European Journal of Immunology

Supporting Information

for

Eur. J. Immunol. 10.1002/eji.200737263

Four primordial immunoglobulin light chain isotypes, including λ and κ, identified in the most primitive living jawed vertebrates

Michael F. Criscitiello and Martin F. Flajnik


```
CGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTC
                                                >LDR
GGCACGAGTACACTCGCTGAAGGCAGGCAGTTCTCTGACAGAAATGATGCG
 A  R  V  H  S  L  K  A  G  S  S  L  T  E  M  M  R

GATCTTGTGCTTTCTCTGTTCTCTGGCACTGTGGCTGGGCTGTGGACAGTG
 I  L  C  F  L  C  S  L  A  L  W  L  G  C  G  Q  C
    >FR1
TCAGACTCTGACTCAGCCTCCAGACCTGACTGTGATGCCGGGAAACGCCGC
 Q  T  L  T  Q  P  P  D  L  T  V  M  P  G  N  A  A
            >CDR1                               >FR2
CACACTCAGCTGTAACATTGGAACTGTTGATCGCTATTACGTTGCCTGGTA
 T  L  S  C  N  I  G  T  V  D  R  Y  Y  V  A  W  Y
                                                >CDR2
CAAGCAGACCCCAGGATCAGTTCCTCAGTGGATCCTGTACTACTATCACTC
 K  Q  T  P  G  S  V  P  Q  W  I  L  Y  Y  Y  H  S
                                  >FR3
TCTCAGTGCCCCGACCTATGGGTCTGGGTTTAGCAGCGACCGTTTTACATC
 L  S  A  P  T  Y  G  S  G  F  S  S  D  R  F  T  S

GACCGTCAACAGCGGCCACAACATTTACAACTTAATCATCAAAAACGTGAA
 T  V  N  S  G  H  N  I  Y  N  L  I  I  K  N  V  K
                                  >CDR3
GGTCGATGATGCCGCCGTCTATTACTGTGGAAAGTGGGTTTCCTCAGCTAG
 V  D  D  A  A  V  Y  Y  C  G  K  W  V  S  S  A  S
          >FR4                              >CONSTANT
TACTAGGGTGTTTGGGCAAGGGACCAAGCTCTTTGTTGCAGCTCAGCAACT
 T  R  V  F  G  Q  G  T  K  L  F  V  A  A  Q  Q  L

CCCTGACCCTTCAGTAAACCTGCTTGGGCCTTCAGCTGAGGAGATGTCCAG
 P  D  P  S  V  N  L  L  G  P  S  A  E  E  M  S  S

TAAAGGAGCTGGTACATTAGTTTGTCTGGTCAGCAAACTGTCCATGGGCTT
 K  G  A  G  T  L  V  C  L  V  S  K  L  S  M  G  F

TGCTGCAGTCAGCTGGACAGTGGATGGGAGTCTGACGAGCAGTGAGGTCCA
 A  A  V  S  W  T  V  D  G  S  L  T  S  S  E  V  Q

AACCAGCTTGGTGTCCCGCGACCCGGACAACAGCTTCAGTCTCAGCAGCTA
 T  S  L  V  S  R  D  P  D  N  S  F  S  L  S  S  Y

CCTGACTGTCCCTGGCAGGGATTGGAGCAGTGGGAAGGTGTACTCCTGCAC
 L  T  V  P  G  R  D  W  S  S  G  K  V  Y  S  C  T

AGTGCAACAAGGGACAGCTTCAAAAACAACCGCAACGATCTCACAGTCACG
 V  Q  Q  G  T  A  S  K  T  T  A  T  I  S  Q  S  R

CTGTTAAATAAGTTTCATTTGTGTCTGCTGTCTAACCTTAGCCAGAAAATG
 C  *

TCACTTTTCTTCCACCCCTTCACCATTACCCACCCTCTTCCTCTGGAATCC

GCTGCACTTTGG
```

Figure S1. Clone 61.8. Leader, frameworks, CDRs and constant regions are marked.

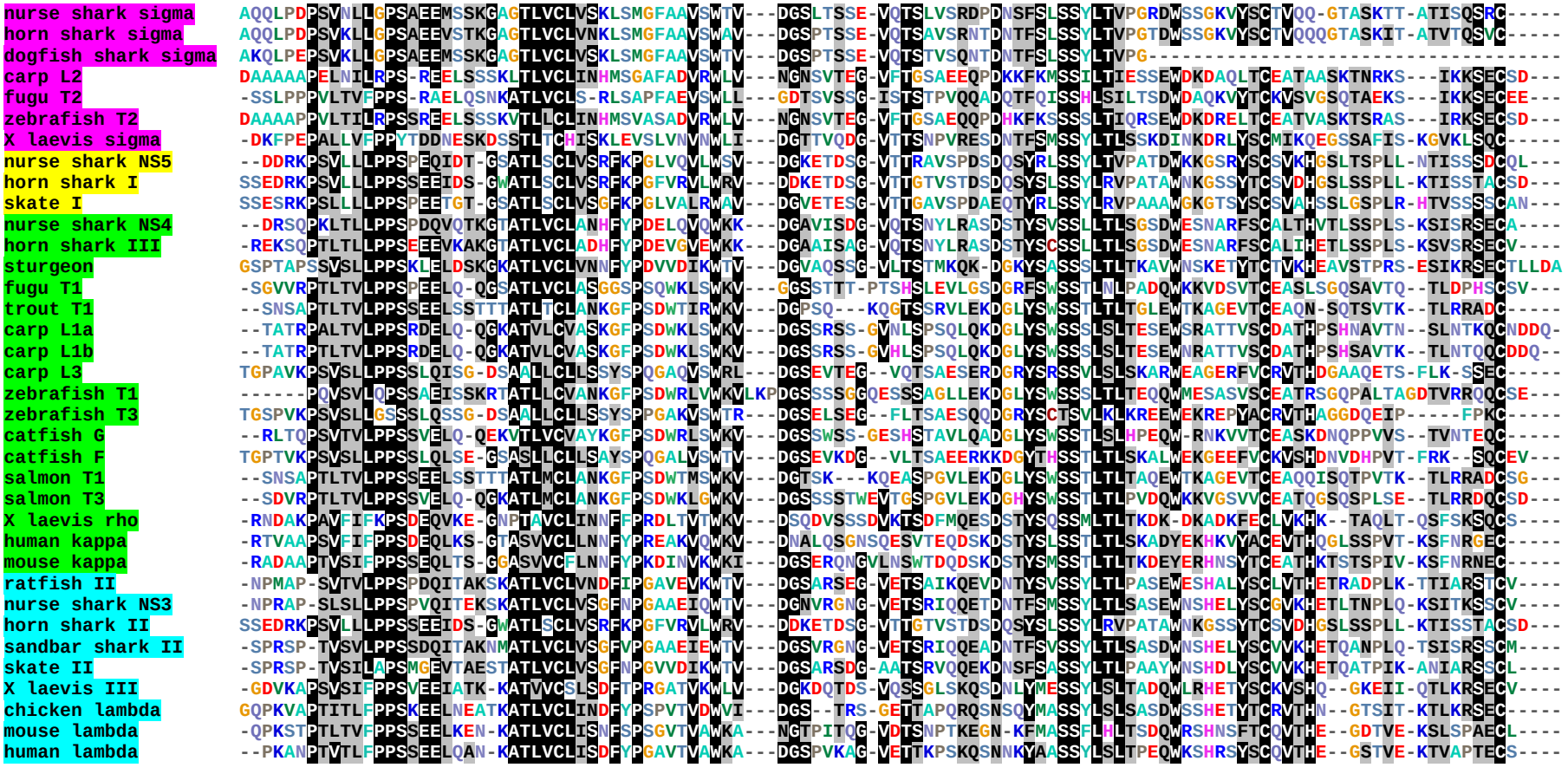
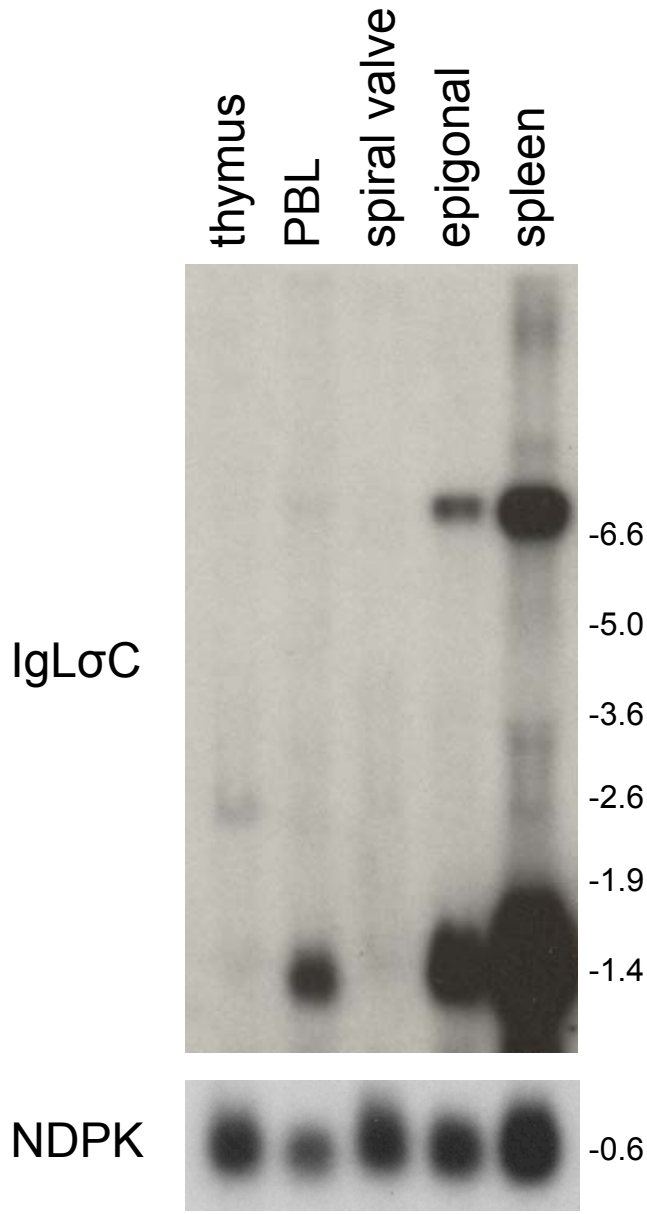


Figure S2. C domain alignment. Hyphens denote gaps. Common names of organisms and original IgL nomenclature are highlighted based on our proposed four isotype scheme; σ = pink, σ' = yellow, κ = green and λ = blue (consistent throughout rest of figures). Accession numbers of sequences used are zebrafish NITR (NM001005577), skate I (L25568), horn shark I (X15315), nurse shark NS5 (xxx), chicken λ (M24403), human λ (AAA59013), mouse λ (AA053422), *X. laevis* TIII (L76575), ratfish II (L25549), skate II (L25566), horn shark II (L25560), nurse shark (xxx), sandbar shark II (M81314), *X. laevis* ρ (XELIGLVAA), horn shark III (L25561), nurse shark NS4 (GSU15144), mouse κ (MUSIGKACN), human κ (S46371), catfish F (U25705), salmon L3 (AF406956), fugu L1 (AB126061), carp L1b (AB073332), zebrafish L3 (AB246193), catfish G (L25533), sturgeon (CAB44624), trout L1 (X65260), salmon L1 (AF273012), zebrafish L1 (AF246185), carp L1a (AB073328), carp L3 (AB073335), trout L2 (AAB41310), carp L2 (AB103558), fugu L2 (DQ471453), zebrafish L2 (AF246183), catfish σ (CK403931), *X. laevis* σ (S78544), *X. tropicalis* σ (scaffold 289), dogfish σ (CX662707), skate σ (CV222129), nurse shark σ (xxx) and horn shark σ (xxx). Solid lines at bottom mark predicted β-strands.

Figure S3 Expression Analysis.
Northern blot of shark RNA probed with IgLσC and nucleotide diphosphate kinase (NDPK) as a loading control. Size markers shown to the right in kilobases.



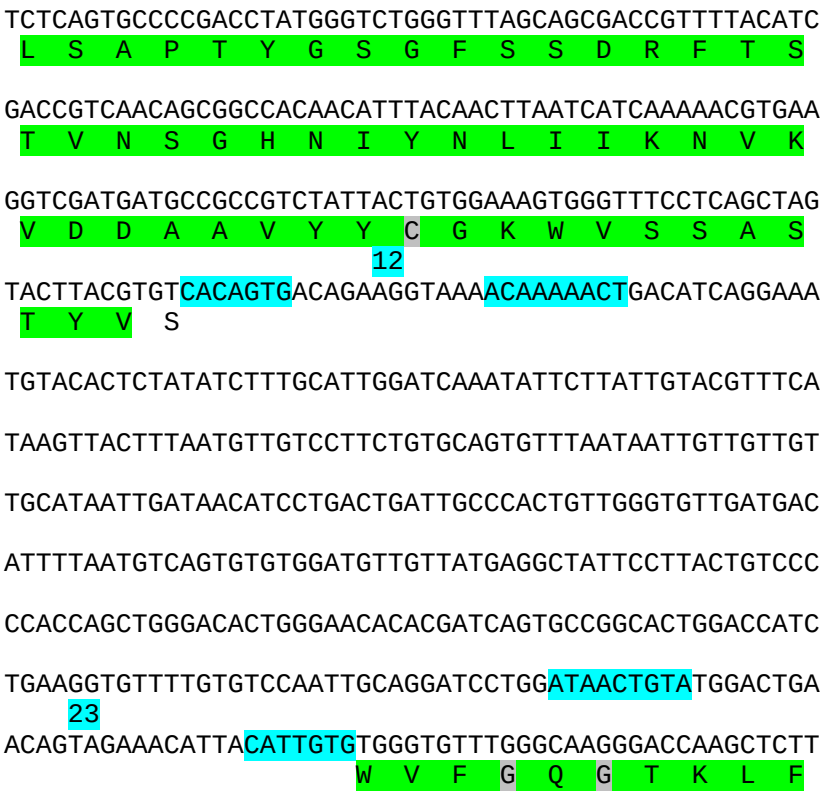
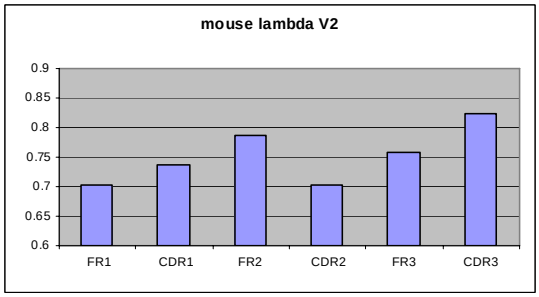
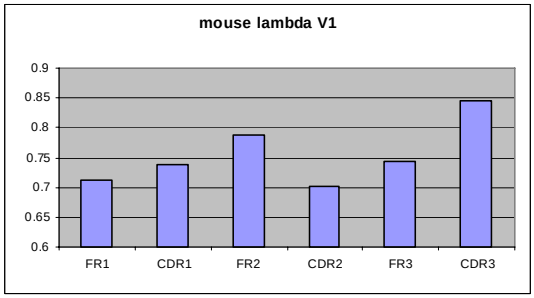
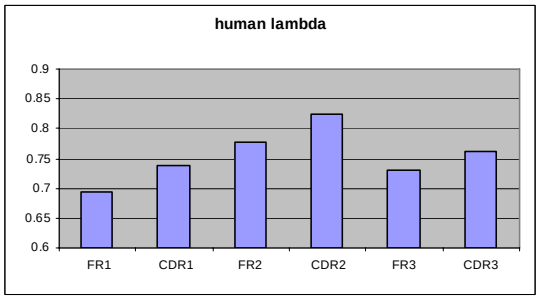
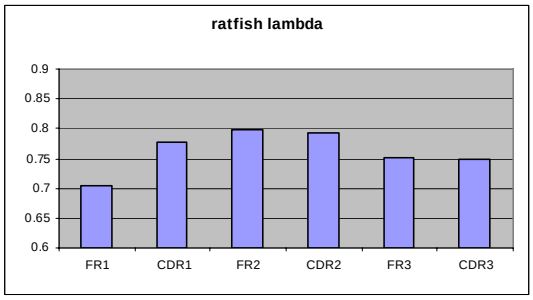


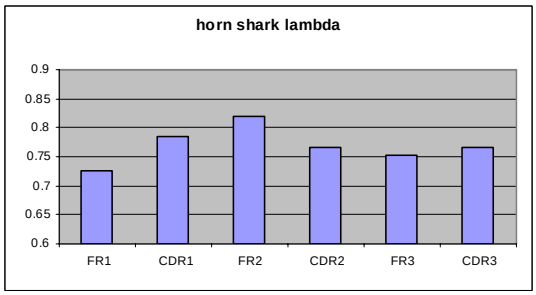
Figure S4. RSS orientation of a nurse shark IgLσ locus. Heptamer and nonamers are highlighted in blue, conserved cysteine of FR3 and GxG of FR4 are highlighted in grey.

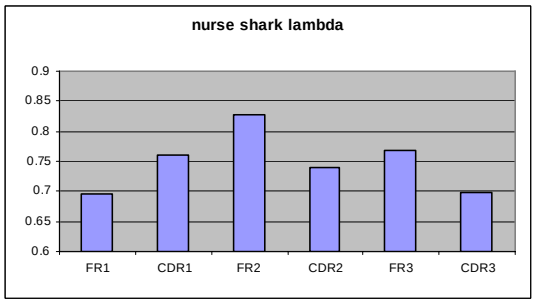

mouse lambda V2
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3

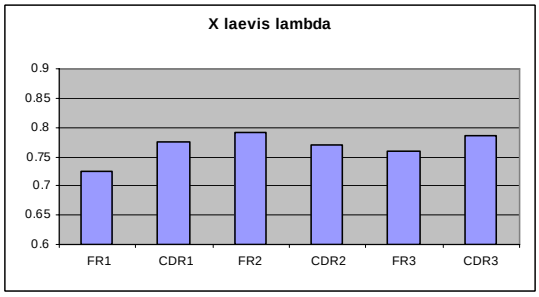

mouse lambda V1
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3

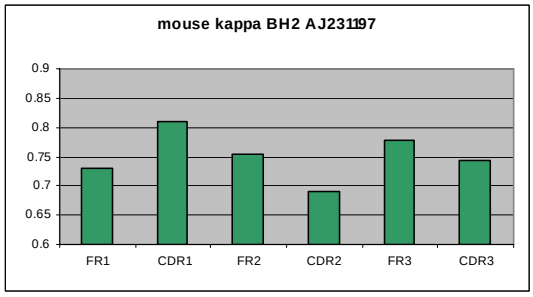

human lambda
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3

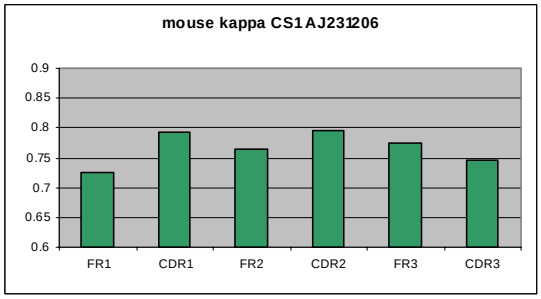

ratfish lambda
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3

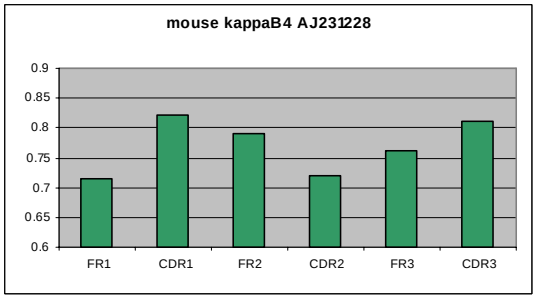

horn shark lambda
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3

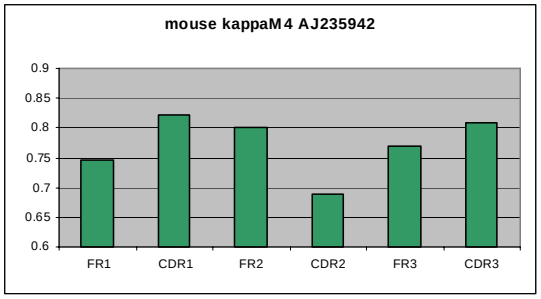

nurse shark lambda
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3

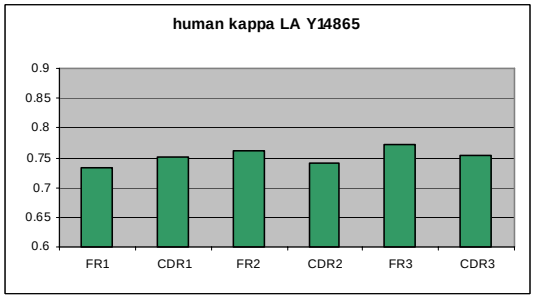

X laevis lambda
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3

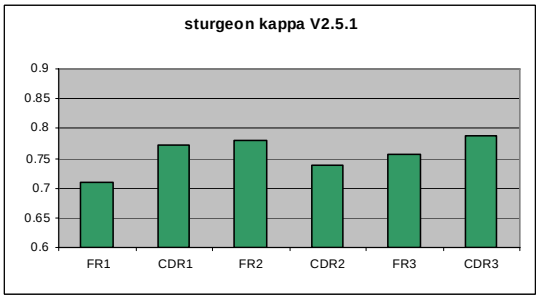

mouse kappa BH2 AJ231197
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3

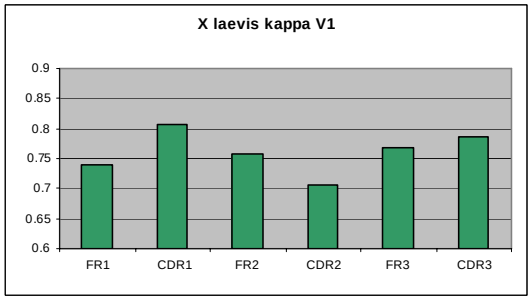

mouse kappa CS1 AJ231206
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3

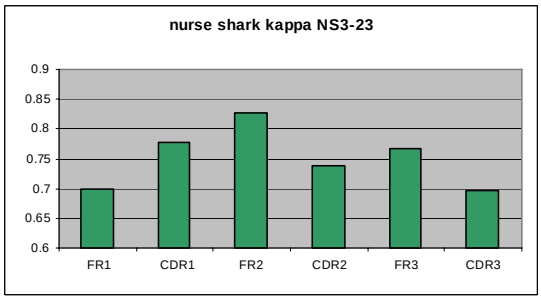

mouse kappaB4 AJ231228
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3

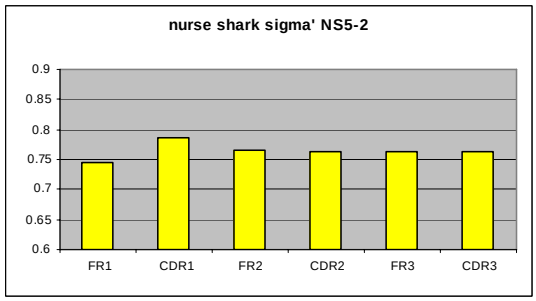

mouse kappaM4 AJ235942
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3

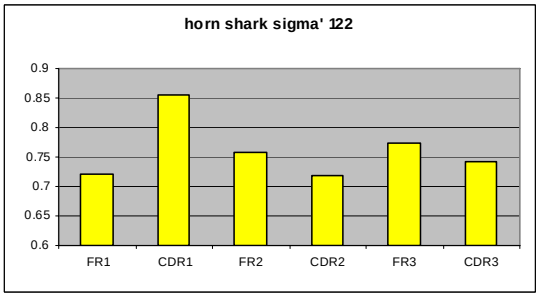

human kappa LA Y14865
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3

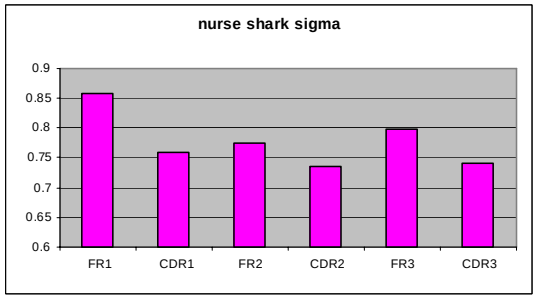

sturgeon kappa V2.5.1
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3

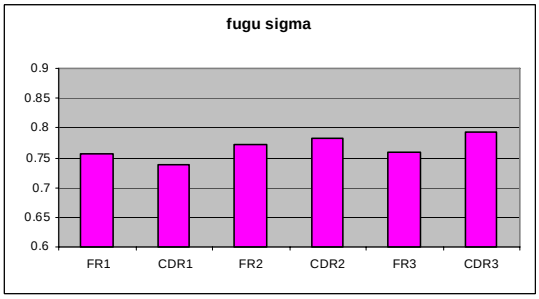

X laevis kappa V1
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3


nurse shark kappa NS3-23
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3


nurse shark sigma' NS5-2
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3


horn shark sigma' 122
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3


nurse shark sigma
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3


fugu sigma
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3

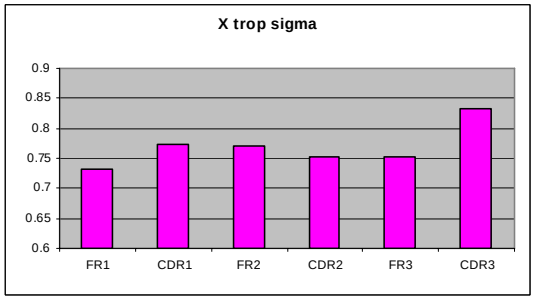

X trop sigma
0.9
0.85
0.8
0.75
0.7
0.65
0.6
FR1
CDR1
FR2
CDR2
FR3
CDR3

Figure S5. Codon volatility analysis of regions of V domains (preceding two pages). Codon volatility scores (as determined by in Sharp *Mol Biol Evol* 2004) were calculated for all codons and averaged for framework and CDR regions based on alignment in Figure 1. For example, the arginine codon “CGA” is least likely to mutate to another amino acid and thus has the lowest volatility score (0.5), whereas any substitution to the tryptophan and methionine codons will result in a change (volatility score of 1.0).

Figure S6. Positive correlation of CDR1L and CDR3H lengths. Ten murine and human known structures of antibody/ligand complexes (reviewed in Padlan 1994) used to compare length of CDRs. Line of best fit drawn with applet at www.mste.uiuc.edu/users/carvell/PlotPoints

	CDR1L	CDR3H
HyHEL-10	11	5
HyHEL-5	10	7
D1.3	11	8
McPC603	17	11
4/4/2020	16	7
AN02	10	6
B13I2	16	10
BV04-01	16	10
17/9	17	11
NC-41	11	10