypes have distinct functions, but so far it is unclear if Ig light (L) chains follow the same pattern. It is usually assumed that form follows function; but if this is true, then why have different IgL isotypes with no known functional differences? In this study we investigate IgH and IgL isotype preferential binding and complementary determining region (CDR) lengths to try to address this question using the African clawed frog, *Xenopus laevis*, as a model. Amphibians exhibit IgH isotype class switch at a single IgH locus and have an additional, more divergent, IgL isotype (σ) plus the two found in mammals (λ and κ). We used quantitative PCR (qPCR) analysis of IgH isotype of B cells sorted by surface IgL isotype expression to find evidence of preferential use of IgL isotype by IgH isotype. We found a relative skewing in the Igκ cells for IgY, in the Igλ cells for IgX, the Igσ cells for IgM, and corroborated published immunoprecipitations showing that IgY and Igσ do not pair with gene expression data of the IgL isotype sorted cells. Our data also suggests that the exaggerated CDR1 of IgHV families III and VII and the long CDR2 of Igσ may cramp IgH CDR3, making the IgHV III/VII-Igσ pairing less common. While these data do not resolve the conundrum of multiple IgL isotype maintenance in vertebrates, they do show that in a tetrapod with several IgH and several IgL isotype options, IgL isotype use is not random.

DEDICATION

This is dedicated to the ones I love. To my father, who I miss dearly; my mother, who has stood by me through it all; and to my love, whose unending patience and kindness exceeds that of a Saint. I love you all.

And in memory of Dr. Ashley Peterson who lit a path to guide my way. Your light and smile shall live on in our hearts forever.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Michael Criscitiello, co-chair Dr. Jane Welsh, and committee member Dr. Scott Dindot for their guidance and support throughout the course of this research.

Thanks also go to my friends and colleagues in the lab and the department faculty and staff for making my time at Texas A&M University a great experience

This work was supported by awards from the NIH (AI56963) to MFC and (R01OD0549) to MFF and YO, and the NSF (IOS1257829). Thad Deiss was supported in part by a graduate merit fellowship. The work of Dr. Christina Du and Dr. Sara Mashoof was funded by individual student research grants through the Texas A&M College of Veterinary Medicine and Biomedical Sciences.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION AND LITERATURE REVIEW

Antibodies, or Immunoglobulins, are the crucial antigen receptor of the vertebrate humoral adaptive immune system. They protect the host from pathogens in many ways including neutralization, opsonization, and activation of the complement cascade. The structure of an antibody consists of two heavy and two light peptide chains covalently connected by disulfide bonds. The N-terminal domains of both Ig light (IgL) and Ig heavy (IgH) are greatly variable for antigen binding and the constant (C) terminal domains are nearly constant within isotypes. There are three CDRs that are part of the V domain. CDR3 is the most diverse and includes the H VDJ gene segments. Many IgH isotypes have been described in vertebrates but all seem to be orthologous with either IgM, IgD, IgT (the dedicated mucosal isotype of teleost (bony) fish (1)), IgX/A (the dedicated mucosal type of tetrapods (2)), or the IgY/IgG/IgE family (reviewed in (3)). Four IgL isotypes have been described. Igλ and Igκ are used by most vertebrate groups, whereas σ and σ-cart are only found in fishes and amphibians (4).

The constant regions of IgH isotypes impart functional distinction to antibodies, but why have multiple IgL isotypes? The extra loci certainly complicate haplotype exclusion, although it has been suggested that the additional loci provide a larger canvas upon which receptor editing can paint, possibly making the confounding path to clonal receptor expression on a lymphocyte worthwhile for central B cell tolerance (5). Additional means of diversification of the antibody repertoire seems a logical explanation for IgL isotypes. But that would favor simpler diversification of divergent V

families at one locus as is seen with IgH isotype, and this mechanism is not supported by comparative data for IgL isotypes (6). If distinct IgL isotype loci are being maintained evolutionarily, discrete physiology should be expected of their products. There is little evidence for distinct roles for mammalian Igκ and Igλ, but hints can be found in the literature (e.g., (7)). In *X. laevis*, IgL isotype (Igσ) only associated with two of the three IgH isotypes (IgM and IgX), displaying a propensity in this amphibian of IgLσ for the two T-independent IgH isotypes expressed in the intestine (8-10). Work in skate showed a large disparity in the ratio of IgL isotype expression in the intestine (11). We have suggested that the principal selection component could be the heterodimerization requirements of the two V domains (12). These ideas were first tested in an animal with clear IgH/IgL isotype preference demonstrated at the protein level (8), *X. laevis*.

2. METHODS

2.1 Animals and cell harvest

The African clawed frog, *X. laevis*, served as the model for this study. The original frogs were acquired from *Xenopus* Express (Brooksville, FL). Outbred *X. laevis* were spawned and maintained as previously described (13). Briefly successive generations were bred using human chorionic gonadotropin hormone (Sigma-Aldrich, St Louis, MO). Two XenoPlus (Techniplast, Buguggiate, Italy) recirculating aquatic husbandry systems were used to support the frogs. A sinking pellet diet was used for adult frogs and a powdered diet for tadpoles (*Xenopus* Express). The room was kept on a 12 hour light cycle at a temperature of 23°C.

Four unimmunized adult frogs age 10-12 months were selected for each of three cell sort experiments and euthanized with MS-222 (tricaine methyl sulfonate, Argent, Redmond WA) overdose. Spleens were dissected and dissociated with tweezers on a steel mesh in a small petri dish filled with 2mL Magnetic Activated Cell Sorting (MACS), Miltenyi Biotech, San Diego CA) buffer (0.5% bovine serum albumin, 2mM EDTA, in phosphate buffered saline adjusted for amphibian salinity (65:35 mammalian PBS to water)). Splenocytes were counted manually with a hemocytometer or with a Cellometer Auto 1000 (Nexcelom Bioscience, Lawrence MA). All frog procedures and care were approved by the Texas A&M Institutional Animal Care and Use Committee (Animal Protocol 2011-303).

2.2 Cell sorting

Anti-frog IgL monoclonal antibodies (mAb) used for sorting were developed in the laboratory of Louis Du Pasquier: mouse anti-frog Igκ 409B8, mouse anti-frog Igλ 1E9, and mouse anti-frog Igσ 13B2 (8). Splenocytes were diluted in primary antibody (600uL in 1:10 amphibian PBS) and incubated for one hour on ice inverting every 15 minutes. They were then washed with 10mL of MACS buffer and centrifuged for 10 minutes at 10°C at 1000rpm twice. The cells were resuspended in MACS buffer (80uL per 10^7 starting cells). Goat anti-mouse IgG MicroBeads (Miltenyi Biotech) were added (20uL per 10^7 starting cells), mixed well, and incubated for one hour on ice inverting every 15 minutes. They were then washed with 2mL MACS buffer per 10^7 starting cells and centrifuged twice as before. Cell pellets were resuspended in 500uL MACS buffer per 10^7 starting cells. Cells were passed through a MACS Pre-Separation Filter (Miltenyi Biotec) pre-wet with 1mL MACS buffer to remove clumps. A MACS LS separation column (Miltenyi Biotec) was placed in a magnetic stand and wet with 3mL MACS buffer. Cells were loaded 3mL at a time and gravity-fed into 15mL collection tubes on ice. Cells on magnetized column were washed with 3mL MACS buffer three times. Flow through was collected as negative populations. Columns were then removed from the magnetic stand, 5mL MACS buffer was applied to column, and cells were immediately pushed with plunger into a new collection tube for positive sort. A second round of MACS with a new column was used to double purify the IgL isotype positive populations. Cells were again counted to determine the final number of negative and positive cells. Cell loss was high in the two rounds of selection, usually around 80%.

2.3 Quantitative PCR

Using the RNeasy kit (Qiagen, Germantown MD) per the manufacturer's instructions, RNA was purified from the sorted cells. The first strand complementary (c) DNA was synthesized using random hexamer priming with SuperscriptIII (Life Technologies, Grand Island NY), and both RNA and cDNA were measured for quality and quantity using the NanoDrop 2000c spectrophotometer (Life Technologies). Standard PCR was performed to evaluate the representative quality of the cDNA using the primers in Table 3. The PCR conditions were 30 cycles, 30s, 95°C denaturation, 30s annealing and 1m extension at 68°C, with initial 3m denaturation and final 5m extension, where annealing was adjusted for each amplicon to 5°C below the lower of the Tms listed in Table 3.

The qPCR reactions were performed with 50ng of cDNA with the recommended 5x HOT FIREPol Eva Green HRM Mix (without ROX, Solis Biodyne, Tartu Estonia) per manufacturer's instructions. The qPCR samples were cycled 45 times with the annealing temperature set to 55°C on a LightCycler 480 (Roche, Basel Switzerland), followed by melting curve analysis. The Roche LightCycler software was utilized for raw data acquisition and calculation of Ct (threshold cycle) values. Changes in gene expression were estimated using the $2^{-\Delta\Delta Ct}$ method (14), with β_2-microglobulin as the reference gene. IgL isotype expression in positive and negative sorts was subjected to median normalization to create normalization ratios for IgH isotype expression accounting for sort efficiency (15).

2.4 Sampling of IgH CDR3 from IgL isotype sorts by plasmid cloning

Standard Taq DNA polymerase (New England BioLabs, Ipswich MA) was used in 25uL reactions with primers in Table 3 then transferred to a C1000 Thermal Cycler (Bio-Rad). The products were run on a 0.8% agarose gel and the desired DNA products were identified with ultraviolet light and cut from the gel for cloning. Topoisomerase cloning and transformation were executed following the TOPO TA Cloning user manual (Life Technologies). White or light blue colonies were selected for overnight cultures for plasmid minipreps. Plasmid DNA was purified using the ZR Plasmid Miniprep™-Classic kit (Zymo Research, Irvine CA) followed by *Eco*RI (Promega, Madison WI) digestion. The digested samples were evaluated on the NanoDrop 2000c spectrophotometer and analyzed on an agarose gel. BigDye V3.1 (Life Technologies) sequence reactions were made with M13 forward primers and cycled per the manufacturer's recommendations. The BigDye XTerminator Purification Kit™ (Life Technologies) was used to prepare the products for sequencing by the DNA Technologies Core Lab in the Department of Veterinary Pathobiology at Texas A&M University.

2.5 Immunogenetic analysis

Sequence data was managed in the Geneious bioinformatic suite version 7.1 (Biomatters, Auckland New Zealand). Lengths of CDR3 were calculated as the exclusive number of amino acids between the conserved cysteine of the YxC motif of the variable (V) and the first glycine of the GxG motif of the joining (J) (16). IgH variable (V), diversity (D) and joining (J) gene sequences were called as previously

assigned (17-19) for *X. laevis* and limits of genomic coding sequence was confirmed in the JGI genomic assembly version 7.1 by finding the conserved heptamer (CACAGTG) and nonamer (ACAAAAACC) motifs of the recombination signal sequences (Table 4). Four bases of identity were required to count in a D segment versus N/P addition, which most likely skews the N/P count higher. Sequences were submitted to the National Center for Biotechnological Information.

2.6 Molecular modeling

Hypothetical single-chain fragment variable molecules (scFv) were created in silico to model the tertiary and quaternary interactions between various frog IgH and IgL isotype, V family, and CDR3 length combinations. These were made by bridging frog IgLV and IgHV in two possible orientations: with a linker at the carboxyl terminus of the IgHV to the amino terminus of the IgLV and vice versa, with the IgL-linker-IgH oriented constructs yielding better results as judged by QMEAN scores (20). Linkers of varying length and amino acid content were also tried before settling on "GGGSGGGSGGGS" (21), again for better QMEAN scores. Structural templates were chosen by SWISS-MODEL (22, 23) and the resulting amino acid alignments visually inspected but no manual adjustments were necessary. All models used had global mean quality estimation (GMQE) scores of 0.75 or above. Coordinates of homology models were visualized in Geneious and images generated in that software suite.

2.7 Flow cytometry

Adult *X. laevis* splenocytes ($5x10^5$ cells/per treatment) were stained with 1:100 dilutions of one of the mAbs described above: mouse anti-frog Igκ 409B8, mouse anti-

frog Igλ 1E9, and mouse anti-frog Igσ 13B2 (8). Cells were then washed 3x with staining buffer before staining with 1:100 FITC labeled rabbit anti-mouse IgG_1 (Sigma, Saint Louis MO) for 30min at 4°C. All samples were washed and resuspended in 300 μL of staining buffer containing 0.1% sodium azide and examined by flow cytometry on a BD LSR II instrument (BD Biosciences, San Jose CA) in the Flow Cytometry Core Facilities of Veterinary Pathobiology or the University of Maryland Department of Microbiology and Immunology. Fifty thousand events were collected, gated for live cells, and analyzed using the FlowJo software (Tree Star Inc., Ashland OR).

3. RESULTS

A surface IgL isotype sorting strategy was employed to study the IgH isotypes, V(D)J gene segments, and CDR3 that were found with the IgL κ, λ and σ isotypes of frog. Although other isotypes are encoded in the IgH locus of *X. laevis* and expressed at low but appreciable levels (24, 25), we focused this study on the IgM, IgX and IgY isotypes dominantly expressed in the spleen and mucosa (26-28 and reviewed in 29).

3.1 IgH/IgL isotype pairing

IgLκ and IgLσ isotypes enriched at least 12-fold by double-MACS with mouse anti-frog Igκ 409B8 and mouse anti-frog Igσ 13B2, respectively, against the other IgL isotypes, while Igλ was enriched over 23-fold against Igκ, although only about seven-fold against IgLσ isotype by the mouse anti-frog Igλ 1E9 mAb (Figure 1A). qPCR of IgH isotypes in these sorted cells showed that IgM dominates in all IgL isotype sorts, but the relative amounts of the IgH isotypes measured did have significant differences. Igκ's use is skewed to IgY, Igλ's to IgX, and Igσ's to IgM (Figure 1B). High C_t values for IgY of the Igσ sort (all >33) contributed to the large standard error in that IgH/IgL isotype pairing for which we failed to recover a single clone (Figure 1B).

3.2 IgH CDR3 length with different IgL isotypes

In order to examine the V(D)J rearrangements of the IgH isotype that were found in cells sorted by IgL isotype, individual 5' RACE PCR products were cloned and sequenced for each of the nine IgH(M, X, Y)/IgL(κ, λ, σ) isotype combinations, with the exception of IgY/Igσ isotype pairing for which no bands were amplified and no clones

were captured (even from “blind” gel excision and cloning from the appropriate migration region of gel lane for the amplicon). A total of 304 cloned IgH amplicons were sequenced, of these 108 had unique sequences in the IgH CDR3 and were assumed to be the product of a unique B lymphocyte clone (Table 1). The amino acid translations of the unique CDR3 clones are displayed in Figure 7, along with the V and J segments used. The assignment of nucleotide origin to germline genomic V, D or J sequences was manually annotated (Figure 8).

Alignment of the CDR3 nucleotide sequences to genomic V(D)J elements allowed discernment of nucleotides not germline encoded (non-template (N) and palindromic (P) nucleotides). Significant differences were not found in the length of CDR3H in different IgH/IgL isotype combinations, although this analysis was limited by the number of clones available from pairings such as IgX and IgY with Igσ (Figure 2A). More or less N and P nucleotides contributing to the CDR3 length also showed no proclivity for particular IgH/IgL isotype pairings (Figure 2B), although the number of retained N and P nucleotide additions after exonuclease activity did correlate with longer CDR3 length (average value x, +/- std dev, plots shown in Figure 9). Of the 108 unique clones analyzed, three were not used in CDR3 length analysis due to incomplete V(D)J rearrangement resulting in (D)JC sequences without V segments. Clone 180731 was amplified with an IgX primer but contained an IgY C region sequence and was analyzed as such.

In the entire analysis only one unique CDR3 sequence appeared in two different sorts and was counted in both sets, the IgY clones 020719 from Igκ sort and 060701

from the Igλ sort. This was probably due to the imperfect nature of the sort. Two point mutations in the V gene segment distinguish the clones, presumably from somatic hypermutation as some V's were more heavily mutated away from genomic sequences than others.

3.3 *X. laevis* IgH isotype V, D and J use

We looked for bias in the use of particular IgH V, D or J segments in pairings with particular IgL isotypes. No significant biases were found in the use of D and J elements in cells using particular IgL isotypes (Figure 10). D3, J3, J4 and J6 were not used and D4 only rarely. As mentioned above, some sequences lacked discernable D contributions to CDR3 while others used more than one D. We found no sequences with more than two Ds rearranged.

Although no striking biases were seen with IgHV family use among IgH or IgL isotypes (Figure 3A), an interesting observation was made with IgH V families III and VII, which have considerably longer CDR1 lengths than the other V families of *X. laevis* (Figure 3B) (17). The IgLσ isotype sorted cells showed a lower frequency of these two IgH V families (Table 2). Interestingly, the CDR3 lengths of IgHV family III and VII using rearrangements from the Igλ and Igκ sorts is somewhat longer (mean of 9.0 amino acids) than those IgHV family III and IgHV family VII rearrangements from Igσ (mean of 8.7 amino acids).

3.4 IgL isotype exclusion

Flow cytometry was performed to more quantitatively determine IgL isotype frequencies suggested by the MACS and RNA recovery (Figure 5 and Table 5). In Figure 5 A-C each IgL isotype was gated (on the top dot plot) then measured for expression of each IgH (bottom three dot plots). Figure 5A is an IgLxIgH isotype analysis gated on Igκ positive cells and analyzed for percent IgH positive cells. Igκ was 98.3% positive for IgM, .54% positive for IgX, and .71% positive for IgY. Figure 5B is gated on λ positive cells and was 98.6% positive for IgM, 1.55% positive for IgX, and .74% positive for IgY. Figure 5C is gated on Igσ positive cells and was 84.7% positive for IgM, 7.19% positive for IgX, and .57% positive for IgY. In Figure 5 E-G each IgH was gated (top dot plot) then measured for expression of each IgL isotype (bottom three contour maps). This was the reciprocal test of what was done in Figure 5 A-C. Figure 5E is an IgHxIgL isotype analysis gated on IgM positive cells and analyzed for percent IgL isotype positive cells. IgM was 60.9% positive for Igκ, 23.5% positive for λ, and 1.6% positive for Igσ. Figure 5F is gated on IgX positive cells and was 26.4% positive for Igκ, 34.8% positive for λ, and 32.3% positive for Igσ. Figure 5G gated on IgY positive cells and was 64.3% positive for Igκ, 15.3% positive for λ, and 1.72% positive for Igσ. Figure 5D and 5H are bar graphs of the previously described dot plots and contour maps. Figure 5I is a light scatter gating for lymphocytes, then single cells. Fifty thousand events were collected, gated for live cells, and analyzed as previously described.

Similar relative ratios of Igλ, Igκ and Igσ, respectively, were found as by cell recovery (Table 5). Igσ was expressed much less overall except for IgX. There is a

notable disparity between the lower Igκ expression and higher Igσ expression for IgY. The number of Igκ, Igλ, and Igσ binding cells do not suggest double staining. This allowed us to discount the possibility of cells having more than one IgL isotype expressed. Thus, the spleen B cells of *X. laevis* appear to be isotypically excluded for IgL isotype.

4. DISCUSSION

4.1 IgL isotype use is far from random in *X. laevis*

Bias in IgH/IgL isotype pairing and differences in CDR lengths show that isotype use is far from random in *X. laevis*. The data in Figure 1B suggest a bias in IgH/IgL isotype pairing in a species with canonical tetrapod class switching and an additional IgL isotype than is found in mammals: a proclivity for IgX/Igλ, IgY/Igκ and Igσ pairing with IgM and IgX but not IgY. A model of this IgH/IgL isotype skewing can be seen in Figure 6. Similar ratios were found using flow cytometry to confirm the IgH/IgL isotype bias (Figure 5). These findings support a previous immunoprecipitation study where Igσ was found to only associate with two of the three IgH isotypes (IgM and IgX), showing a preference for the two T-independent IgH isotypes expressed in the intestines (8). Work in skate showed a large disparity in the ratio of IgL isotype expression based on life stage and location (11). Sharks are biased for Igκ expression (30). If allowed some assumptions (e.g. mRNA coinciding with protein levels), we can interpret these data as reason to expect distinct function in the IgL isotypes of some vertebrates.

Clearly IgL isotype use is not random as described in our qPCR data and flow cytometry and by immunoprecipitation and qPCR and northern blotting in these two previous studies. The next question is where does the preferential binding occur? Our study only examined what happened in the spleen. Is this skewing also present in primary lymphoid tissues, like bone marrow, or did it occur after antigen recognition and class switch recombination (CSR) in secondary lymphoid tissues, like the spleen? We

propose that the bias occurs after antigen exposure and CSR in the secondary lymphoid tissues (Figure 6).

B cell precursors originate in the bone marrow. IgH rearranges first, expressing IgM, then IgL isotypes rearrange. Immature B cells are tested for self-recognition. If self- reactive, it can rearrange V region genes (receptor editing) or die. After antigen exposure in the germinal center, B cells with high affinity surface Ig for the antigen will receive activation signals from antigen on follicular dendritic cells to proliferate and differentiate vs those with low affinity will not receive survival signals and die.

Awkward VH-VL or C1H-CL combinations may be caused by steric hindrance and are less likely to occur or survive. Longer CDR1 lengths found in IgHV families III and VII may cause steric hindrance and may explain families being found at a lower frequency with the long CDR2 of Igσ compared to the other IgL isotype sorts. (Table 2 and Figure 3). This steric hindrance is corroborated by threading molecular modeling (Figure 4). The space filling model of VH family III employing IgH from IgLσ isotype sort shows crowding of CDR3 IgH by long Igσ CDR2 and VHIII CDR1. As mentioned earlier, the CDR3 lengths of IgHV family III and VII using rearrangements from the Igλ and Igκ sorts is somewhat longer (mean of 9.0 amino acids) than those IgHV family III and IgHV family VII rearrangements from Igσ (mean of 8.7 amino acids), suggesting a further compensation for steric hindrance brought about by the longer IgH CDR1 and IgL CDR2.

Comparing our data on IgH/IgL isotype bias and CDR3 lengths to that of antigen naïve yet mature B cells in the bone marrow could solve the mystery of where (primary

or secondary lymphoid tissue) and when (before or after antigen exposure and CSR) isotype skewing occurs and lead us closer to finding possible functions for IgL isotypes.

4.2 Hints of IgL isotype distinct physiology

The use of IgH V, D, and J to create diverse combinations of CDR3 is critical to formation of antigen binding sites that protect against countless different pathogens. Post-metamorphosis adult frogs were chosen as more rearrangement diversity (31) and longer CDR3H lengths had been noted in *X. laevis* adults by resolving labeled products on sequencing gels (32). Our sequencing data found CDR3H lengths averaging 10-11 codons similar to previous sequencing gel findings (32). Corresponding CDR3H lengths are found in trout, mice, and human (33, 34). This conservation in size may account for reduced antibody diversity in lower vertebrates. The structure of IgH isotype is similar among vertebrates but the arrangement of the IgH V, D, and J genes differ. Lower vertebrates have a cluster arrangement that does not allow as much combinatorial diversity as the one with multiple repeats of single IgH V, D, and Js (27).

An aside to our central hypothesis is that originally distinct IgL isotypes evolved to pair with products of distinct IgH chain loci in cartilaginous fish. But the IgL isotypes have since (in tetrapods with class switch at a single IgH locus) maintained discrete CDR 1 and 2 lengths to allow at least one isotype to yield a favorable paratope topology when paired with IgH CDR3s of diverse length. The length of CDR3H in *X. laevis* is long for IgH CDR1 and IgL CDR2 in the Igσ sorts but the opposite pattern is seen in the Igκ and Igλ sorts (8, 11). This crowds the CDR3H and is less compatible.

Is it possible that the camelids and sharks that employ some IgH without IgL isotypes (35), or bovidae that use ultralong CDR3 domains (36), are ways of freeing the IgH from the need for IgL isotype fit, and in the process opening up new paratope design space? In this work we have focused on the pairing of the IgL V domain with IgH V, but distinct function could be found in the C domains of IgL isotype as they are in IgH. Alternatively, it has been suggested that steric hindrance could make certain IgL isotype C and IgH isotype C1 domains incompatible (8). The retention of multiple IgL isotypes compounds allelic exclusion with isotype exclusion (5), so there must be advantages that make it worthwhile. In Figure 2, CDR3H lengths were not significantly different amongst IgH/IgL isotype combinations although this analysis was limited by the number of clones available. N and P nucleotides contributing to the CDR3 length also showed no bias in IgH/IgL isotype pairing. Adult *X. laevis*, but not tadpoles, have N and P nucleotides to make CDR3 more diverse (37). There was a correlation between longer CDR3 length and number of retained N and P nucleotide additions after exonuclease activity (Figure 9). No significant bias in the use of D and J segments in IgH of particular IgH isotype and IgL isotype sort were found through sequencing and cloning (Figure 10).

Glycine residues in trout CDRH3 sequences resemble that of *X. laevis* and may contribute to a more flexible conformation in the CDR3 loop allowing greater mobility. This has been hypothesized to help these antibodies bind to a greater variety of antigens (33).

4.3 Evolution of IgL isotype physiology

The evolution of Ig isotypes in different classes of vertebrates may be important in the understanding of IgL isotype functions. Jawless fish, do not have immunoglobulins, but instead use variable like receptors for their adaptive immune response (38). IgH, IgM and IgD are found in most jawed vertebrates from cartilaginous fish to mammals. IgY arose in amphibians and later evolved into mammalian IgG and IgE. IgX also emerged in amphibians and is orthologous to the mucosal antibody IgA (2) in reptiles, birds, and mammals. Igλ and Igκ are found in most vertebrates. So far Igσ has only been described from amphibians, teleosts, and cartilaginous fish. σ-cart was discovered in cartilaginous fish, hence the name, but was later found in bony fish as well. Interestingly, there is an overall trend of an increasing number of IgH and decreasing number of IgL isotypes as more recent classes of vertebrates evolved.

Diversity in antigen recognition is crucial for survival and depends upon selection of V genes from IgH and IgL isotypes during B cell development. A decrease in IgL isotype during vertebrate evolution seems counterintuitive. It would appear more logical and advantageous for a greater number of IgL isotypes to be present to give more IgH/IgL combinatorial diversity if that is the role of IgL isotypes. Instead less IgL isotypes are employed for some reason as vertebrates evolved. Perhaps other mechanisms have emerged to compensate for the loss of diversity caused by less IgL isotype V genes to work with?

The platypus is an interesting creature telling of an evolutionary crossroads. Its genome reveals connections to mammals, reptiles, birds, amphibians, and fish (39).

Although the platypus is related to vertebrates, it does not have Igσ. However, highly diverse V Igλ repertoire and CDR lengths may compensate for the lack of IgL isotype diversity in platypus and birds (39).

Other interesting examples of mechanisms that evolved to compensate for the loss of IgL isotype are seen in nurse sharks and camelids. Two very different vertebrate groups have developed antibodies absent of IgL isotype that are highly functional and expressed in equal amount to conventional antibodies with both IgH and IgL isotypes (35). The camel IgH antibody developed from a normal IgH V gene fairly recently whereas the nurse shark Ig is novel and must have been created during an evolutionary event approximately half a billion years ago. Sharks have a multiple cluster Ig VDJ arrangement (VDJ)n compared to other vertebrates that have a translocon organization (VnDnJn). This might give the shark more flexibility to create new Ig loci and thus antibody structures. Camels have adapted to extreme heat and sharks contain a high concentration of urea in their blood that may induce environmental stress contributing to the formation of antibodies absent of IgL isotype (35). It must have been very difficult to create antibodies without IgL isotype (single domain V) since it is has only been described in camelids and sharks so far, but once this mutation appeared, it was seen as highly advantageous because it creates a more diverse projecting single variable domain paratope therefore possibly protecting against more recessed epitopes. Due to the unique conformation of these IgH only antibodies, they are highly effective against certain antigens that antibodies with IgL isotype may not be able to bind to. For example, the IgH only antibodies in camels infected with trypanosomes bind with high affinity to this

parasite (35). Perhaps more species with antibodies absent of IgL isotype will be discovered. Studying the animals that have naturally evolved this unique variation will help us better understand IgL isotype usage and possibly find better treatments for infectious diseases caused by pathogens whose neutralizing epitopes do not bind to conventional antibodies and may have a higher affinity for antibodies without IgL isotype. Hopefully, more variations in IgL isotype use will be found in the future and may help us understand the function of IgL isotypes.

5. CONCLUSIONS

We hypothesized that IgL isotypes evolved different functions based on preferences in IgH isotype binding and CDR lengths. Our findings are supportive of this hypothesis. Although more work needs to be done to find possible functions of IgL isotypes, our results for IgH/IgL isotype pairing, IgH CDR3 lengths with different IgL isotypes, *X. laevis* IgH VDJ use, and flow cytometry shed some light on this enigma. For IgH/IgL isotype pairing we found preferential binding between IgX/Igλ, IgY/Igκ and Igσ pairing with IgM and IgX but not IgY. IgH CDR3 length with different IgL isotypes averaged 10-11 codons similar to previous studies (40). We did not find evidence for CDR3H skewing with IgL isotype due to the low number of some clones. N and P nucleotides contributing to the CDR3 length showed no proclivity for particular IgH/IgL isotype pairings, although the number of retained N and P nucleotide additions after exonuclease activity did correlate with longer CDR3 length. *X. laevis* IgH VDJ rearrangements showed no significant bias in use of D and J for cells using particular IgL isotypes. IgHV family use did not show significant bias either, but IgH V families III and VII with longer germline CDR1 lengths than other V families paired with lower frequency to Igσ sorted cells suggestive of steric hindrance causing IgH/IgL isotype pairing bias. Finally, flow cytometry suggested IgL isotype exclusion in spleen B cells of *X. laevis*. These findings will help pave the way for future studies towards discovering functions of IgL isotypes.

The applications of this study will help shed light on the origins and physiological role of mammalian Igλ and Igκ isotypes, their association with IgH, relation to heterodimeric paratopes, and role in IgL isotype associated diseases.

REFERENCES

1. Zhang YA, et al. IgT, a primitive immunoglobulin class specialized in mucosal immunity. Nat Immunol.2010;11:827-835.

2. Mashoof S, et al. Ancient T-independence of mucosal IgX/A: gut microbiota unaffected by larval thymectomy in *Xenopus laevis*. Mucosal immunology.2013;6:358-368.

3. Sun Y, Wei Z, Li N, Zhao Y. A comparative overview of immunoglobulin genes and the generation of their diversity in tetrapods. Dev Comp Immunol.2013;39:103-109.

4. Edholm ES, Wilson M, Bengten E. Immunoglobulin light (IgL) chains in ectothermic vertebrates. Dev Comp Immunol.2011;35:906-915.

5. Nemazee D, Martensson A, Verkoczy L. Haplotype exclusion and receptor editing: irreconcilable differences? SeminImmunol.2002;14:191-198.

6. Ota T, Sitnikova T, Nei M. Evolution of vertebrate immunoglobulin variable gene segments. CurrTopMicrobiolImmunol.2000;248:221-245.

7. Sverremark E, Rietz C, Fernandez C. Kappa-deficient mice are non-responders to dextran B512: is this unresponsiveness due to specialization of the kappa and lambda Ig repertoires? IntImmunol.2000;12:431-438.

8. Hsu E, Lefkovits I, Flajnik M, Du PL. Light chain heterogeneity in the amphibian *Xenopus*. MolImmunol.1991;28:985-994.

9. Hadji-Azimi I. Structural studies of the *Xenopus* 19S immunoglobulin and 7S immunoglobulin and two immunoglobulin-like proteins. Immunology.1975;28:419-429.

10. Hsu E, Du Pasquier L. Studies on *Xenopus* immunoglobulins using monoclonal antibodies. MolImmunol.1984;21:257-270.

11. Miracle AL, et al. Complex expression patterns of lymphocyte-specific genes during the development of cartilaginous fish implicate unique lymphoid tissues in generating an immune repertoire. IntImmunol.2001;13:567-580.

12. Criscitiello MF, Flajnik MF. Four primordial immunoglobulin light chain isotypes, including lambda and kappa, identified in the most primitive living jawed vertebrates. EurJImmunol.2007;37:2683-2694.

13. Du CC, Mashoof SM, Criscitiello MF. Oral immunization of the African clawed frog (*Xenopus laevis*) upregulates the mucosal immunoglobulin IgX. Veterinary immunology and immunopathology.2012;145:493-498.

14. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods.2001;25:402-408.

15. Livak KJ, et al. Methods for qPCR gene expression profiling applied to 1440 lymphoblastoid single cells. Methods.2013;59:71-79.

16. Rock EP, Sibbald PR, Davis MM, Chien YH. CDR3 length in antigen-specific immune receptors. J Exp Med.1994;179:323-328.

17. Haire RN, Amemiya CT, Suzuki D, Litman GW. Eleven distinct VH gene families and additional patterns of sequence variation suggest a high degree of immunoglobulin gene complexity in a lower vertebrate, *Xenopus laevis*. J Exp Med.1990;171:1721-1737.

18. Litman GW, et al. Evolutionary development of immunoglobulin gene diversity. AdvExpMedBiol.1991;292:11-17.

19. Hsu E, Schwager J, Alt FW. Evolution of immunoglobulin genes: VH families in the amphibian *Xenopus*. Proc Natl Acad Sci U S A.1989;86:8010-8014.

20. Biasini M, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic acids research.2014;42:W252-258.

21. Huston JS, et al. Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli. Proc Natl Acad Sci U S A.1988;85:5879-5883.

22. Kopp J, Schwede T. The SWISS-MODEL Repository: new features and functionalities. Nucleic acids research.2006;34:D315-318.

23. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics.2006;22:195-201.

24. Ohta Y, Flajnik M. IgD, like IgM, is a primordial immunoglobulin class perpetuated in most jawed vertebrates. ProcNatlAcadSciUSA.2006;103:10723-10728.

25. Zhao Y, et al. Identification of IgF, a hinge-region-containing Ig class, and IgD in *Xenopus* tropicalis. ProcNatlAcadSciUSA.2006;103:12087-12092.

26. Schwager J, Mikoryak CA, Steiner LA. Amino acid sequence of heavy chain from *Xenopus laevis* IgM deduced from cDNA sequence: implications for evolution of immunoglobulin domains. Proc Natl Acad Sci U S A.1988;85:2245-2249.

27. Schwager J, Grossberger D, Du Pasquier L. Organization and rearrangement of immunoglobulin M genes in the amphibian *Xenopus*. EMBO J.1988;7:2409-2415.

28. Mussmann R, Du PL, Hsu E. Is *Xenopus* IgX an analog of IgA? EurJImmunol.1996;26:2823-2830.

29. Du PL, Schwager J, Flajnik MF. The immune system of *Xenopus*. AnnuRevImmunol.1989;7:251-275.

30. Iacoangeli A, et al. Biased Immunoglobulin Light Chain Gene Usage in the Shark. J Immunol. 2015 Sep 4. pii: 1501426. [Epub ahead of print]

31. Schwager J, Burckert N, Courtet M, Du Pasquier L. The ontogeny of diversification at the immunoglobulin heavy chain locus in *Xenopus*. EMBO J.1991;10:2461-2470.

32. Desravines S, Hsu E. Measuring CDR3 length variability in individuals during ontogeny. J Immunol Methods. 1994 Feb 10 ;168(2):219-25.

33. Roman T, De Guerra A, Charlemagne J. Evolution of specific antigen recognition: size reduction and restricted length distribution of the CDRH3 regions in the rainbow trout. Eur J Immunol. 1995 Jan;25(1):269-73.

34. Lee A, Desravines S, Hsu E. IgH diversity in an individual with only one million B lymphocytes. Dev Immunol. 1993 ;3(3):211-22.

35. Flajnik MF, Deschacht N, Muyldermans S. A case of convergence: why did a simple alternative to canonical antibodies arise in sharks and camels? PLoS Biol.2011;9:e1001120.

36. Wang F, et al. Reshaping antibody diversity. Cell.2013;153:1379-1393.

37. Schwager J, et al. The ontogeny of diversification at the immunoglobulin heavy chain locus in *Xenopus.* EMBO J. 1991 Sep ;10(9):2461-70.

38. Cannon JP, et al. Variable domains and a VpreB-like molecule are present in a jawless vertebrate. Immunogenetics.2005;56:924-929.

39. Johansson J, et al. High variability in complementarity-determining regions compensates for a low number of V gene families in the lambda light chain locus of the platypus. Eur J Immunol. 2005 Oct ;35(10):3008-19.

APPENDIX A

FIGURES

Figure 1. IgH isotype message is not equal in cells sorted for IgL isotype. A. Quantification of enrichment of anti-IgL isotype MACS cells by qPCR. Note greater scale of fold enrichment on Igλ sort. **B.** Enrichment-normalized values of IgH isotype message in IgL isotype sorted cells. IgY/σ heterodimers were undetectable by traditional cloning of the sorted cells. Error bars denote standard error of the mean.

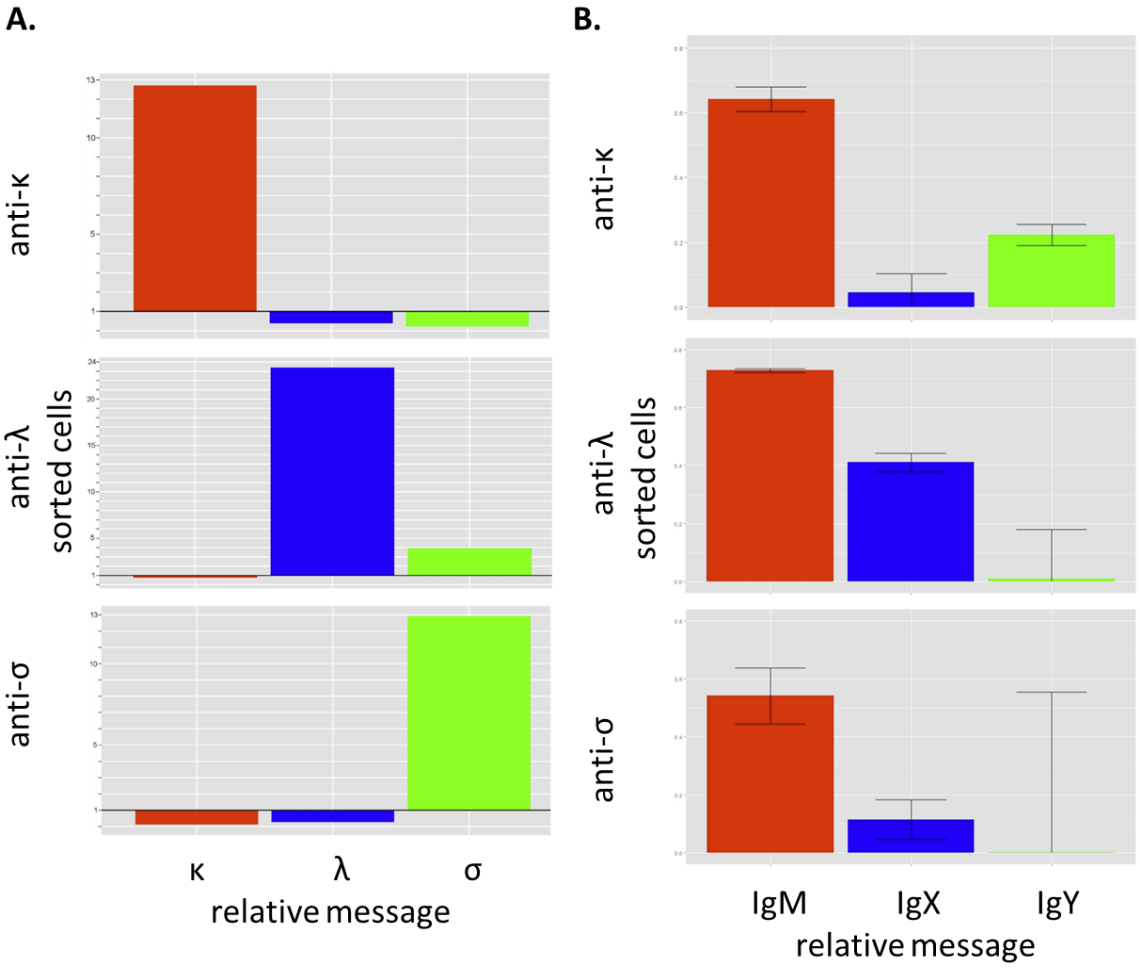
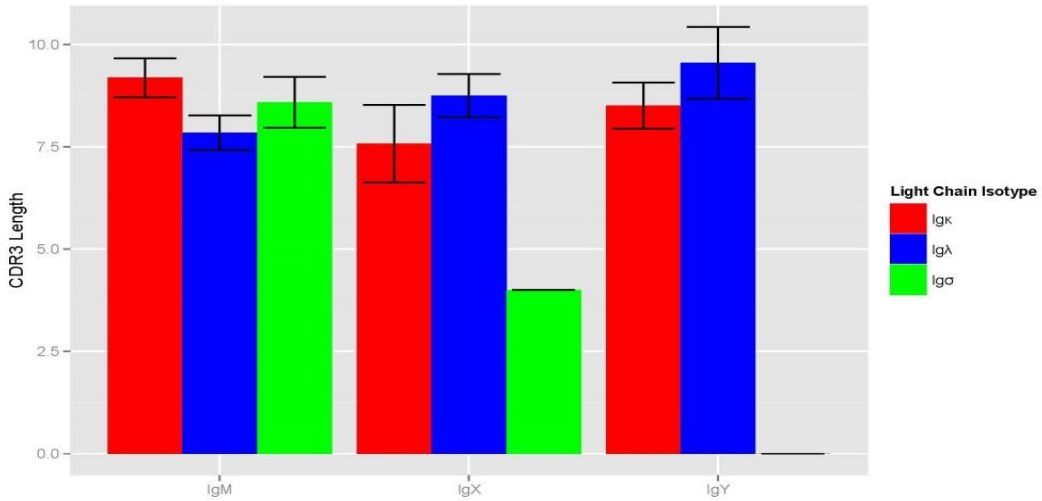


Figure 2. CDR3H lengths not significantly different amongst isotypes. A. CDR3 length in amino acids between conserved cysteine of YxC motif encoded by the V segment and first conserved glycine of GxG diglycine bulge encoded by the J, measured exclusively, minus four. Data averaged from 104 non-redundant clones in Figure 7, with the standard error of the mean. **B.** Number of N and P nucleotides contributing to the CDR3H, computed in Figure 8 from JGI *X. laevis* genome assembly 7.1.

A.



B.

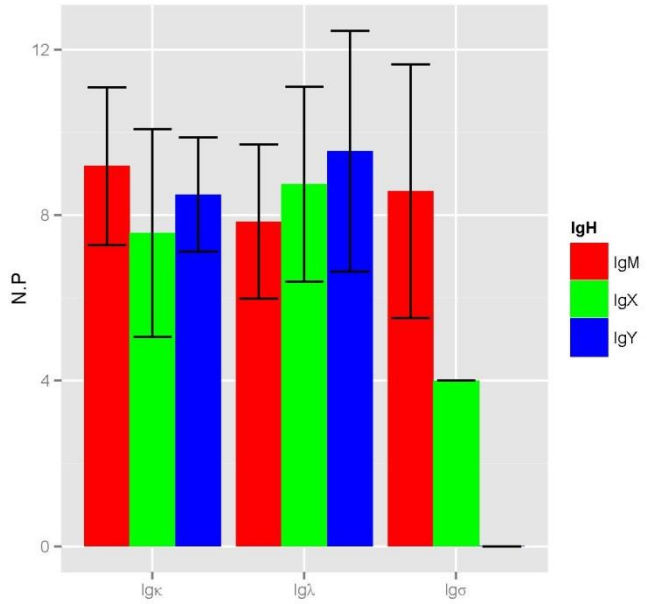
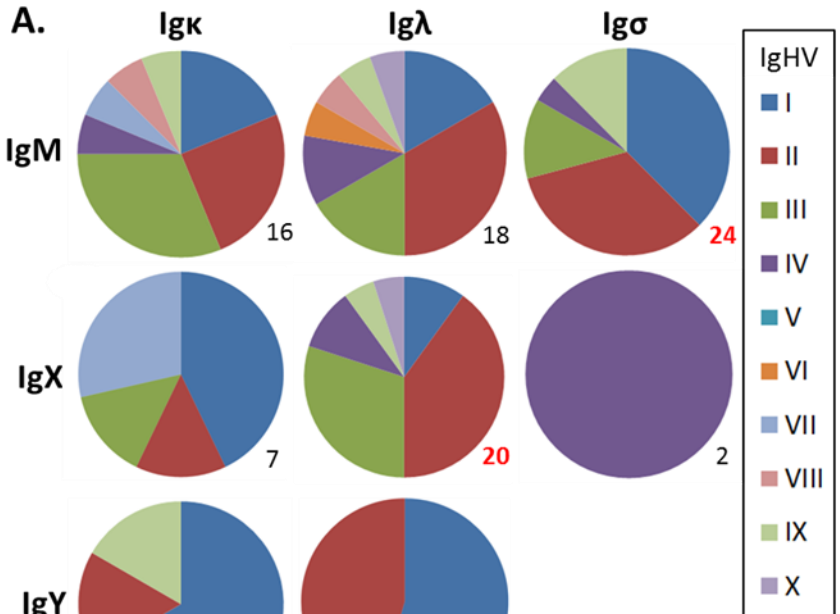


Figure 3. IgH V families III and VII used preferentially with Igκ, not Igσ. Pie chart shows percentage of clones employing the 11 different IgHV families of *X. laevis* in our clones (n for each isotope heterodimer shown to the bottom right of each chart, in red for biased pairings). **B.** Disparate lengths of CDR1H in VH III and VII versus other families. Conserved cysteine forming intradomain disulfide is highlighted in green of framework (FR) 1 as are conserved WYRQ and glycine of FR2. Shared residues in the long CDR1 of V III and VII are highlighted in yellow (adapted from (17)). IgH families V and XII were not expressed in this sampling.



B.

FAMILY	FR1	CDR1	FR2
III	TCKVTGASLT	DSSKIACVQ	WVRQPVG
VII	VCKVSGALIT	DGSKIHAVD	YIRQFSG
I	SCTASGFTFS	STWMY	WRVQTPG
II	TCTVSGFEIT	NYGLS	WIRQPPG
IV	SCKVSVSVTS	QHWA	WTWQAPG
VI	ACKVSVSVSS	YYWT	WVRQTPG
VIII	QCEVSGYNIN	DHHMH	WYRQAPG
IX	SCKTSGYTFT	DYWIH	WIQQVPG
X	TCRGSGFDFS	HYGMH	WIRQTPG

Figure 4. SCFV homology modeling supports CDRH1-CDRL2 steric hindrance constraining CDR3H. A. Space filling model of VH family III employing IgH from IgLσ sort shows crowding of CDR3 heavy by long Igσ CDR2 and VHIII CDR1. View of antigen binding surface from antigen perspective with coloring as follows: CDR1H red, CDR2H green, CDR3H blue, CDR1L cyan, CDR2L magenta, and CDR3L yellow. **B.** Trace backbone view of same heterodimer modeling with CDRs highlighted the same as in A. perspective from the IgHV side of the molecule rotated slightly towards the CDR2 side of IgL isotype.

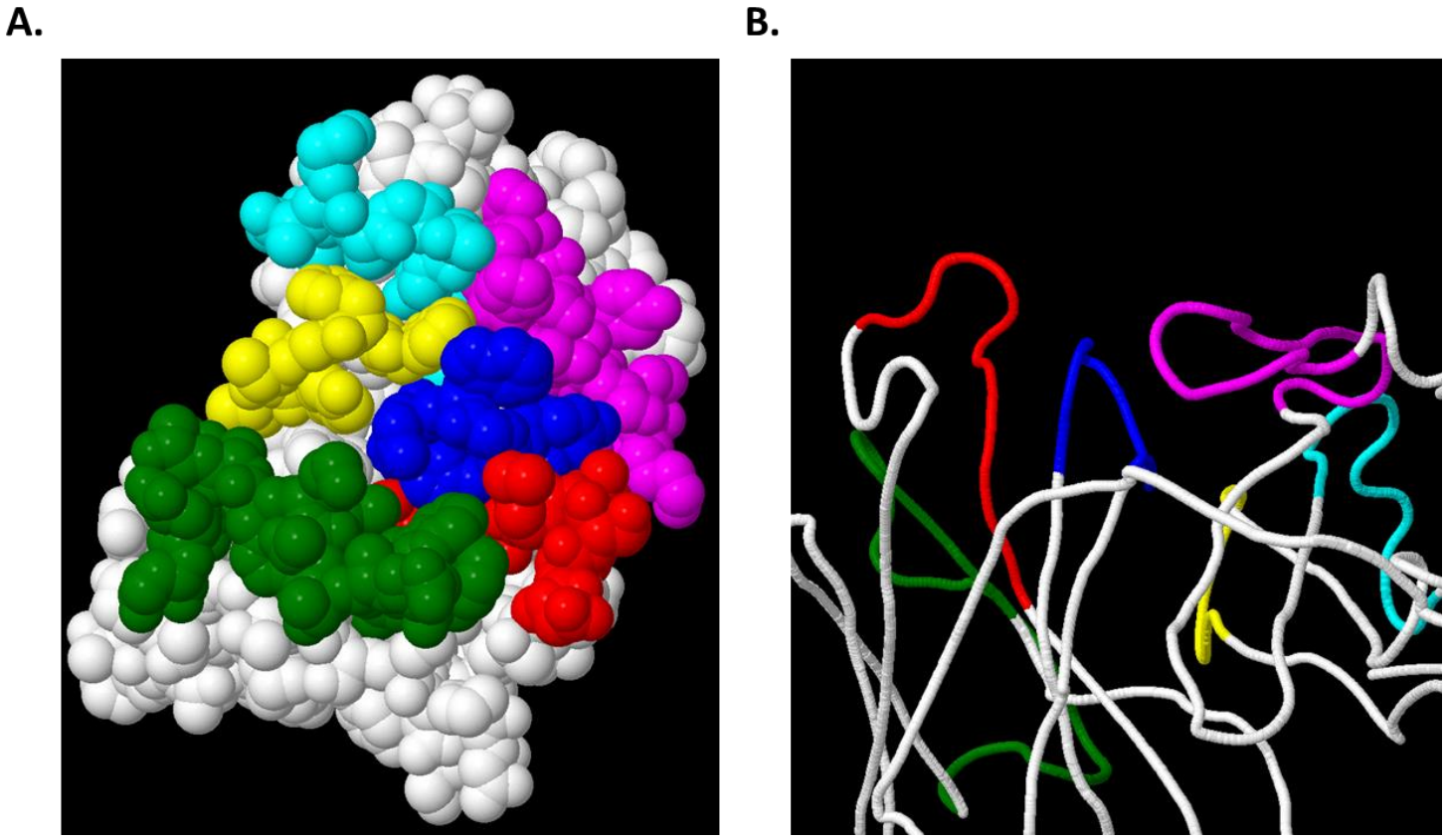
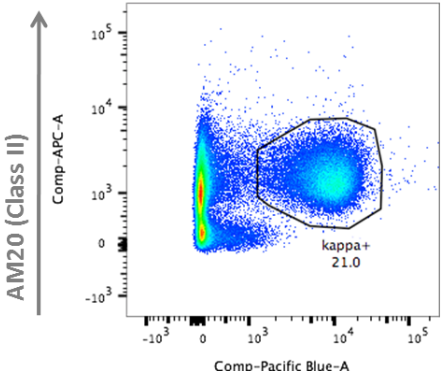
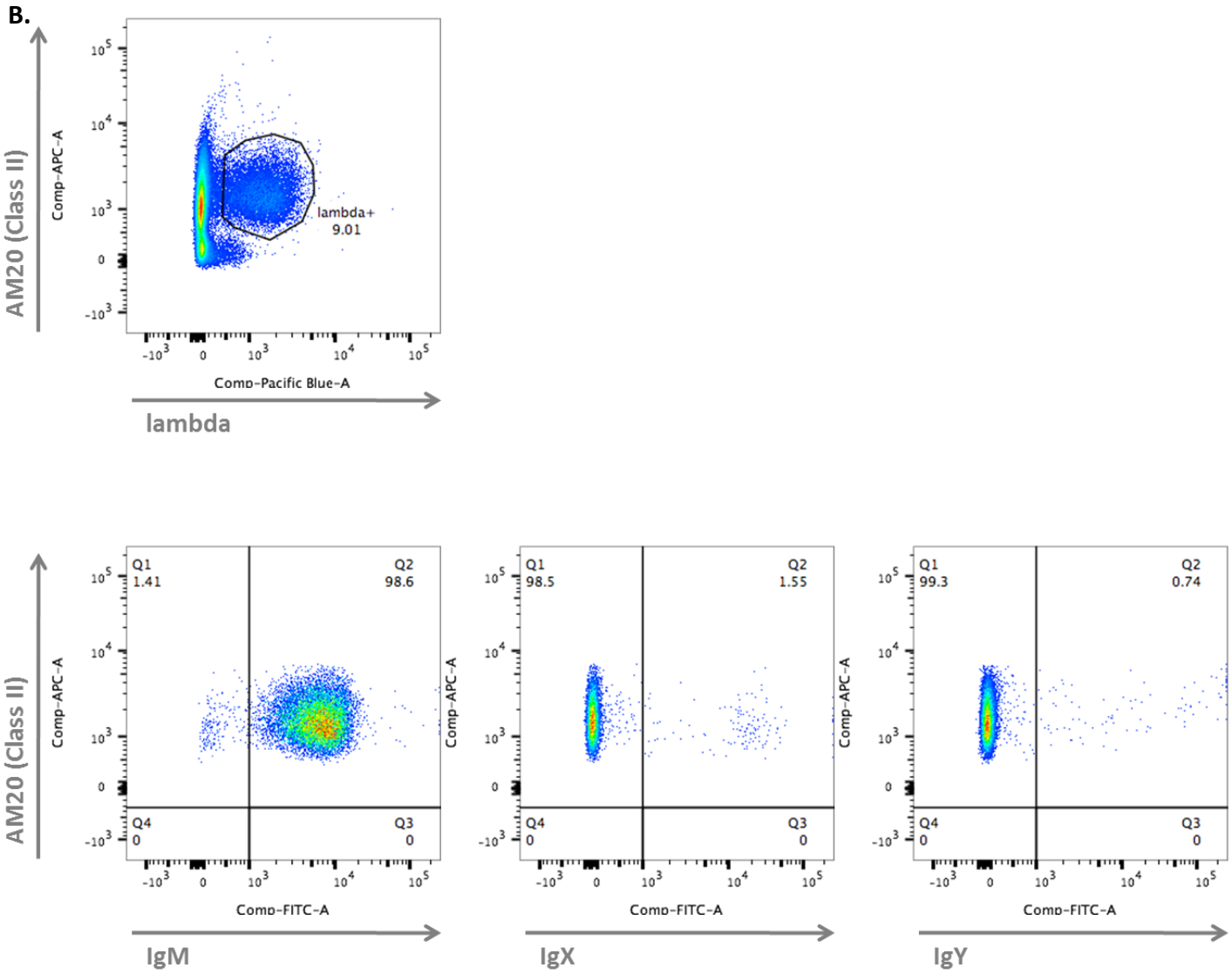
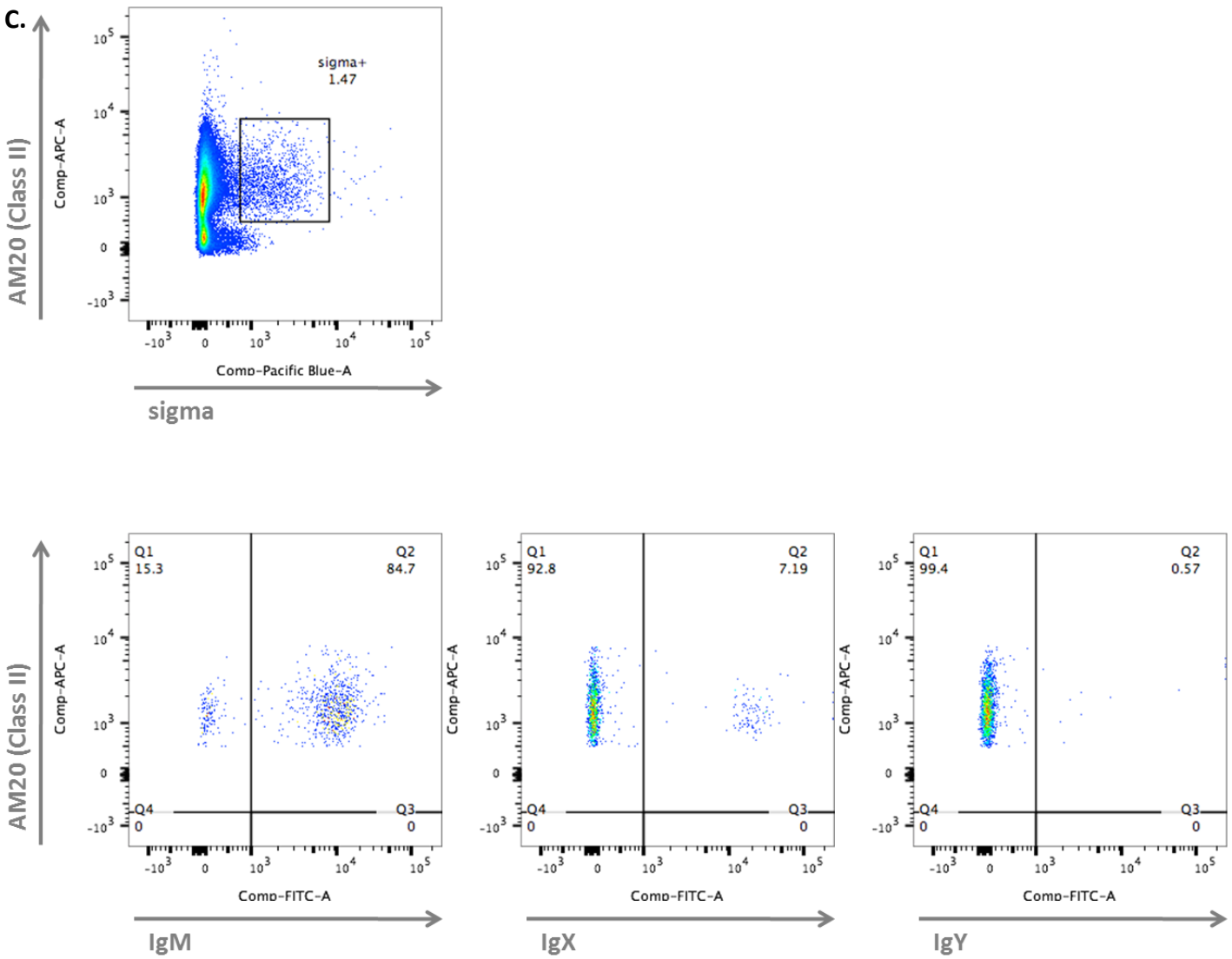


Figure 5. **Flow cytometry.** 5A-C are IgL vs. IgH isotype analysis gated for positive IgL isotype and analyzed for percent IgH isotype positive. 5E-G are the reciprocal analysis of IgL vs. IgH isotype gated for positive IgH isotype and analyzed for percent IgL isotype positive. 5D and H are bar graphs of the scatter plots and contour maps in 5A-C and 5E-F. 5I is a light scatter gating for live lymphocytes.



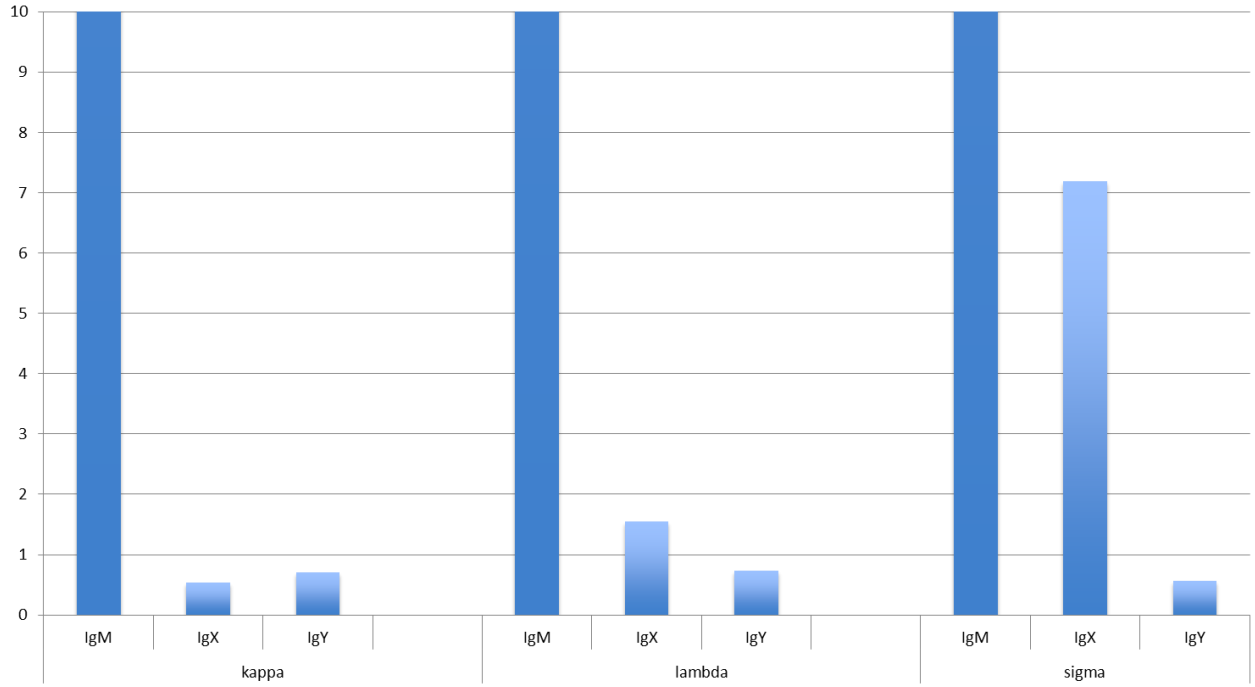
B.

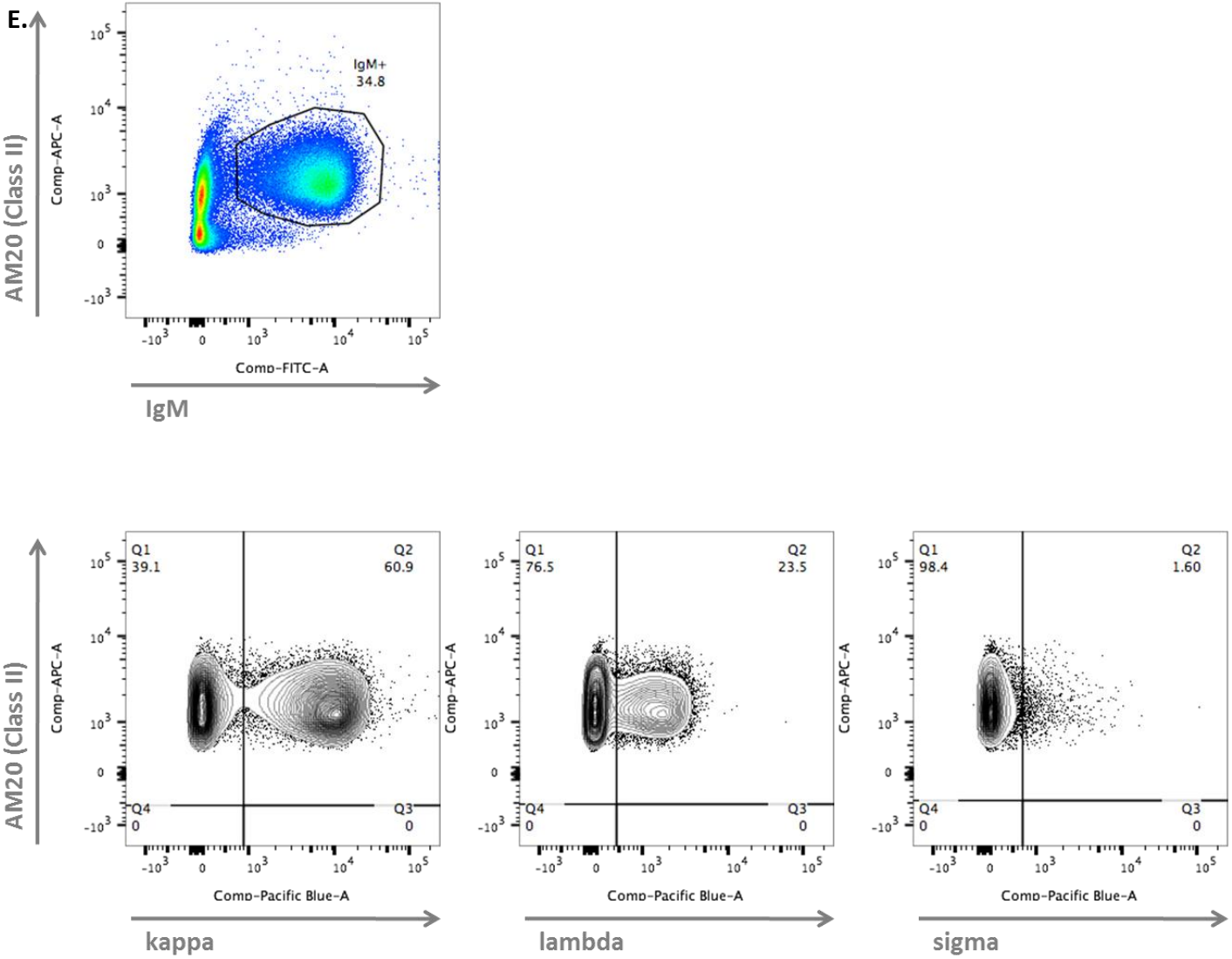

AM20 (Class II)
Comp-APC-A
lambda+
9.01
Comp-Pacific Blue-A
lambda
Q1
1.41
Q2
98.6
Q4
0
Q3
0
Comp-FITC-A
IgM
Q1
98.5
Q2
1.55
Q4
0
Q3
0
Comp-FITC-A
IgX
Q1
99.3
Q2
0.74
Q4
0
Q3
0
Comp-FITC-A
IgY

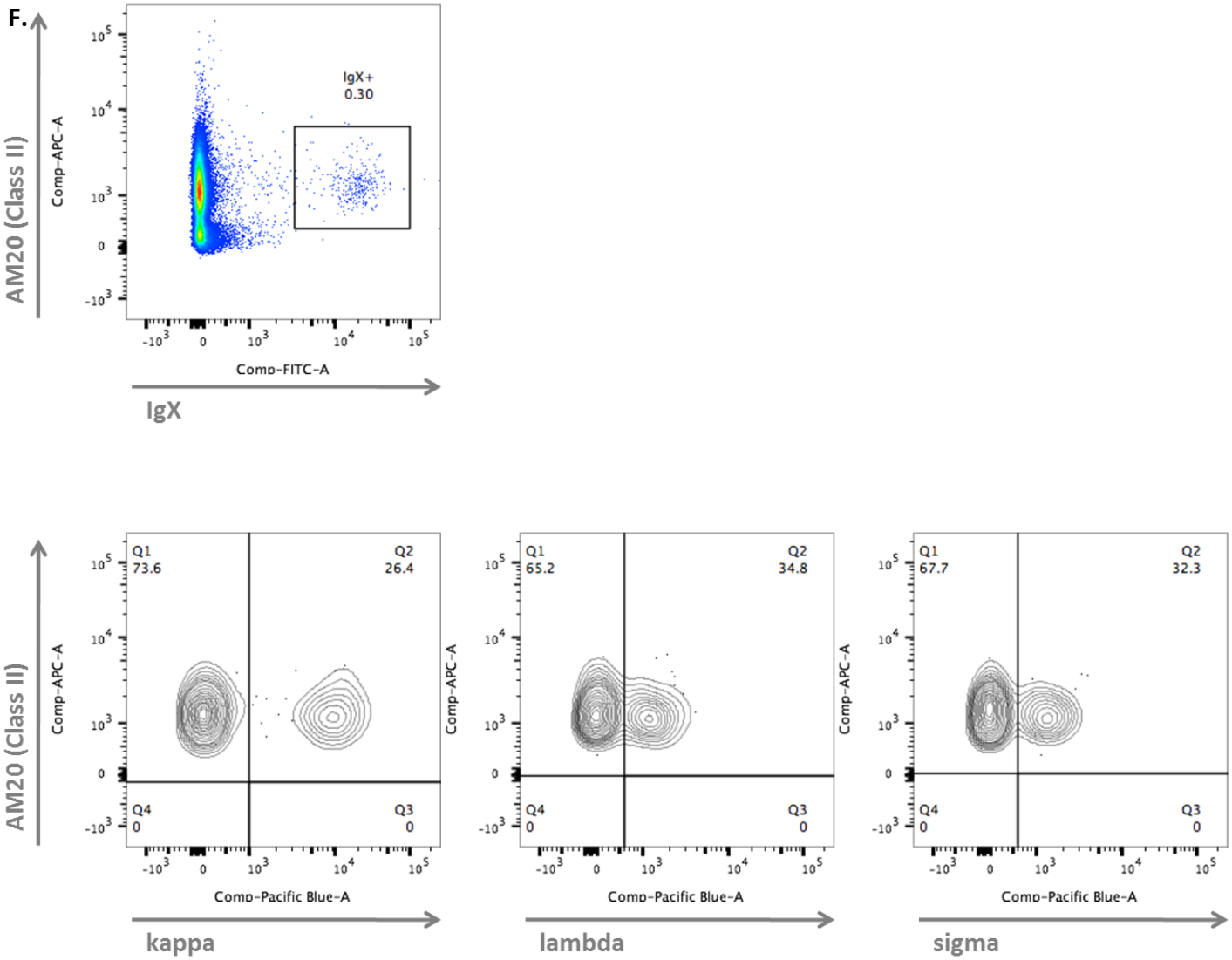

C.
AM20 (Class II)
Comp-APC-A
sigma+
1.47
Comp-Pacific Blue-A
sigma
Q1
15.3
Q2
84.7
Q4
0
Q3
0
Comp-FITC-A
IgM
Q1
92.8
Q2
7.19
Q4
0
Q3
0
Comp-FITC-A
IgX
Q1
99.4
Q2
0.57
Q4
0
Q3
0
Comp-FITC-A
IgY

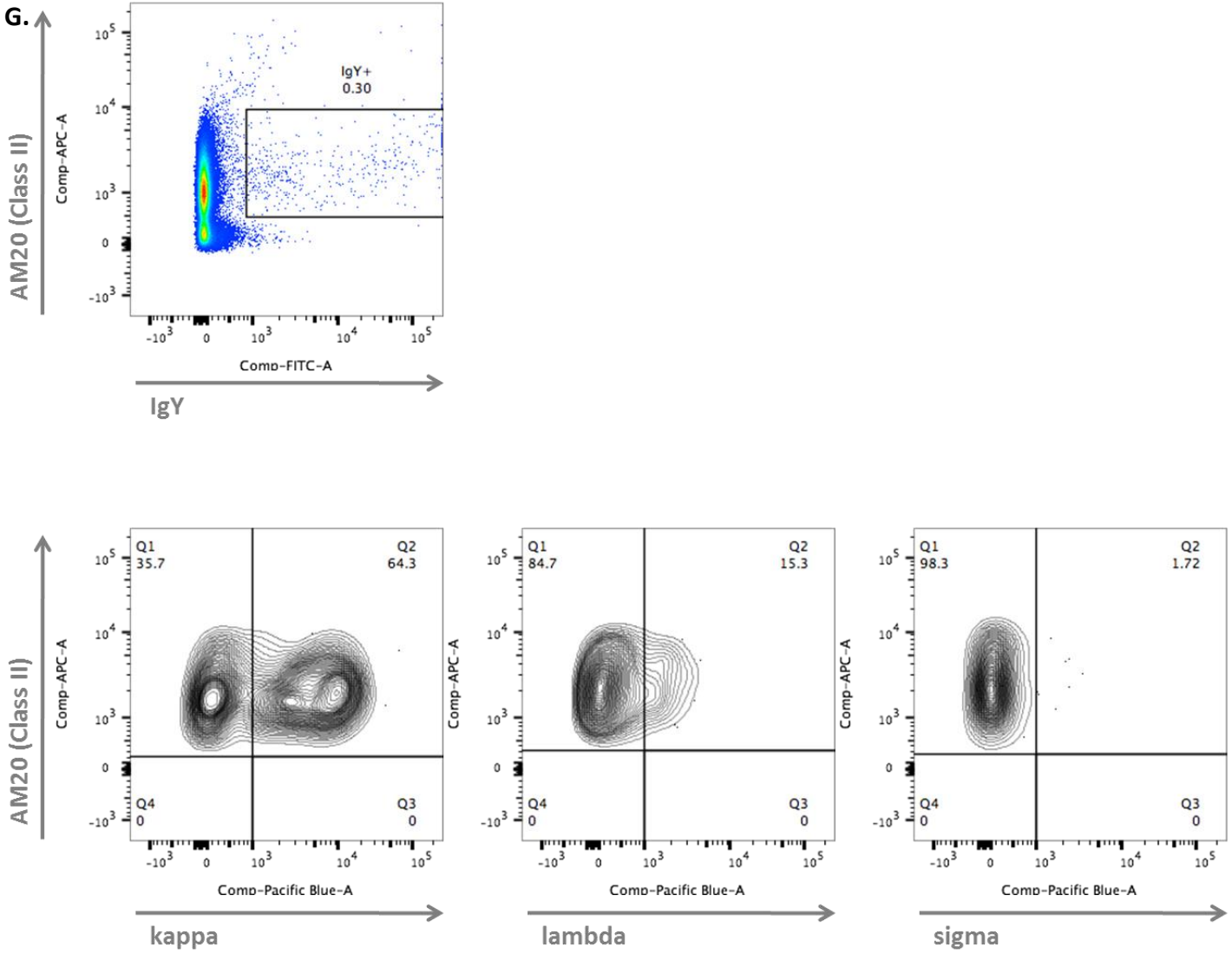
D.

kappa	IgM	98.3
	IgX	0.54
	IgY	0.71
lambda	IgM	98.6
	IgX	1.55
	IgY	0.74
sigma	IgM	84.7
	IgX	7.19
	IgY	0.57



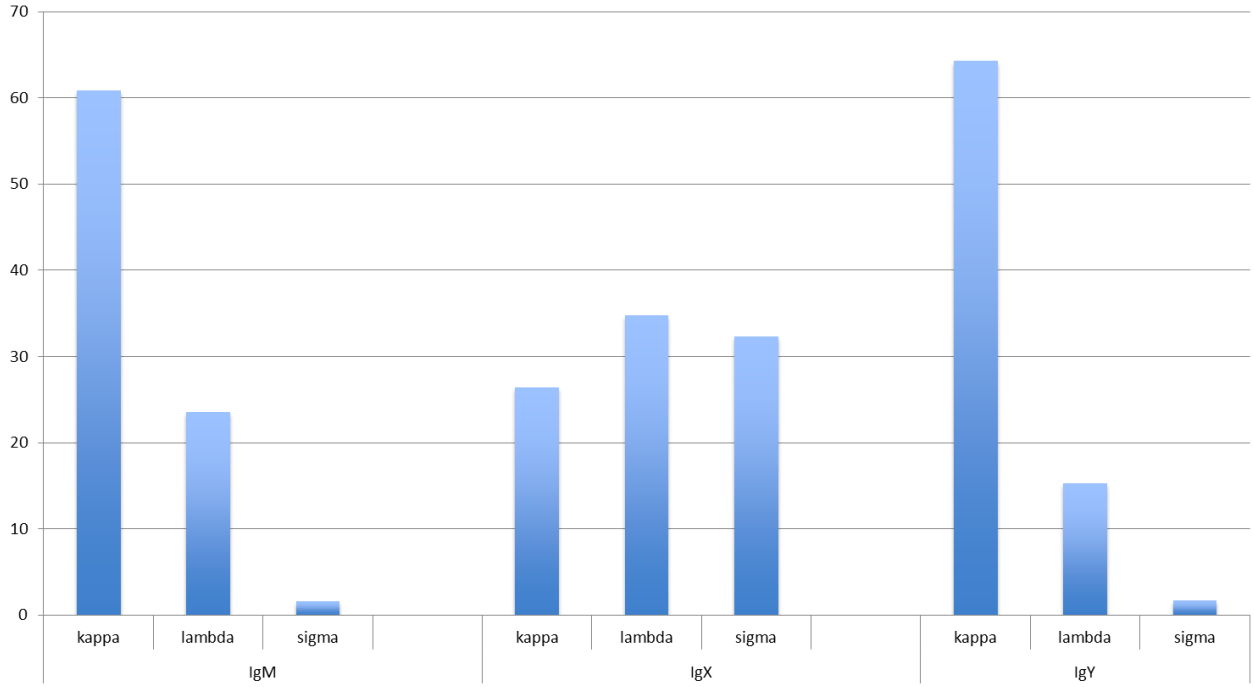

E.
AM20 (Class II)
Comp-APC-A
IgM+
34.8
Comp-FITC-A
IgM
Q1
39.1
Q2
60.9
Q4
0
Q3
0
Comp-Pacific Blue-A
kappa
Q1
76.5
Q2
23.5
Q4
0
Q3
0
Comp-Pacific Blue-A
lambda
Q1
98.4
Q2
1.60
Q4
0
Q3
0
Comp-Pacific Blue-A
sigma


F.
AM20 (Class II)
Comp-APC-A
IgX+
0.30
Comp-FITC-A
IgX
Q1
73.6
Q2
26.4
Q4
0
Q3
0
Comp-Pacific Blue-A
kappa
Q1
65.2
Q2
34.8
Q4
0
Q3
0
Comp-Pacific Blue-A
lambda
Q1
67.7
Q2
32.3
Q4
0
Q3
0
Comp-Pacific Blue-A
sigma


G.
AM20 (Class II)
Comp-APC-A
IgY+
0.30
Comp-FITC-A
IgY
Q1
35.7
Q2
64.3
Q4
0
Q3
0
Comp-Pacific Blue-A
kappa
Q1
84.7
Q2
15.3
Q4
0
Q3
0
Comp-Pacific Blue-A
lambda
Q1
98.3
Q2
1.72
Q4
0
Q3
0
Comp-Pacific Blue-A
sigma

H.

IgM	kappa	60.9
	lambda	23.5
	sigma	1.6
IgX	kappa	26.4
	lambda	34.8
	sigma	32.3
IgY	kappa	64.3
	lambda	15.3
	sigma	1.72



I.

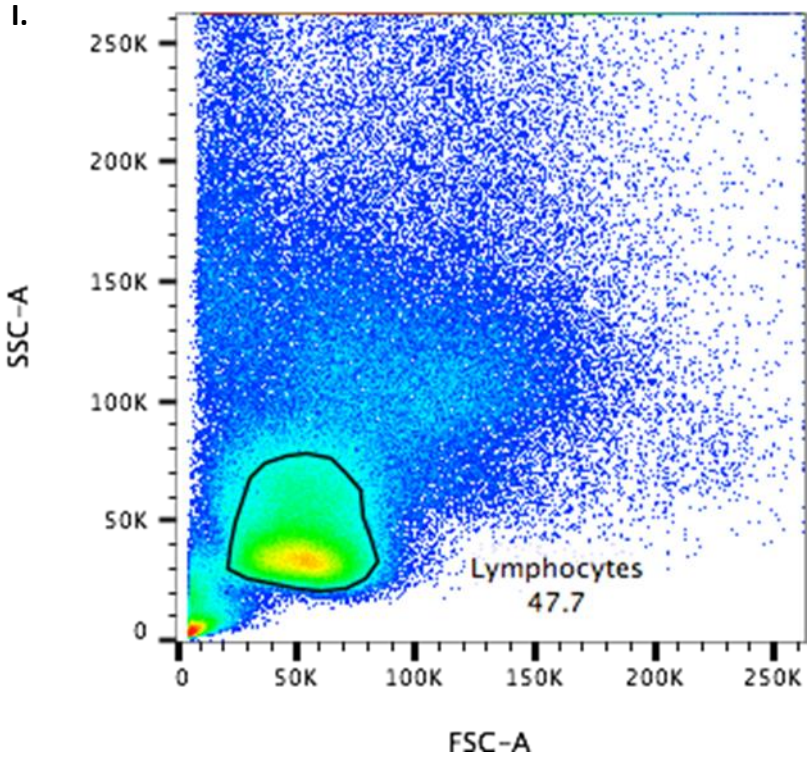

250K
200K
150K
100K
50K
0
SSC-A
0
50K
100K
150K
200K
250K
FSC-A
Lymphocytes
47.7

Figure 6. Model of development of isotype skewing. This model shows the bias in a IgL isotype for IgH isotypes. Each IgL isotype is given a different color (λ green, κ blue, σ red). Each IgH is represented by a circle. The circles decrease in size to show a decrease in binding preference of the IgL isotype to that IgH. λ, κ, and σ have the highest affinity for IgM in the primary lymphoid tissue (blue box). In the secondary lymphoid tissue (grey box), after exposure to antigens (labeled Ag), IgL isotypes indicate a bias for the second circle over the third circle. The yellow x is placed over the arrow between IgX and IgY because we failed to recover a single clone for the σ sorted IgY pairing.

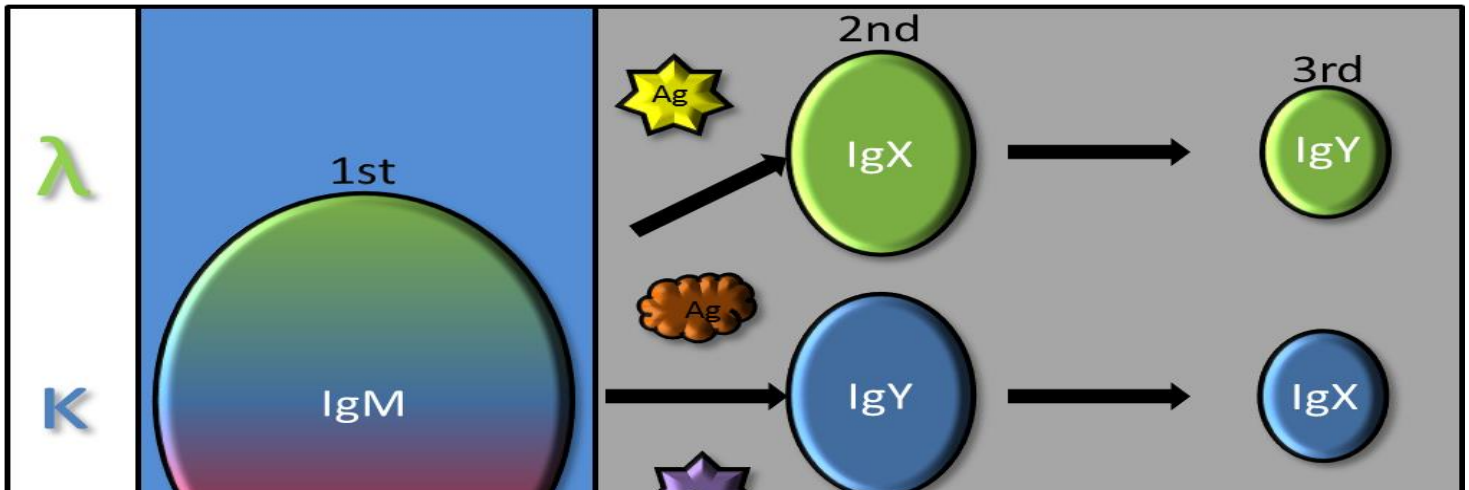


Figure 7. Unique CDR3 amino acid sequences cloned from IgL isotype sorts. Clone names followed by IgHV family, amino acid sequence with partial aligned V, J and C encoded sequences, followed by the J family at the far right. Conserved cysteine of YxC motif of V and first conserved glycine of the diglycine bulge encoded by the J segment highlighted in grey. Amino acids predicted to be encoded by at least two nucleotides of D segments highlighted magenta for D5, yellow for D1, green for D2, blue for D6, and red for D4. Clones with the same CDR3 as those in alignment are indicated under each set, somatic hypermutations do cause mutations outside of CDR3 in these sequences.

```
IgM
IgL κ sort                                                                  <IgM C
510719   III QSLQGRITVSRDTNKGEVYLKLTGMKPEETAVYYCA   REALWSGV   YYAFDYWGAGTMVTVTSATSNPPSLF J5
520719  VIII ISESFKDRVTPSTSGSTAQLRINKLSSSDTATYYCAR    GAYGG     YDFAYWGQGTMVTVTSATSNPPSLF J2
020729   III QNLQGRITVSRDTNKGEVYLKLTGMKPEETAVYYC     TAEALAG     PFDYWGQGTMVTVTSATSNPPSLF J7
040729     I DSVKGRFTISKDNNNNKLYLQMNNLQTEDTAVYYCAS   DLHWGGS    YAFDYWGAGTMVTVTSATSNPPSLF J5
050729   III QSLQGRITVSRDTNKGEVYLKLTGMKPEETAVYYC     TGRTLAGS     FDYWGQGTMVTVTSATSNPPSLF J7
060729    IX PSYQGRCHISTDNSQGTAFLQLNNLKVEDTAMYYCAR    DLGWEG      FAYWGQGTMVTVTSATSNPPSLF J2
070729    II ADSLNRVTITKDNGKKQVYLQMTGMEVKDTAMYYCAR     EGN        GDYWGQGTMVTVTSATSNPPSLF J7
090729   VII PDLKSRLTLSRDTAKNEAYLEISGMTAGDTAMYYCAK   HGGVTEG     YFEHWGQGTMVTVTSATSNPPSLF J8
100729     I DSFKGRFTISRDNNNNKLYLQMNNLQTEDTVVYYCAR   DMGSTSG    YWFDYWGQGTMVTVTSATSNPPSLF J7
110729    IV SSFQSRVTFTRDTSKNEIYLQMTSMKSEDSGTYYCA     ISLG    DYAYFDIWGPGTTVTVTSATSNPPSLF J1
130729   III QTLQGRITVSRDTNKGEVYLKLTGMKQEETAVYYCA     GAGVA     YYFDYWGQGTMVTVTSATSNPPSLF J7
170729     I DSVKGRFTISRDNNNNKLYLQMNNLQTEDTAVYYCTR    YIPASP      FDYWGQGTMVTVTSATSNPPSLF J7
180729    II DSLKNRVTITRDTGKKQVYLQMTGMEVKDTAMYYCARD   LGVGA       FAYWGQGTMVTVTSATSNPPSLF J2
011202    II DSLKSRVTITRDTGKKQVYLQMNGMEVKDTAMYYCAR    EGLEWV      FDYWGQGTMVTVTSATSKSPSLF J7
021202    II DSLKNRVTITKDNGKKQVYLQMTGMEVKDTAMYYCAR   VGWGGSS     AFDYWGAGTMVTVTSATSNPPSLF J5
031202   III QSLQGRITVSRDTNKGEVYLKLTGMKPEETAVYYCARE   APASG    YYAFDYWGAGTMVTVTSATSNPPSLF J5

040729 is the same as:
080729            041202

090729 is the same as:
120729            160729

020729 is the same as:
190729
180729 is the same as:
```

IgM
IgL λ sort

```
                                                                                    <IgM C
020724    IX PSYQGRCHISTDNSQSTAFLQLNNLKVEDTAMYYCAR              /GDYWGQGTMVTVTSATSNPPSLF J7
031113  VIII ISESFKERVTPSTSGSTAQLRISKLSSSDTATYYCAT      VRGP     YFAYWGQGTMVTVTSATSKSPSLF J2
041113    IV PSFQSRITLTRDTSKNEISLQMTSMKSADSGTYYCAR     YRGVA     HFDYWGQGTMVTVTSATSKSPSLF J7
071113     X KSVEGRLVITRNNAEQVTFMELKNLVYQDTAVYYCTR      RI      NAFDYWGAGTMVTVTSATSKSPSLF J5
091113    IV PSFQSRVTLTRDTSKNEISLQMTSMKSADSGTYYCTR     YRGVA     HFDYWGQGTMVTVTSATSKSPSLF J7
111113    IX PSYQGRCHISTDNSQSTGFLQLNNLKVEDTAMYYCARS    SGVGAY     FEHWGQGTMVTVTSATSKSPSLF J8
141113    VI PAFQNRVTLTRDTAKNEIYLAVSSMRSEDSGTYYCA      ALAGGP     FAYWGQGTMVTVTSATSNPPSLF J2
131121     I DSVKGRFTISRDNNNNNLYLQMNNLQTEDTAVYYCARD     RGGS    GYFEHWGQGTMVTVTSATSNPPSLF J8
191121    II DSLKNRVTITKDNGKKQVYLQMTGMEVKDTAMYYCARE    IGGVTA     FDYWGQGTMVTVTSATSNPPSLF J7
241121   III TTVQGRLTLSRDTNKGEVYFKLTEAKTEESATYYCAR     LAHWGG   NWFDYWGQGTMVTVTSATSKSPSLF J7
011125    II DSLKSRVTITRDTGKKQVYLQMNGMEVKDTAMYYCAR      YRGT     SFAYWGQGTMVTVTSATSKSPSLF J2
021125    II DSLKNRVTITKDNGKKQVYLQMTGMEVKDTAMYYCARD     TEVQ      LDYWGQGTMVTVTSATSKSPSLF J7
061125    II DSLKNRVTITKDNGKKQVYLQMTGMEVKDTAMYYCARD     ASGY    RYFEHWGQGTMVTVTSATSKSPSLF J8
121125    II DSLKNRVTITKDNGKKQVYLQMTGMEVKDTAMYYCARD    PLAGT    AWFDYWGQGTMVTVTSATSNPPSLF J7
131125     I DSFKGRFTLSRDNNNNKLYLQMNNLQTEDTAVYYCARD     KGN      HFDYWGQGTMVTVTSATSKSPSLF J7
181125     I DSFKGRFTISRDNNNNNLYLQMNNLQTEDTAVYYC       KWSGSG    YFDYWGQGTMVTVTSATSNPPSLF J7
031126    II DSLKNRVTITKDNGKKQVYLQMTGMEVKDTAMYYC      VRGSRYSS   DFAYWGQGTMVTVTSATSNPPSLF J2
061126   III QSLQGRITVSRDTNKGEVYLKLTGMKQEETAVYYCA      SCRYG    NWFDYWGQGTMVTVTSATSKSPSLF J7
121126   III TTVQGRLTLSRDTNKGEVYFKLTEAKTEESATYYCARQ    RLAGTAG    FAYWGQGTMVTVNSATSKSPSLF J2
```

020724 is the same as:

030724	040724	060724	080724	090724	100724
110724	120724	130724	140724	150724	160724
170724	180724	190724	200724	230724	270724
280724	290724	300724			

041113 is the same as:

101113	081121	141121	171121	201121	211121
141125	161125	211125	221125	161126	201126

071113 is the same as:

051121	151121	041125	191126

091113 is the same as:

161121	221121	031125	071125	081125

021125 is the same as:
091125 151125

011125 is the same as:
111125

131125 is the same as:
231125

241121 is the same as:
181126 241126

121125 is the same as:
221126

IgM

IgL σ sort <IgM C

```
021009    II DSQKSRVTITRDTGKKQVYLQMTGMEVKDTAMYYCA    EAYGGGRG   NWFDYWGQGTMVTVTSATSNPPSLF J7
031009    b  PSYQGRCHISTDNSQSTAFLQLNNLKVEDTAMYYCAR              /GDYWGQGTMVTVTSATSNPPSLF J7
071009    IX PSYQGRCHISTDNSQSTGFLQLNNLKVEDTAMYYCAR      A      SHFDYWGQGTMVTVT*ATSNPPSLF J7
081009    II DSLKNRVTITKDNGEKQVYLQMTGMEVKDTAMYYCAR    YAGGTY    NAFDYWGAGTMVTVTSATSKSPSLF J5
111009    II DSLKNRVTITKDNGKKQVYLQMTGMEVKDTAMYYCA     KGYGG    YNGFDYWGAGTMVTVTSATSNPPSLF J5
151009   III QSLKGRITVSRDTNKGEVYLKLTGMKPEETAVYYCA    RGLTGVG      LAYWGQGTMVTVTSATSNPPSLF J2
191009    II DSLKNRVTITKDNGKKQVYLQMTGMEVKDTAMYYCARE    GGSF      AFDYWGAGTMVTVTSATSNPPSLF J5
241009     I DSVKGRFTISRDNNNNKLYLQMNNLQTEDTAVYYCAR    PTSGYP      FAYWGQGTMVTVTSATSNPPSLF J2
011001     I DSVKGRFTISKDNNNNKLYLQMNNLQTEDTAVYYCTRE    GGGS      WFDYWGQGTMVTVTSATSNPPSLF J7
011014                                          LFLSVWGT     GAFDYWGAGTMVTVTSATSNPPSLF J5
041014    II DSLKNRVTITRDTGKKQVYLQMTGMEVKDTAMYYCAR    PYASG    YYAFDYWGAGTMVTVTSATSNPPSLF J5
111014     I DSVKGRFTISRDNNNNKLYLQMNNLQTEDTAVYYCAR     SPGV    YYAFDYWGAGTMVTVTSATSNPPSLF J5
121014     I DSVKGRFTISRDNNNNKLYLQMNNLQTEDTAVYYCAT      ED   IYYAYFDIWGPGTTVTVTSATSNPPSLF J1
131014     I DSVKGRFTISRDNNNNKLYLQMNNLQTEDTAVYYCAT    GWGSNS     YFEYWGQGTMVTVTSATSNPPSLF J8
161014    IX PSYQGRCHISTDNSQSTGFLQLNNLKVEDTAMYYCAR      GGI     YAFDYWGAGTMVTVTSATSNPPSLF J5
191014     I DSVKGRFTISRDNNNNKLYLQMNNLQTEDTAVYYCARD   GAWGSK    DYFDYWGQGTMVTVTSATSNPPSLF J7
201014    IX PSYQGRCHISTDNSQSTAFLQLNNLKVEDTAMYYCAR              /YWGQGTMVTVTSATSN/PSLF J7
221014     I DSVKGRFTISRDNNNNKLYLQMNNLQSEDTAVYYCTR   GGVASGYA  YAYFDIWGPGTTVTVTSATSNPPSLF J1
011016    II DSLKNRVTITRDTGKKQVYVQMTGMEVKDTAMYYCA      TGVGG    AYFDIWGPGTTVTVTSATSNPPSLF J1
```

```
021016   III TTVEERLTLSRDPNKGEVYFKLTEARTEESATYYCARH    ARGWQ     NFDYWGQGTMVTVTSATSNPPSLF J7
081016    II DSLKNRVTITKDNGKKQVYLQMTGMEVKDTAMYYCARD   EERGGSA       YWGQGTMVTVTSATSNPPSLF J2
091016     I DSVKGRFTISRDNNNNKLYLQMNNLQTEDTAVYYCAR     WAGVG   DAYFDIWGPGTTVTVTSATSNPPSLF J1
101016   III TTVEERLTLSRDPNKGEVYFKLTEARTEESATYYCARH    ARGWQ     DFDYWGQGTMVTATSATSNPPSLF J7
201016     I DSVKGRFTISKDNNNNKLYLQMNNLQTEDTAVYYCAS    DLHWGGS   YAFDYWGAGTMVTVTSATSNPPSLF J5
241016    IV SSFQSRVTFTRDTSKNEIYLQMTSMKSEDSGTYYCAR     YASGY   RRAFDYWGAGTMVTVTSATSNPPSLF J5
121009                                                 GERG     AFDYWGAGTMVTVTSATSNPPSLF J5
```

031009 is same as:

041009	051009	061009	091009	101009	131009
141009	171009	181009	201009	211009	231009
031014	051014	141014	211014	241014	021014
111016	121016	151016	181016	211016	

241009 is same as:

011002	031001	051001

081009 is same as:

010924	020924	030924	040924	050924	060924
070924	080924	090924	100924	110924	120924
130924	140924	150924	160924	170924	180924
190924	200924	210924	220924	230924	240924

011014 is the same as:

061014	071014	091014	101014	151014	171014
181014	231014	031016	041016	071016	141016
171016	221016	161009	221009	011009	

221014 is the same as:

051016

191014 is the same as:

061016

161014 is the same as:

131016	231016

091016 is the same as:
161016

221014 is the same as:
191016

IgX
IgL κ sort

```
                                                                                                  <IgX C
150729   III QSLKGRITVSRDTNKGEVYLKLTGMKPDETAVYYCA       RAE       LDYWGQGTMVTVTSVTASAPSVF J7
011205     I DSVKGRFTITRDNNNNKLYLQMNNLQTEDTAVYYCAR     SSGV    NWYFEHWGQGTMVTVTSVTASAPSVF J8
021205     I DSVKGRFTISRDNNNNKLYLQMNNLQTEDTAVYYCTRD     SGG       FDYWGQGTMVTVTSVTASAPSVF J7
031205    II DSLKNRVTITRDTGKKQVYLQMNGIEVKDTAMYYCAR     GLRGV      FDYWGQGTMVTVTSVTASAPSVF J7
061205   VII PDLKSRLTLSRDTVKNEAYLEISGMTAGDTAMYYCAK   HGLLEWDYA    FDYWGAGTMVTVTSVTASAPSVF J7
091205     I DSVKGRFTISRDNNNNKLYLQMNNLQTEDTAVYYCATE     MGG       FAYWGQGTMVTVTSVTASAPSVF J2
101205   VII PGLKSRLTLSRDTAKNEDYLEISGMTAGDTAMYYCAKQ    FTGL    GSYFDYWGQGTMVTVTSVTASAPSVF J7
```

150729 is the same as:
160729 170729 180729

011205 is the same as:
041205 111205

021205 is the same as:
071205

061205 is the same as:
081205

101205 is the same as:
111205

IgX
IgL λ sort

```
                                                                                                  <IgX C
010701    II DLMKNRVKITKDNGKKEVYLQMTGMEVKDTAMYYCTR   TRTLTYTGK   WFDYWGQGTMVTVTSVTASAPSVF J7
020701    IV PSFQSRVTLSRDTSKNEISLQMTSMKSEDSGTYYCAR              HDFAYWGQGTMVTVTSVTASAPSVF J2
040701    II DTLKNRVTITRDTGKEQVYLQMNGMEVKDTAMYYCAR   YMFWSGTNV    FDYWGAGTMVTVTSVTASAPSVF J5
030727     X KSVEGRLVITRNNAEQVTFMELKNLVYQDTAVYYCTR     DSGVGV     FDYWGQGTMVTVTSVTASAPSVF J7
```

```
040727    II DSLKNRVTITKDTGKKQVYLQMNEMEVKDTAMYYCARD    RLGEV   TEAFDYWGAGTMVTVTSVTASAPSVF J5
280727   III QSLKGRITLSRDTNKGEVYLKLTGMKPEETAVYYCAR     DTVGP     AFDYWGAGTMVTVTSVTASAPSVF J5
050727   III TTVQGRLTLSRDTNKGEVYFKLTEAKTEESATYYCAR    YYVSGYK  YAYFDIWGPGTTVTVTSVTASAPSVF J1
080727     I DSVKGRLTISRDNNNNKLYLQMNNLQTEDTAVYYCTR     SGVGP     YFDYWGQGTMVTVTSVTASAPSVF J7
090727   III QSLQGRITVSRDTNKGEVYLKLTGMKREETALYYCTNY    RGGGTS   DYFDYWGQGTMVTVTSVTASAPSVF J7
140727    IX PSYQGRCHISTDNSQSTGFLQLNNLKVEDTAMYYCAR     KERVQP     FAYWGQGTMVTVTSVTASAPSVF J2
190727     I DSVKGRFTISRDNNNNKLYLQMNNLQTEDTAVHYC     AFFCSGSSG   TFDYWGAGTMVTVTSVTASAPSVF J5
260727   III YAMQGRLTLSRDTNKGEVYFKLTETKTEESATYYCARQ     TGVA    NYFDYWGQGTMVTVTSVTASAPSVF J7
300727    II DSLKNRVTITKDTGKKQVYLQMNEMEVKDTAMYYCAS      TGVG     DFAYWGQGTMVTVTSVTASAPSVF J2
030731    II DSLKNRVTITKDNGKKQVYLQMTGMEVKDTAMYYCAR     YAGGTY   NAFDYWGAGTMVTVTSATSKSPSLF J5
070731    II DSLKNRVTITKDNGKKQVYLQMTGMEVKDTAMYYCARD     TEVQ      LDYWGQGTMVTVTSATSKSPSLF J7
100731   III TAVQGRLTLSRDTNKGEVYFKLTEAKTEESATYYCAR     SLTGVA    HFAYWGQGTMVTVTSATSKSPSLF J2
130721    II DSLKSRVTITRNTGKKQVYLQMNGMEVKDTAMYYCAR     EPYGGY   NAFDYWGAGTMVTVTSATSKSPSLF J5
140731    II DTLKNRVTITRNTGKKQVYLQMNGMEVKDTAMYYCAR     YDWVGA    YFEHWGQGTMVTVTSATSKSPSLF J8
190731   III QNLQGRITVSRDTNKGEVYLKLTGMKPEETAVYYC       VGG      FDYWGQGTMVTVTSATSNPPSLF J7
230731    IV PSFQSRITLTRDTSKNEISLQMTSMKSADSGTYYCAR     YRGVA     HFDYWGQGTMVTVTSATSKSPSLF J7
```

020701 is the same as:
030701 070727

030727 is the same as:
020727 100727

040727 is the same as:
160727

050727 is the same as:
200727

140727 is the same as:
210727 220727

080727 is the same as:
270727

190727 is the same as:
010727

IgX

IgL σ sort

```
                                                                                    <IgX C
180919    IV PSFQSRVTLSRDTSKNEISLQMTSMKSEDSGTYYCARH                 DFAYWGQGTMVTVTSVTASAPSVF J2
091024    IV PSFQSRVTLSRDTSKNEISLQMTSMKSEDSGTYYCASH                 DFAYWGQGTMVTVTSVTASAPSVF J2
```

180919 is same as:

```
130919          140919          150919          160919          170919          010919
030919          060919          050919          070919          080919          021022
031022          011024          021024          031024          041024          051024
061024          071024          081024          101024          111024          121024
131024          141024          151024          161024
```

IgY

IgL κ sort

```
                                                                                    <IgY C
010719    IX SSYQGRCHISTDNSQSTAFLQLNNLKVEDTAMYYCA   TEAFGVGP    WYFEHWGQGTMVTVTSATLHAPSVF J8
020719     I DSLKGRFIISRDNNKNNLYLQMNNVQTEDTAVYYCTRE  VRLGVT       FDYWGAGTMVTVTSATLHAPSVF J7
060719     I DSLKNRVTITRDTGKKQVYLQMTGMEVKDTAIYYC     AKTEWDA      FDYWGAGTMVTVTSATLHAPSVF J7
100719     I DLIKGRFTISRDNNNNKLYLQMNNIQTEDTAVYYCARD     GD     HAYFDIWGPGTTVTVTSATLHAPSVF J1
160719    II DYLKNRVTITKDNVKKQVYVQMTGMEVKDTAMYYCAR    LGDWG      DFGYWGQGTMATVTSATLHAPSVF J2
290719     I DSVKGRFTISRDNNNNKLYLQMNNLQTEDTAVYYCTR     CGEG     AYFDIWGPGTTVTVTSATLHAPSVF J1
```

020719 is the same as:

```
030719          040719          050719          070719          080719          110719
120719          170719          190719          200719          210719          250719
280719          300719
```

100719 is the same as:

```
180719          240719
```

010719 is the same as:

```
260719
```

IgY
IgL λ sort

```
                                                                                    <IgY C
050701     II DYLKNRVTITKDNGKKQVYVQMTGMEVKDTAMYYCSR     LGDWG       DFAYWGQGTMVTVTSATLHAPSVF J2
060701      I DSLKGRFIISRDNNKNNLYLQMNNVQTEDTAVYYCTRE    VRLGVT       FDYWGAGTMVTVTSATLHAPSVF J7
070701      I DSLKGRFTISRDNNNNKLYLQMNNLQTEDTAGYYCAR    HVGTLAGP    NWFDYWGQGTMVTVTSATLHAPSVF J7
010709      I DSVKGRFTISRDNNNNKLYLQMNNLQTEDTAVYYCTRD   SGGSPYYA      FDYWGAGTMVTVTSATLHAPSVF J7
030709      I DSVKGRFTISRDNNNNKLYLEMNNLQTEDTAVYYCAS   QHRYGAPPAS   YDFAYWGQGTMVTVTSATLHAPSVF J2
040709     II DYLKNRVTITKDNVKKQVYVQMTGLEVKDTAMYYCAR      LGDWG      DFGYWGQGTMVTVTSATLHAPSVF J2
050709      I DSLKGRFTISRDNNNNKLYLQMNNLQTEDTAVYYCAR    HVGTPAGP    NWFDYWGQGTMVTVTSATLHAPSVF J7
070709      I DSLKGRFTISRDNNNNKLYLQMNNLQTEDTAVYYCTR              GWFDYWGQGTMVTVTSATLHAPSVF J7
080709     II DYLKNRVTITKDTGKKQVYLQMNGMEVKDTAMYYCTR       VS     YGNWFDYWGQGTMVTVTSATLHAPSVF J7
110709      I DSVKGRFTISRDNNNNKLYLQMNNLQTEDTAVYYCAR      LGLNA       FDYWGAGTMVTVTSATLHAPSVF J7
120709                                                    VGG         GQGTMVTVTSATLHAPSVF J7
180731*    II DSLKNRVTITKDNGKKQVYLQMTGMEVKDTAMYYCAS   TLAVTALLPA     FDYWGQGTMVTVTSATLHAPSVF J7
```

*cloned with IgX primer

070701 is the same as:
080701 020709

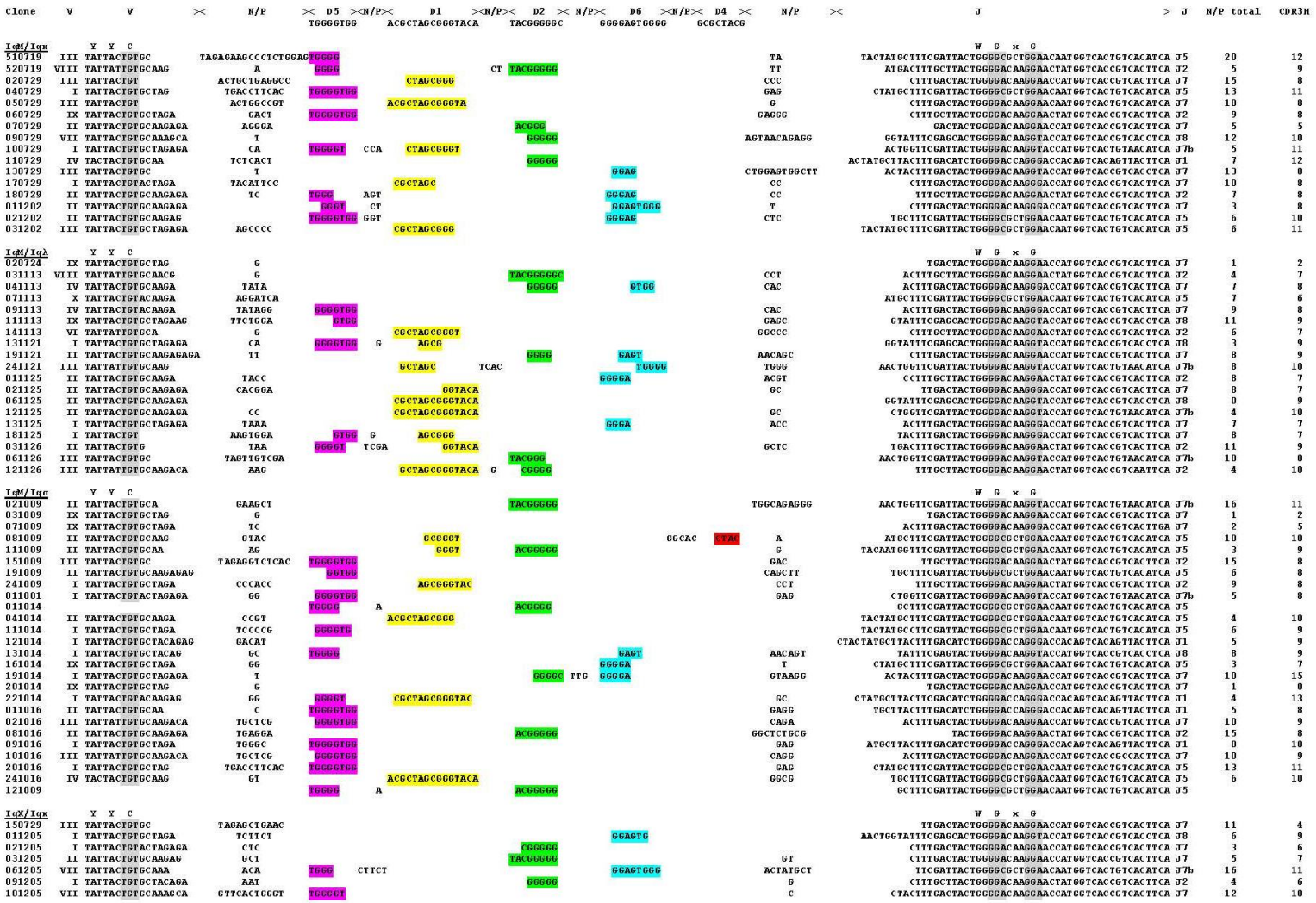
060701 is the same as:
060709 090709

040709 is the same as:
100709

IgY
IgL σ sort

none cloned

Figure 8. Nucleotide alignments of IgH isotype clone CDR3s.



Clone	V	V	><	N/P	><	D5	><N/P><	D1	><N/P><	D2	><	N/P><	D6	><N/P><	D4	><	N/P	><	J	>	J	N/P total	CDR3H
						TGGGGTGG		ACGCTAGCGGGTACA		TACGGGGGC			GGGGAGTGGGG		GCGCTACG								
IgX/Igλ		Y Y C																	W G x G				
010701	II	TATTACTGTACAAGA		ACGAGGACGTTAACATA						TACGGG							CAAG		TGGTTCGATTACTGGGGACAAGGTACCATGGTCACTGTAACATCA	J7b		21	12
020701	IV	TATTACTGTGCAAGACA																	TGACTTTGCTTACTGGGGACAAGGAACTATGGTCACCGTCACTTCA	J2		0	4
040701	II	TATTACTGTGCAAGA						TACATGTTC				TGGAGT	GGGA				CCAATGT		TTTCGATTACTGGGGCGCTGGAACAATGGTCACTGTCACATCA	J5		17	11
030727	X	TATTACTGTACAAGA		GATTCT									GGAGTGGGG				GT		CTTTGACTACTGGGGACAAGGGACCATGGTCACCGTCACTTCA	J7		8	8
040727	II	TATTACTGTGCAAGAGA		TAGACTT		GGTG	AAGTA			ACGG							AG		GCTTTCGATTACTGGGGCGCTGGAACAATGGTCACTGTCACATCA	J5		14	11
280727	III	TATTACTGTGCTAGA		GATACCGT							GGGGC						C		TGCTTTCGATTACTGGGGCGCTGGAACAATGGTCACTGTCACATCA	J5		9	8
050727	III	TATTATTGTGCAAGA		TACTACGT					TAGCG						GCTAC		AAA		TATGCTTACTTTGACATCTGGGGACCAGGGACCACAGTCACAGTTACTTCA	J1		11	12
080727	I	TATTACTGTACTAGA		TC		TGGGGTGG											GACC		CTACTTTGACTACTGGGGACAAGGAACCATGGTCACCGTCACTTCA	J7		6	8
090727	III	TATTACTGT		ACTAATTAT						CGGGGG		G	GTGGG				ACCTCGG		ACTACTTTGACTACTGGGGACAAGGGACCATGGTCACCGTCACTTCA	J7		17	11
140727	IX	TATTACTGTGCTAGAA		AGG					AGCGGGTACA								GCC		CTTTGCTTACTGGGGACAAGGAACTATGGTCACCGTCACTTCA	J2		6	8
190727	I	CATTACTGT		GCTTTTTTCTGCA		GTGG	GAGTT			CGGGG							A		CTTTCGATTACTGGGGCGCTGGAACAATGGTCACTGTCACATCA	J5		19	10
260727	III	TATTATTGTGCAAGACA		GA						CGGGG			GTGG				CAA		ACTACTTTGACTACTGGGGACAAGGGACCATGGTCACCGTCACTTCA	J7		5	9
300727	II	TATTACTGTGCAAG		TAC		TGGGGTGG											GC		GACTTTGCTTACTGGGGACAAGGAACTATGGTCACCGTCACTTCA	J2		5	7
030731	II	TATTACTGTGCAAG		GTAC					GCGGGT	GGCAC					CTAC		A		ATGCTTTCGATTACTGGGGCGCTGGAACAATGGTCACTGTCACATCA	J5		10	10
070731	II	TATTACTGTGCAAGAGA		CACGGA					GGTACA								GC		TTGACTACTGGGGACAAGGGACCATGGTCACCGTCACTTCN	J7		8	7
100731	III	TATTATTGTGCAAGA		TCCCT						TACGGGG			GTGG				CTC		ACTTTGCTTACTGGGGACAAGGAACTATGGTCACCGTCACTTCA	J2		8	9
130731	II	TATTACTGTGCAAG		GGAGCCC						TACGGGG					GCTAC		A		ATGCTTTCGATTACTGGGGCGCTGGAACAATGGTCACTGTCACATCA	J5		8	10
140731	II	TATTACTGTGCAAGA		TACGAC		TGGG							TGGG				AGC		GTATTTCGAGCACTGGGGACAAGGTACCATGGTCACCGTCACCTCA	J8		9	9
190731	III	TATTACTG		C		GTGG											GAGG		CTTTGACTACTGGGGACAAGGAACCATGGTCACCGTCACTTCA	J7		5	3
230731	IV	TATTACTGTGCAAGA		TATAGGG		GGGTGG											CAC		ACTTTGACTACTGGGGACAAGGGACCATGGTCACCGTCACTTCA	J7		10	8
IgX/Igσ		Y Y C																	W G x G				
180919	IV	TATTACTGTGCAAGACA																	TGACTTTGCTTACTGGGGACAAGGAACTATGGTCACCGTCACTTCA	J2		0	4
091024	IV	TATTACTGTGCAAGTCA																	TGACTTTGCTTACTGGGGACAAGGAACTATGGTCACCGTCACTTCA	J2		0	4
IgV/Igκ		Y Y C																	W G x G				
010719	IX	TATTACTGTGCTA		CTGAAGCATTT									GGAGTGGG				ACC		CTGGTATTTCGAGCACTGGGGACAAGGTACCATGGTCACCGTCACCTCA	J8		14	11
020719	I	TATTACTGTACTAGAGA		AGTTCGGTT									GGGAGT				TACT		TTCGATTATTGGGGCGCTGGAACAATGGTCACTGTCACATCA	J7b		13	9
060719	I	TATTACTGT		GCAAAAACTGA		GTGG											GATGCT		TTCGATTACTGGGGCGCTGGAACAATGGTCACTGTCACATCA	J7b		17	7
100719	I	TATTACTGTGCTAGAGA		T									GGGGA				TC		ATGCTTACTTTGACATCTGGGGACCAGGGACCACAGTCACAGTTACTTCA	J1		3	8
160719	II	TATTACTGTGCAAG		GTT						GGGGG		AC	TGGGG						TGACTTTGGTTATTGGGGACAAGGAACTATGGCCACCGTCACTTCA	J2		5	8
290719	I	TATTACTGTACTAGA		TGT									GGGGAG				GGA		GCTTACTTTGACATCTGGGGACCAGGGACCACAGTCACAGTTACTTCA	J1		6	8
IgV/Igλ		Y Y C																	W G x G				
050701	II	TATTACTGTTCAAG		GTT						GGGGG		AC	TGGGG						TGACTTTGCTTATTGGGGACAAGGGACTATGGTCACCGTCACTTCA	J2		5	8
060701	I	TATTACTGTACTAGAGA		AGTTCGGT									TGGG				AGTTACT		TTCGATTATTGGGGCGCTGGAACAATGGTCACTGTCACATCA	J7b		15	9
070701	I	TATTACTGTGCTAGA		CATGTCGGT				ACGCTAGCGGGT									CC		CAACTGGTTCGATTACTGGGGACAAGGTACCATGGTCACTGTAACATCA	J7b		11	12
010709	I	TATTACTGTACTAGAGA							TAGCGG				TGGG				AGCCCTTACTATGCT		TTCGATTACTGGGGCGCTGGAACAATGGTCACTGTCACATCA	J7b		14	11
030709	I	TATTACTGTGCTAG		CCAACACCGA						TACGGGG							CCCCTCCGGCCTCAT		ATGACTTTGCTTATTGGGGACAAGGAACTATGGTCACCGTCACTTCA	J2		25	14
040709	II	TATTACTGTGCAAG		GTT						GGGGG		AC	TGGGG						TGACTTTGGTTATTGGGGACAAGGAACTATGGTCACCGTCACTTCA	J2		5	8
050709	I	TATTACTGTGCTAGA		CATGTCGGTACGCC					AGCGGGT								CC		CAACTGGTTCGATTACTGGGGACAAGGTACCATGGTCACTGTAACATCA	J7b		16	12
070709	I	TATTACTGTACTAGAG		G															CTGGTTCGATTACTGGGGACAAGGTACCATGGTCACTGTAACATCA	J7b		1	4
080709	II	TATTACTGTACAAGAG		TCTCA						TACGG									CAACTGGTTCGATTACTGGGGACAAGGTACCATGGTCACTGTAACATCA	J7b		5	8
110709	I	TATTACTGTGCTAGA		CT					CGGG								CTCAATGCT		TTCGATTACTGGGGCGCTGGAACAATGGTCACTGTCACATCA	J7b		11	7
120709													GTGGG				AGGT		GGACAAGGTACCATGGTCACTGTAACATCA	J7			
180731	II	TATTACTGTGCAAG		TACGT					TAGCGG		T	TACGG		CC		CTAC	TACCTGC		CTTTGACTACTGGGGACAAGGGAACCATGGTCACCGTCACTTCA	J7		15	12

Figure 9. Correlation of N/P nucleotide additions with CDR3 length.

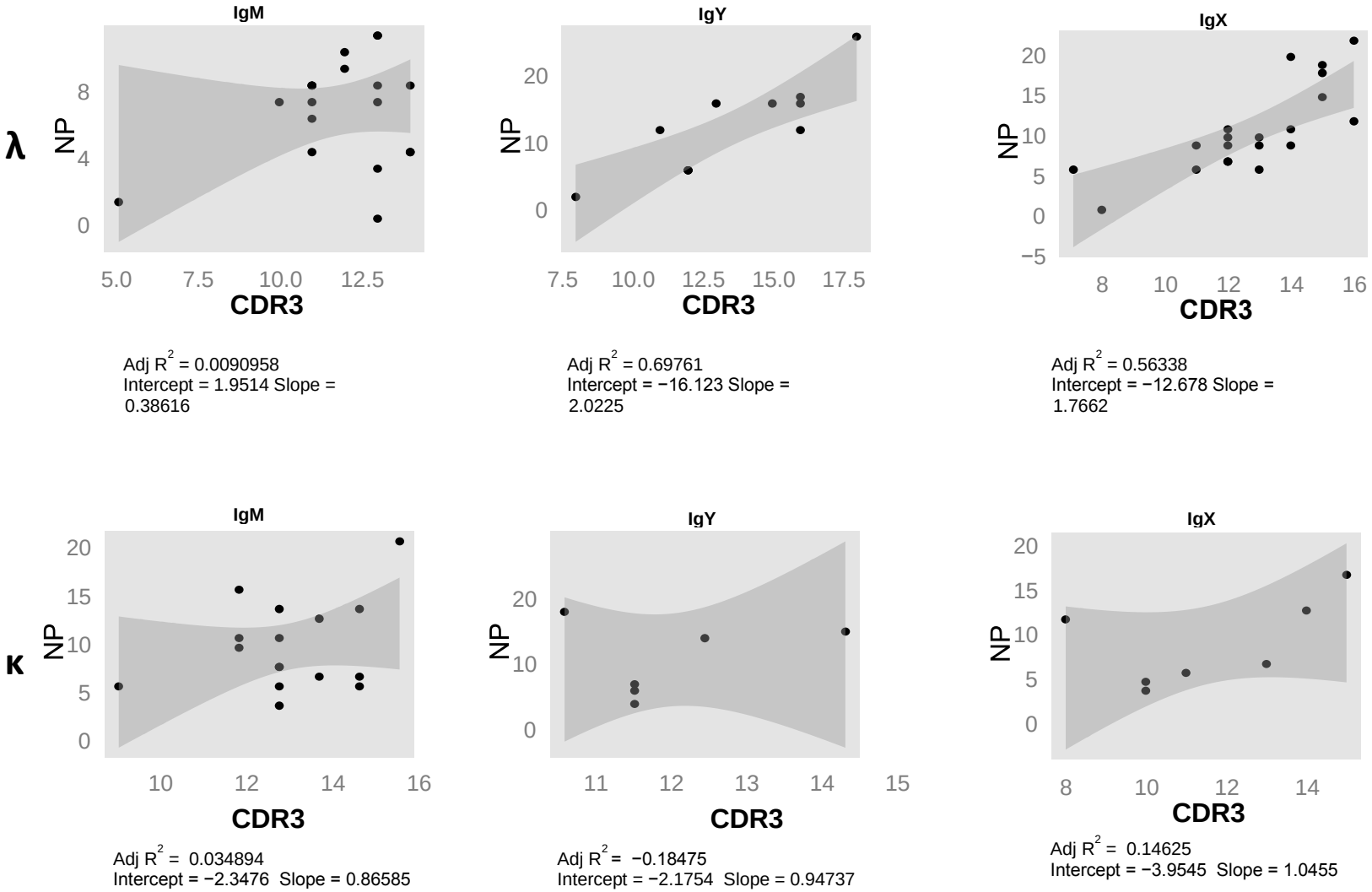
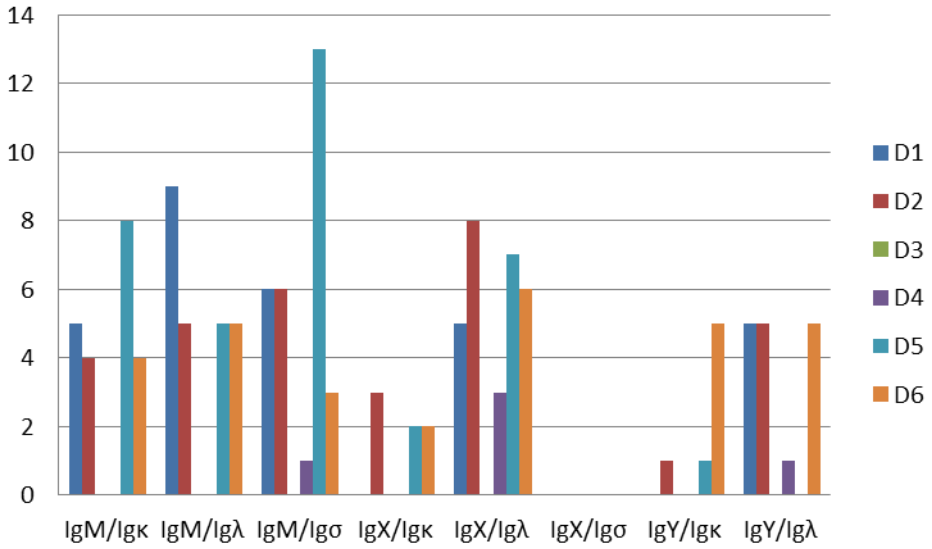
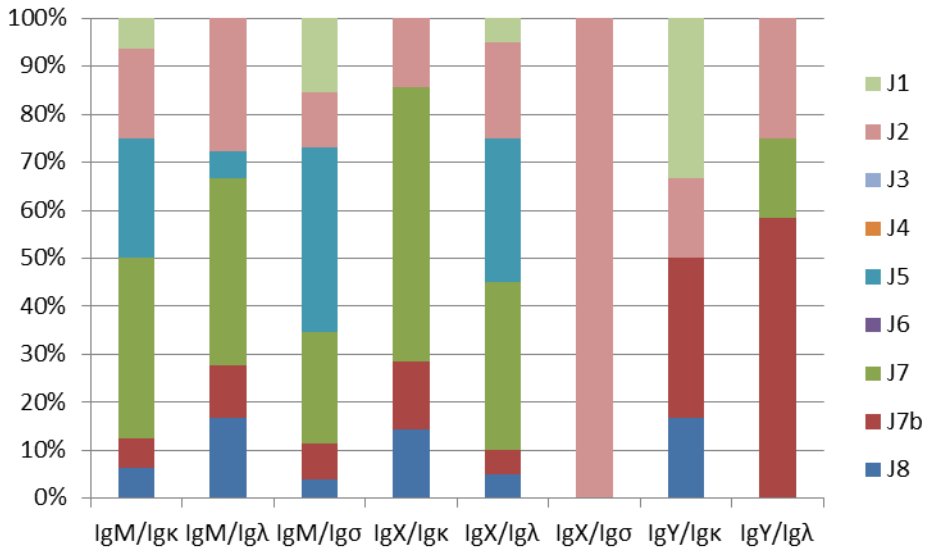


Figure 10. Use of D and J segments in IgH isotype of particular IgH and IgL isotype sort. A. Use of the six germline D segments. Several use two different D's serially in the same CDR3, some use none (or they were exonucleased away). **B.** Use of nine germline J segments (scaled to 100% as each sequence has one J). No IgY/Igσ sequences were cloned.

A.



B.



APPENDIX B

TABLES

Table 1. IgH clones sequenced by IgL isotype sort.

IgH	IgL	Total	Unique	Duplicate CDR3	no V	FS
M	**κ**	22	16	6	-	-
M	**λ**	68	19	49	-	1
M	**σ**	99	26	73	2	2
X	**κ**	15	7	8	-	-
X	**λ**	30	20	10	-	-
X	**σ**	30	2	28	-	-
Y	**κ**	23	6	17	-	-
Y	**λ**	17	12	5	1	-
Y	**σ**	-	-	-	-	-

Table 2. Use of long CDR1 VH families III and VII by IgL isotype.

	Ratio	%
IgLκ	9/29	31.0%
IgLλ	9/49	18.4%
IgLσ	3/26	11.5%

Table 3. Primers used.

Code	Name	5'-3' Sequence	T_m C°
MFC 158	XlB2MF1	AAC ATT AGT CCC CCG GTG G	60.0
MFC 159	XlB2MR1	GGG AGA CCA CAC ATT CCA CT	60.0
MFC 201	IgMC3-F	AAC ACA CAG AGC TGG CTT CA	58.4
MFC 202	IgMC4-R	AGC ATG TCA AGG TGG CAG TT	58.4
MFC 203	IgXC3-F	GTG TTT GTG CTG AGG ACT GG	60.5
MFC 204	IgXC4-R	TAG TTC TTG AGC GGA TGG TG	58.4
MFC 230	IgYC3-F2	CCT GAT CTT CCA TCA CCA	53.8
MFC 231	IgYC4-R4	CCC TCT TCT TCT TCT TCC	53.8
MFC 435	IgLambdaC-F1	T ACA GGT GAC GTG AAA GCC C	60.0
MFC 437	IgLambdaC-R1	AGC GAT GGG TTG TTG GAG AG	60.0
MFC 234	IgKappaC-F1	AGT TCC TCC GAC GTT AAG AC	58.4
MFC 235	IgKappaC-R1	CTC TGT GTC AGT TGT GCT GT	58.4
MFC 236	IgSigmaC-F1	CAG TAA GCC TGG TCA ATG TG	58.4
MFC 237	IgSigmaC-R1	GAA GCC AGG GTC AAG TAA C	57.5

Table 4. Limits of IgH genomic V and J sequences.

V or J segment	Orientation	Last base	Scaffold
VH I	S	839144	29869
VH II	AS	7017527	13576
VH III	AS	6937997	13576
VH IV	AS	6902258	13576
VH VI	AS	6896015	13576
VH VII	AS	6969838	13576
VH VIII	S	1043406	29869
VH IX	S	59820	272406
VH X	S	1297421	29869
J1	AS	6822054	13576
J2	AS	6821835	13576
J5	AS	6820764	13576
J7	AS	6821611	13576
J7b	AS	6820149	13576
J8	AS	6821325	13576

Table 5. Cells and RNA recovered after IgL isotype sorts.

	Starting Cell Number	Final Cell Number	ratio	RNA(ug)
IgLκ	$6.1x10^7$	$5.07x10^6$	.083	.536
IgLλ	$3.84x10^7$	$3.57x10^6$	.092	.110
IgLσ	$6.1x10^7$	$1.15x10^6$	.018	.226