

Genomics and Genetics of Immunoglobulin in the Rat

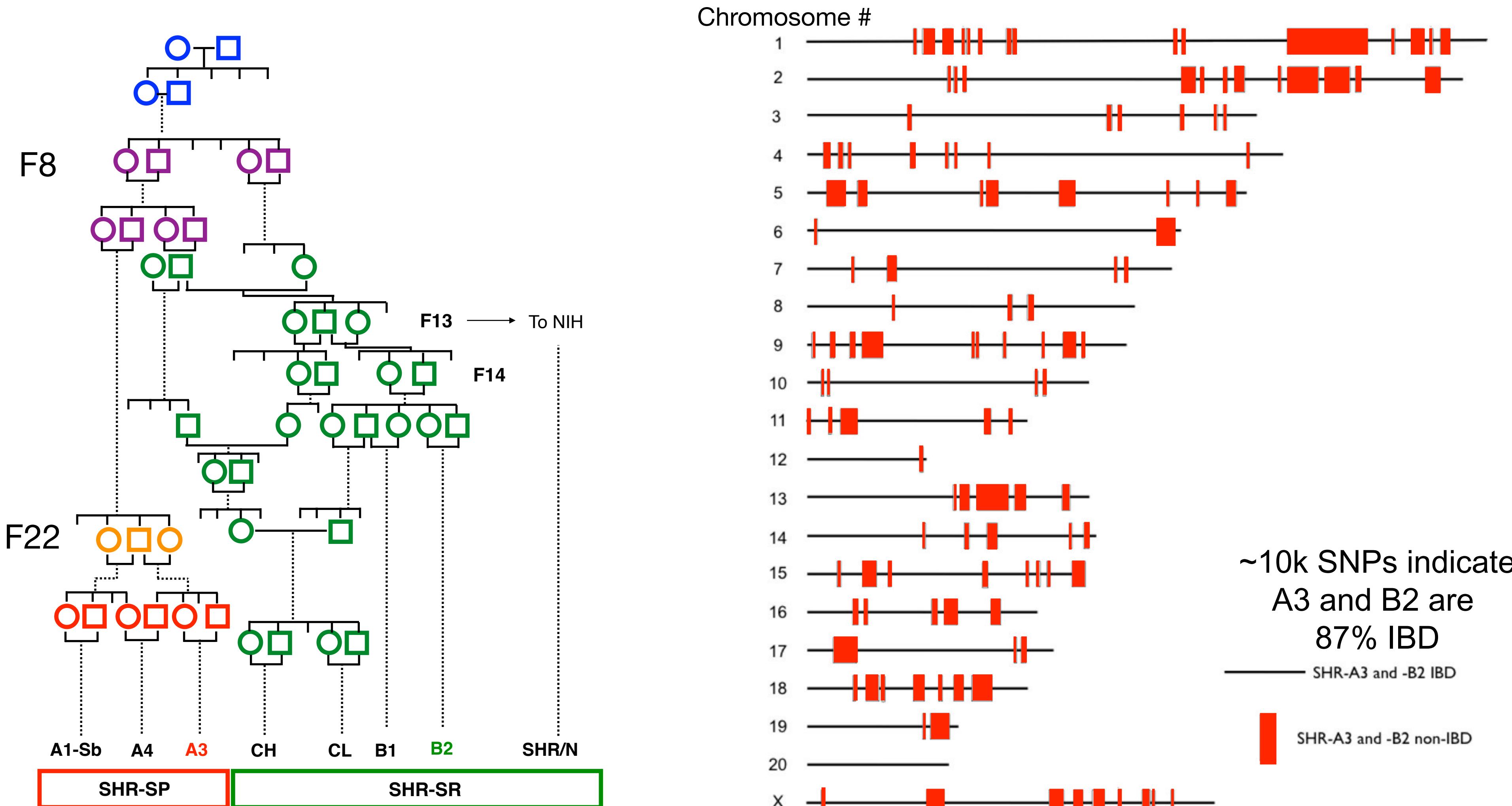
Peter A Doris, Ted Kalbfleisch, Melissa L. Smith

Center for Human Genetics, Brown Foundation Institute of Molecular Medicine,
McGovern Medical School, University of Texas Health Science Center, Houston, TX, USA.

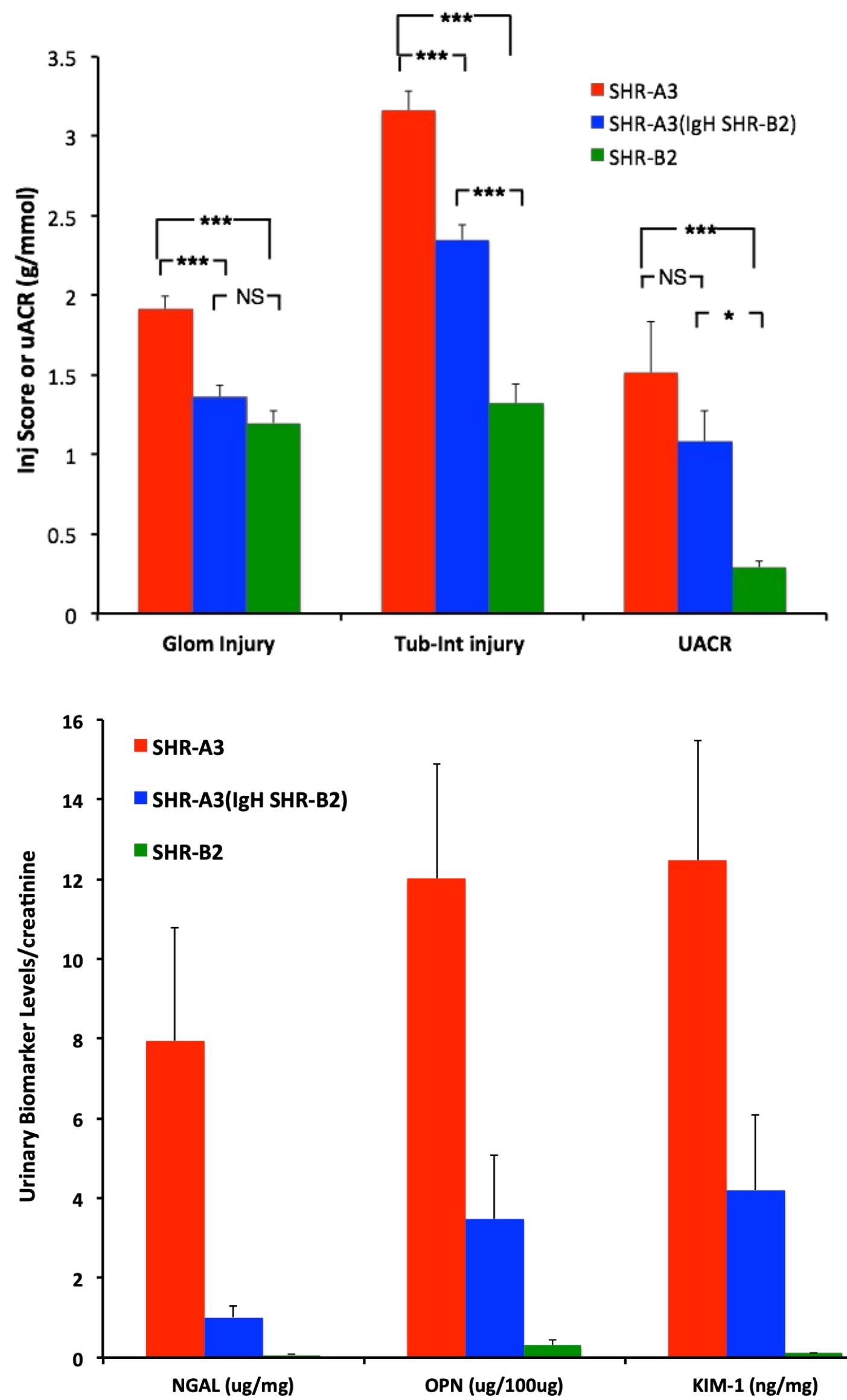
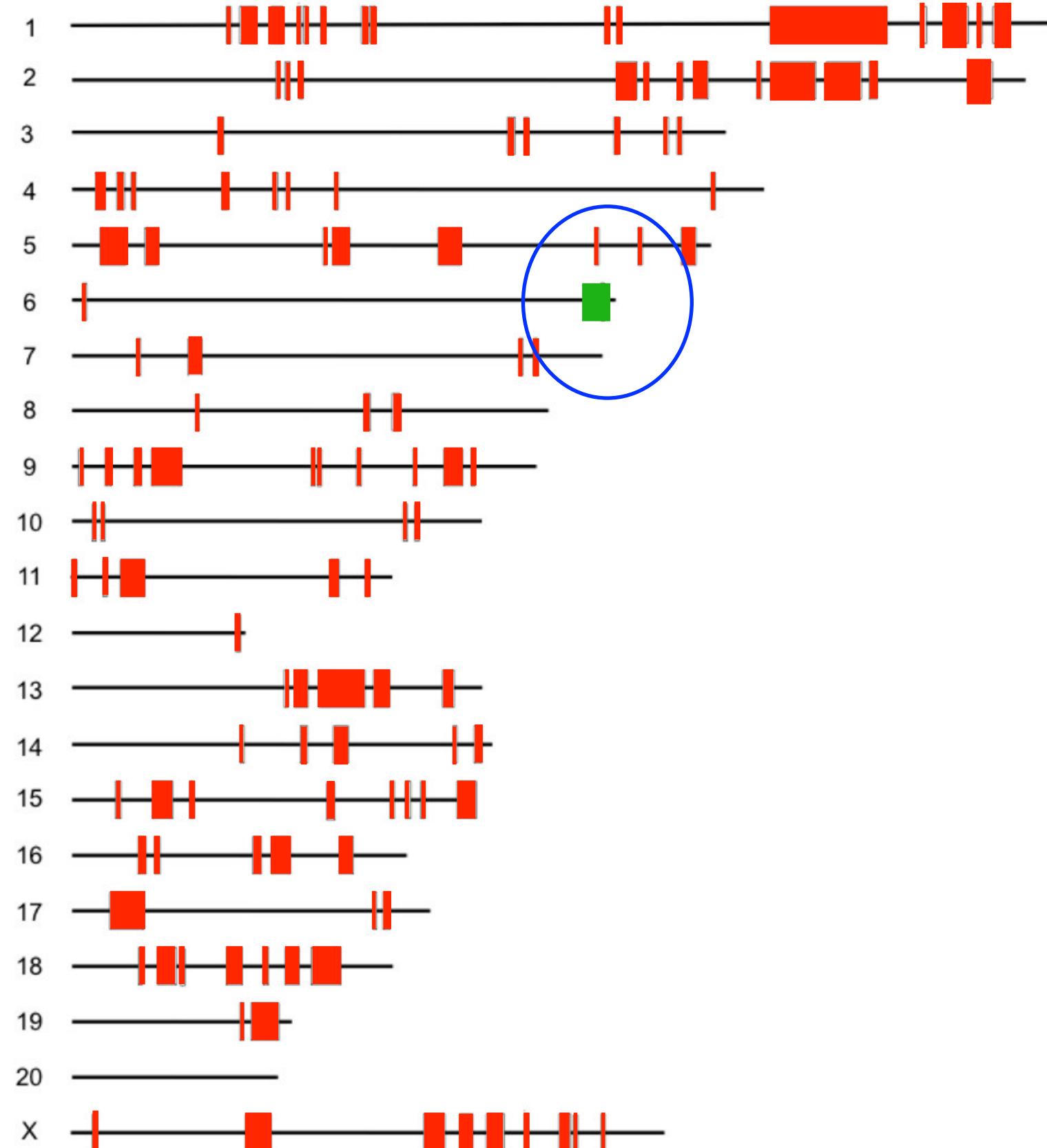
Gluck Equine Genomics Center, University of Kentucky, Lexington, KY, USA
Dept Biochemistry and Molecular Biology, University of Louisville School of Medicine, Louisville, KY

Spontaneously hypertensive rat (SHR)

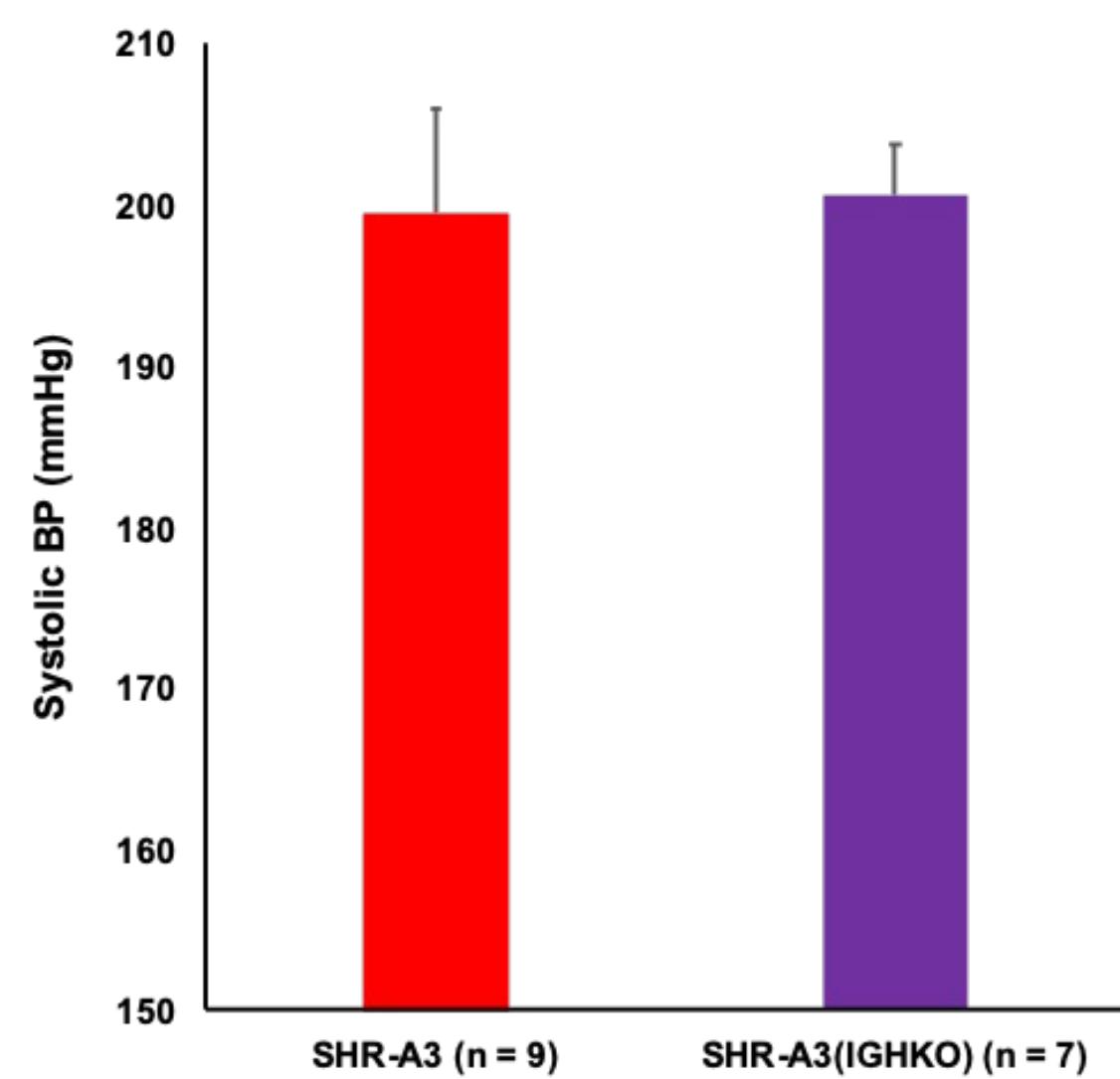
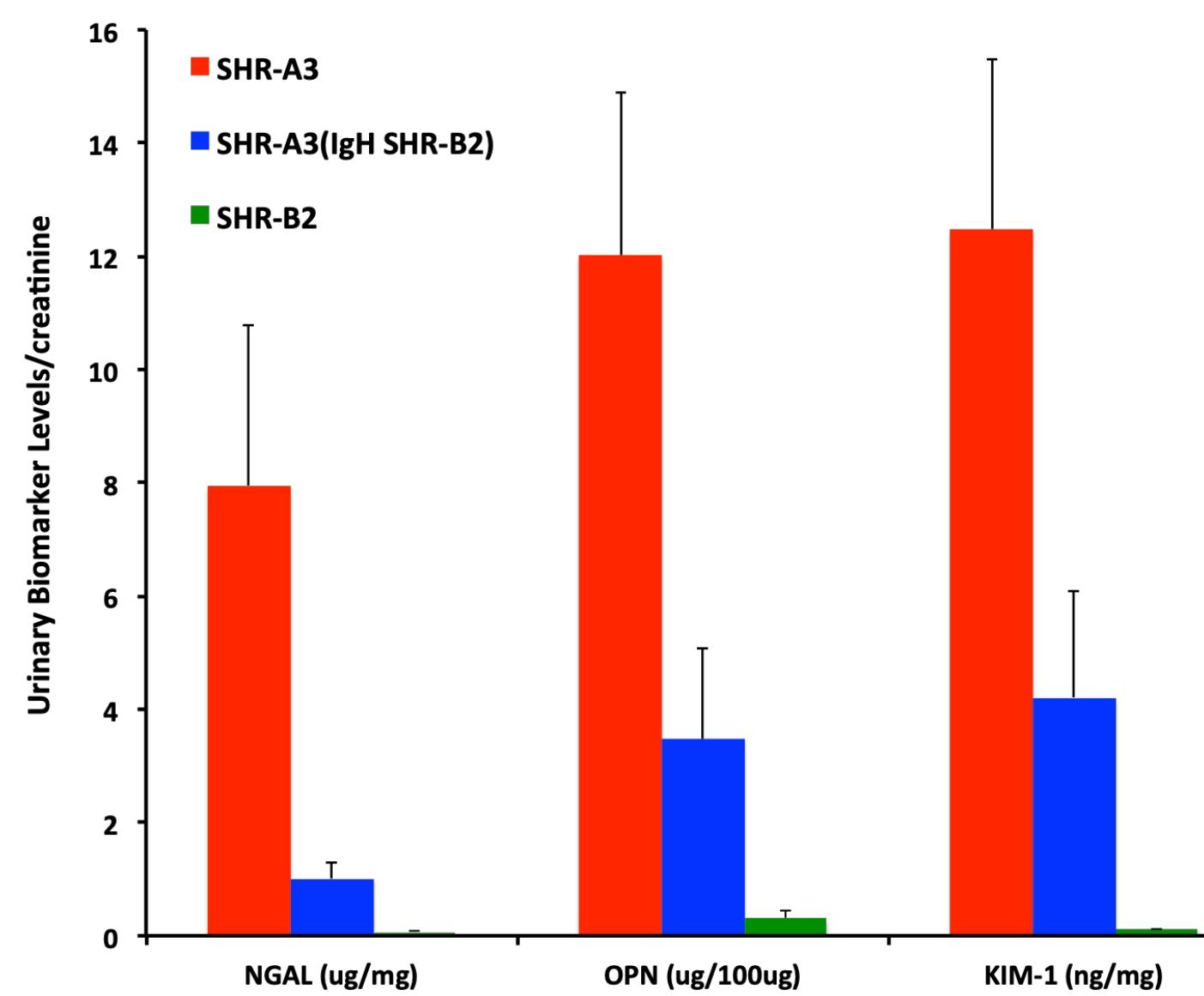
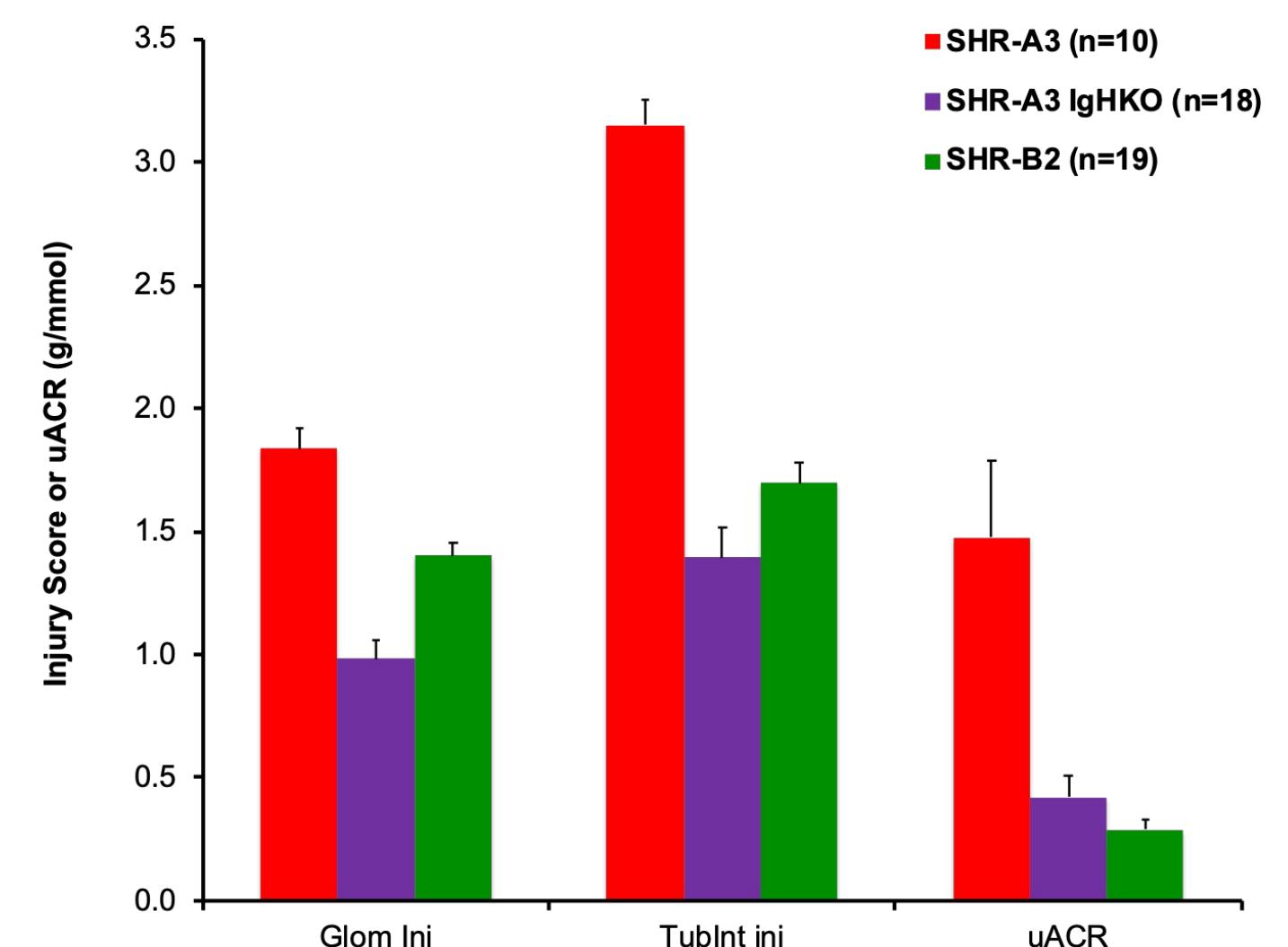
SHR-A3 and SHR-B2 have differential end organ disease susceptibility



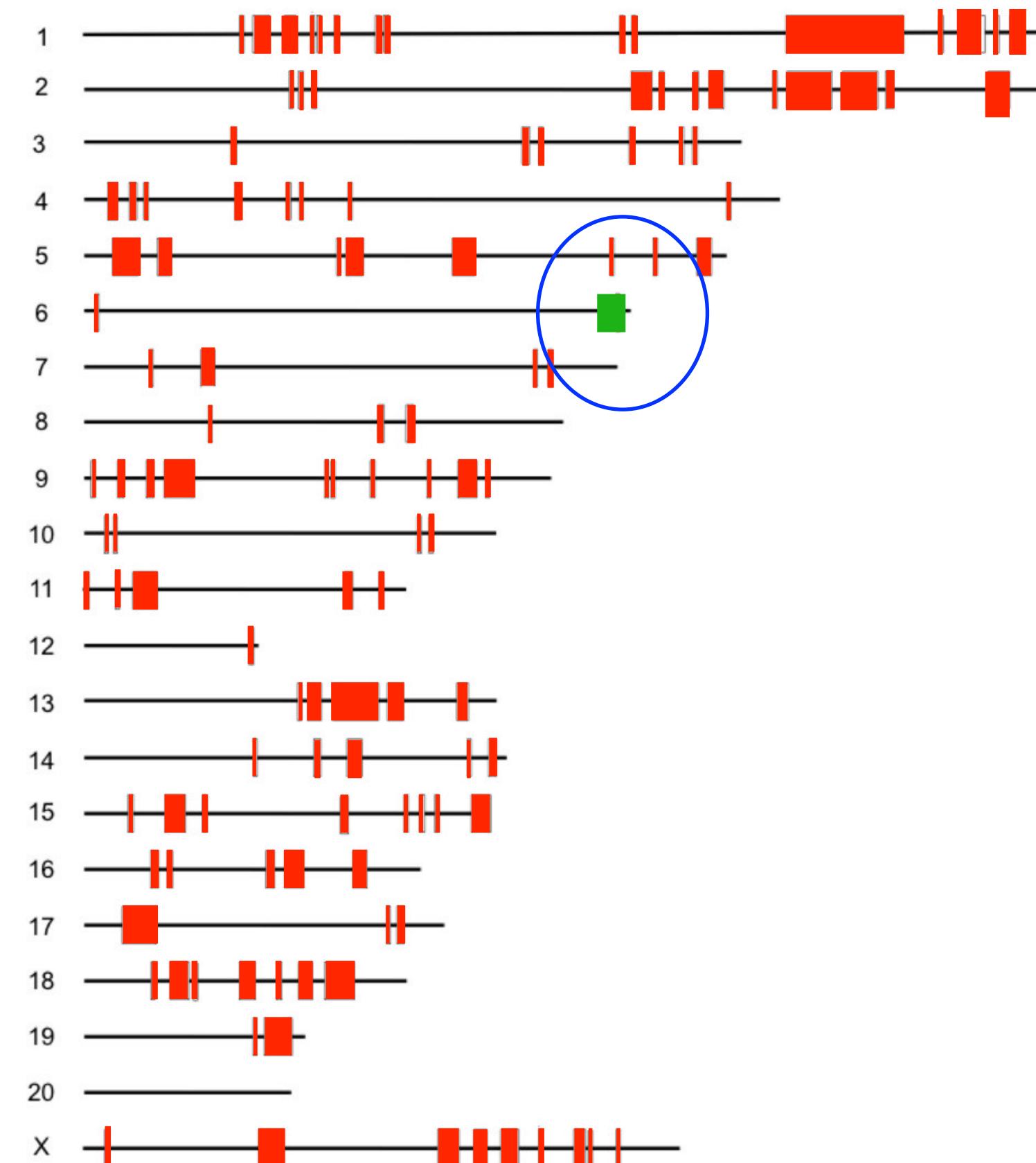
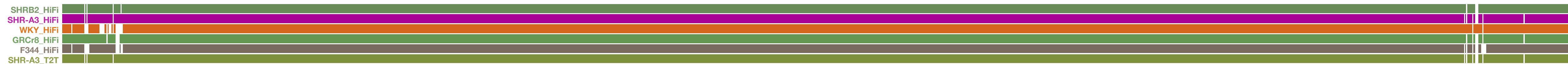
Effect of congenic substitution of IgH locus from SHR-B2 into SHR-A3 on renal injury and biomarkers



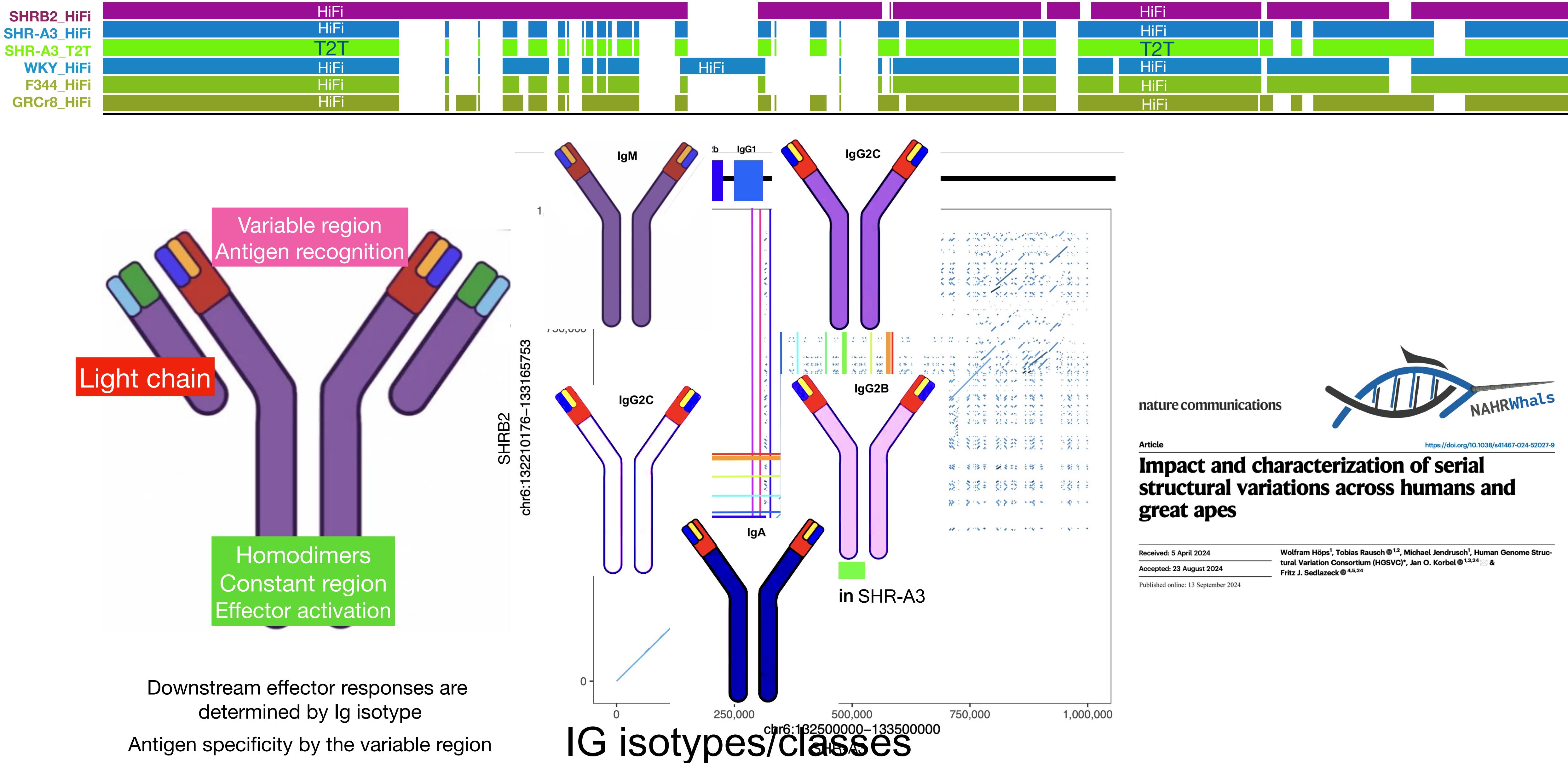
Effect of IgH knockout on renal injury and biomarkers



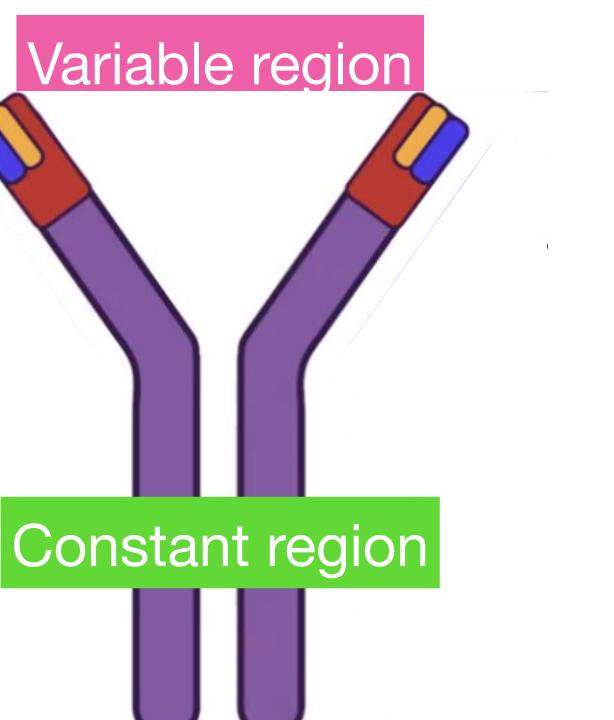
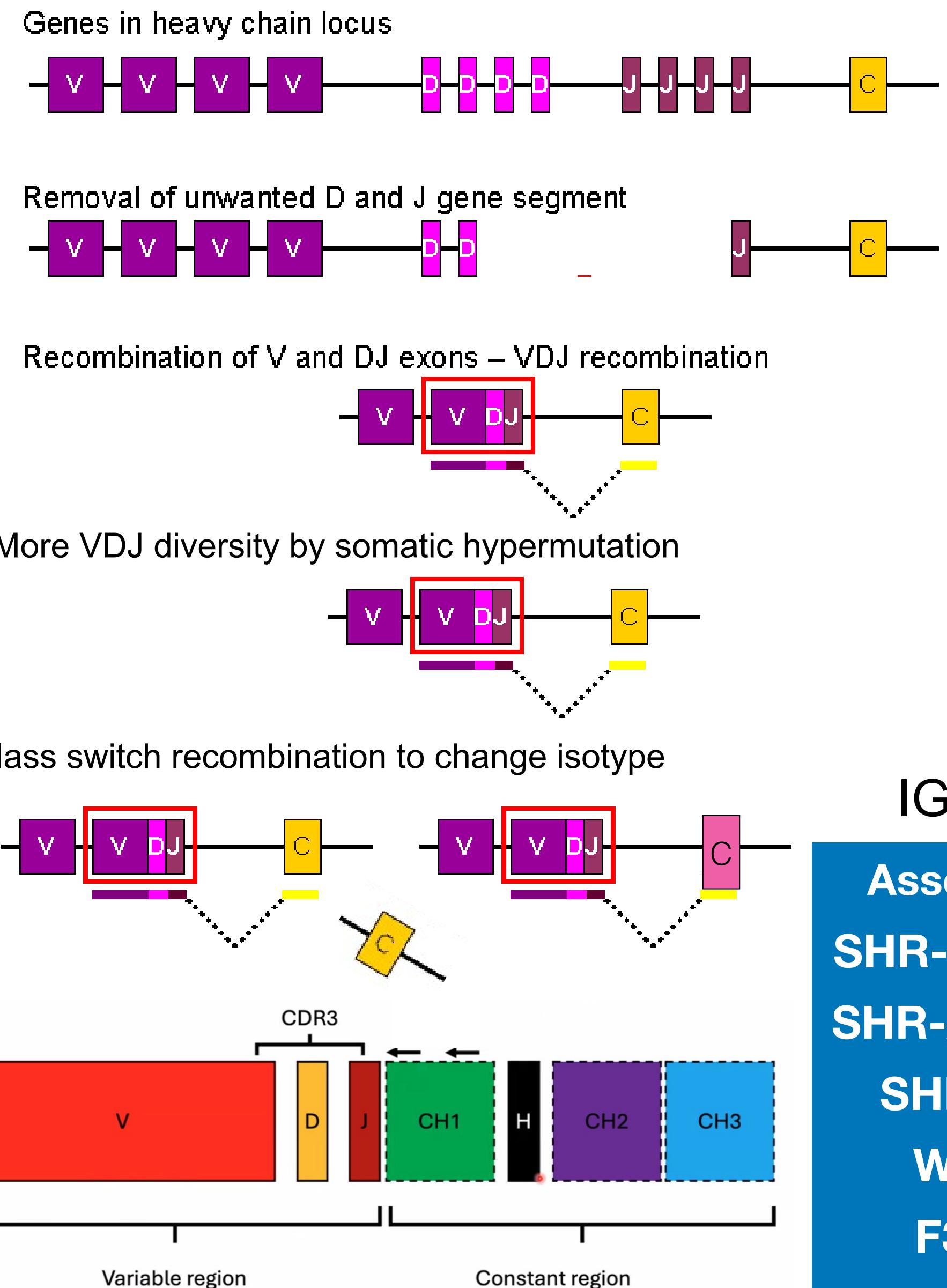
Pangenome visualization of chromosome 6



Lets talk about variation in the Immunoglobulin Heavy Chain



But it gets more complicated: IGH transcripts arise from somatic recombination of gene segments



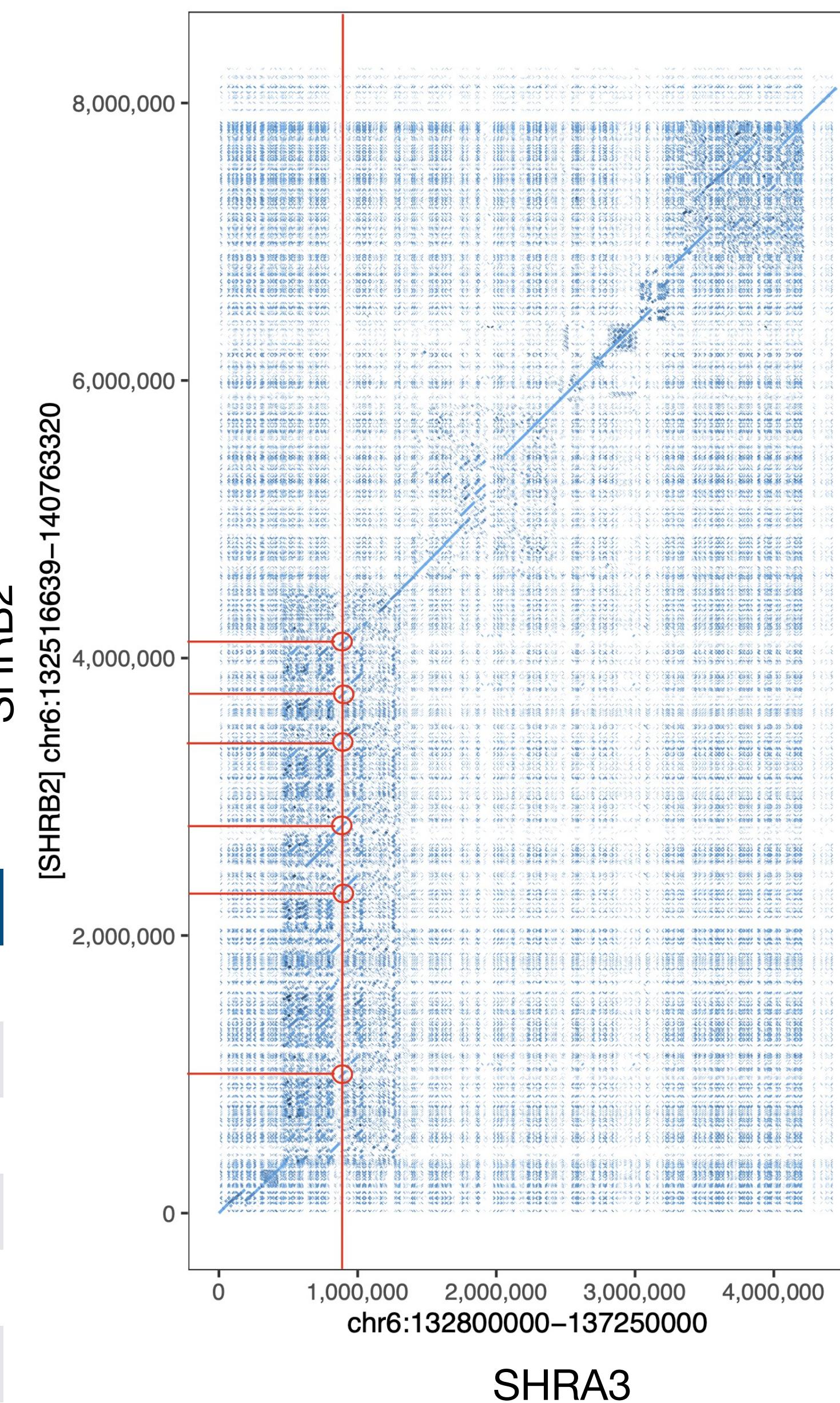
Gene prediction in the immunoglobulin loci

Vikram Sirupurapu,¹ Yana Safonova,^{1,2} and Pavel A. Pevzner¹

¹Computer Science and Engineering Department, University of California San Diego, San Diego, California 92093, USA; ²The Department of Computer Science, Johns Hopkins University, Baltimore, Maryland 21218, USA

IGH VDJ gene prediction from assembly

Assembly	IGHV genes	IGHD genes	IGHJ genes
SHR-A3 T2T	168	33	4
SHR-A3 HiFi	173	33	4
SHR-B2	423	34	4
WKY	313	35	5
F344	245	35	5
GRCr8	226	35	4



So now let's study heritability in a locus that is bonkers complex and variable and exists to respond to antigenic (environmental) inputs

- Haplotype of IGH determines susceptibility or resistance in SHR: one haplotype produces pathogenic antibodies, the other does not.
- Genetic studies require that we maximize G and limit E
- Trans effects arising in SHR-A3 or SHR-B2 genetic background may influence IGH function

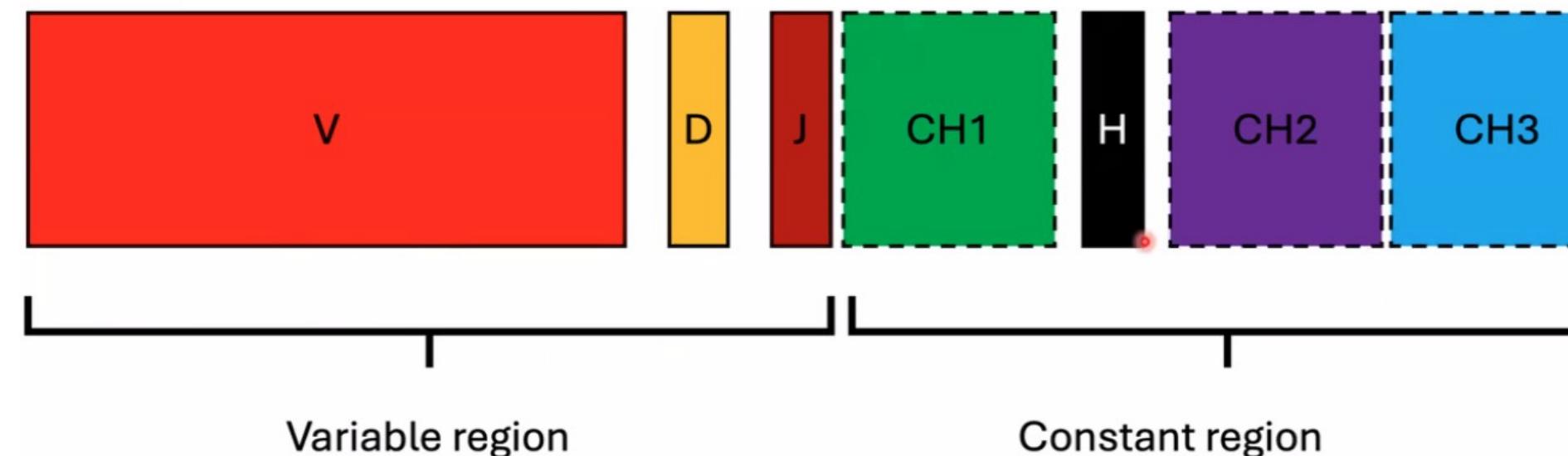
Eliminate environmental and trans-acting effects on IGH using a congenic strain with the SHR-A3 background and both IGH haplotypes

Genome of SHRA3(IGH-A3/IGH-B2)



Kinnex (IsoSeq) sequencing provides single molecule RNA reads

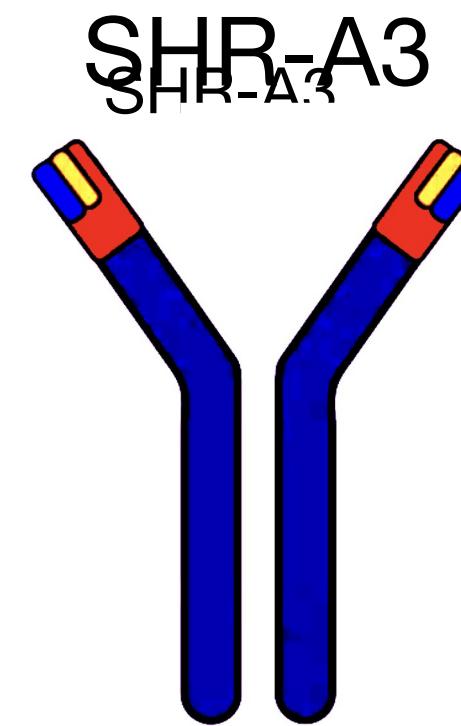
each full length IGH read is ~1.5kb



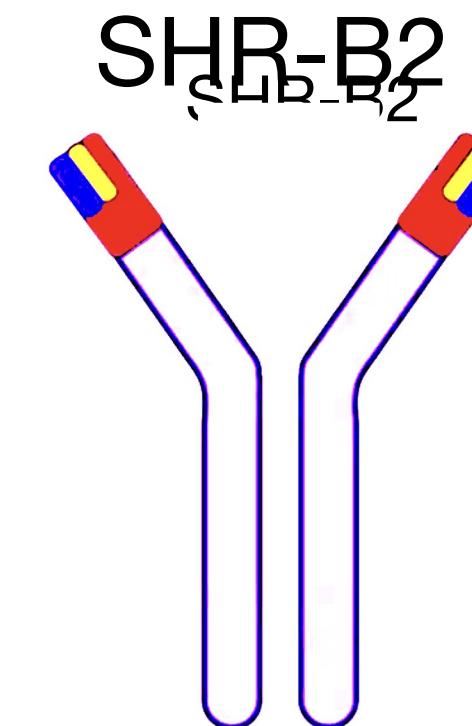
80-100Gb of mRNA sequence per Revio flow cell

Sequencing of mesenteric lymph nodes: yield ~800Mb of IGH reads per flow cell
About 500,000 IGH reads per flow cell

Recover strain specific IGH sequences by blat



sort by haplotype
sort by isotype

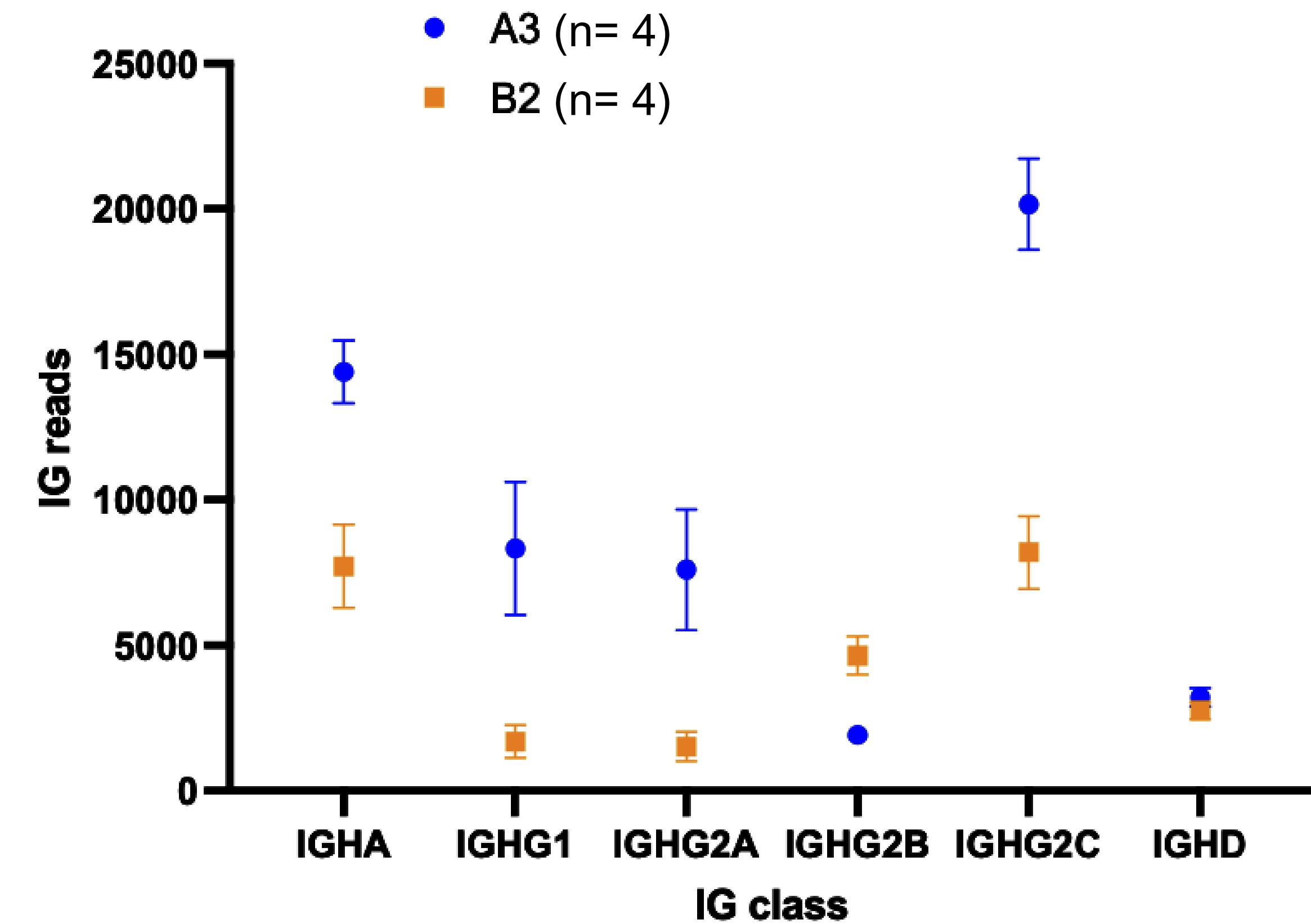


IGH haplotype affects expression of Constant genes

Genome of SHRA3(IGH-A3/IGH-B2)



Both IGH haplotypes in SHRA3(IGH-A3/IGH-B2) have
1) same antigen exposure
2) same trans-acting variant input



IgBLAST a tool to identify VDJ gene utilization in IG reads

National Library of Medicine
National Center for Biotechnology Information

IgBLAST A tool for immunoglobulin (IG) and T cell receptor (TR) V domain sequences

Documentation: Introduction Reference FAQs Stand-alone IgBLAST Ig Germline Genes Other Resources

Analyze immunoglobulin (Ig) sequences Analyze T cell receptor (TR) sequences

Enter Query Sequence Retrieve recent results Save search parameters Reset page

Enter sequence(s) ? Clear

Or, upload local sequence file Choose File No file chosen

Germline gene databases

Organism for query sequence: Rat

Germline V gene Database: IMGT rat V genes (F+ORF+in-frame P)

Germline D gene Database: IMGT rat D genes (F+ORF)

Germline J gene Database: IMGT rat J genes (F+ORF)

Search Parameters

Program: blastn

V gene mismatch penalty: -1

Min D gene nucleotide matches: 5

D gene mismatch penalty: -2

Alignment extension: Extend alignment at 5' end, Extend alignment at 3' end, Allow to overlap

V(D)J genes overlap

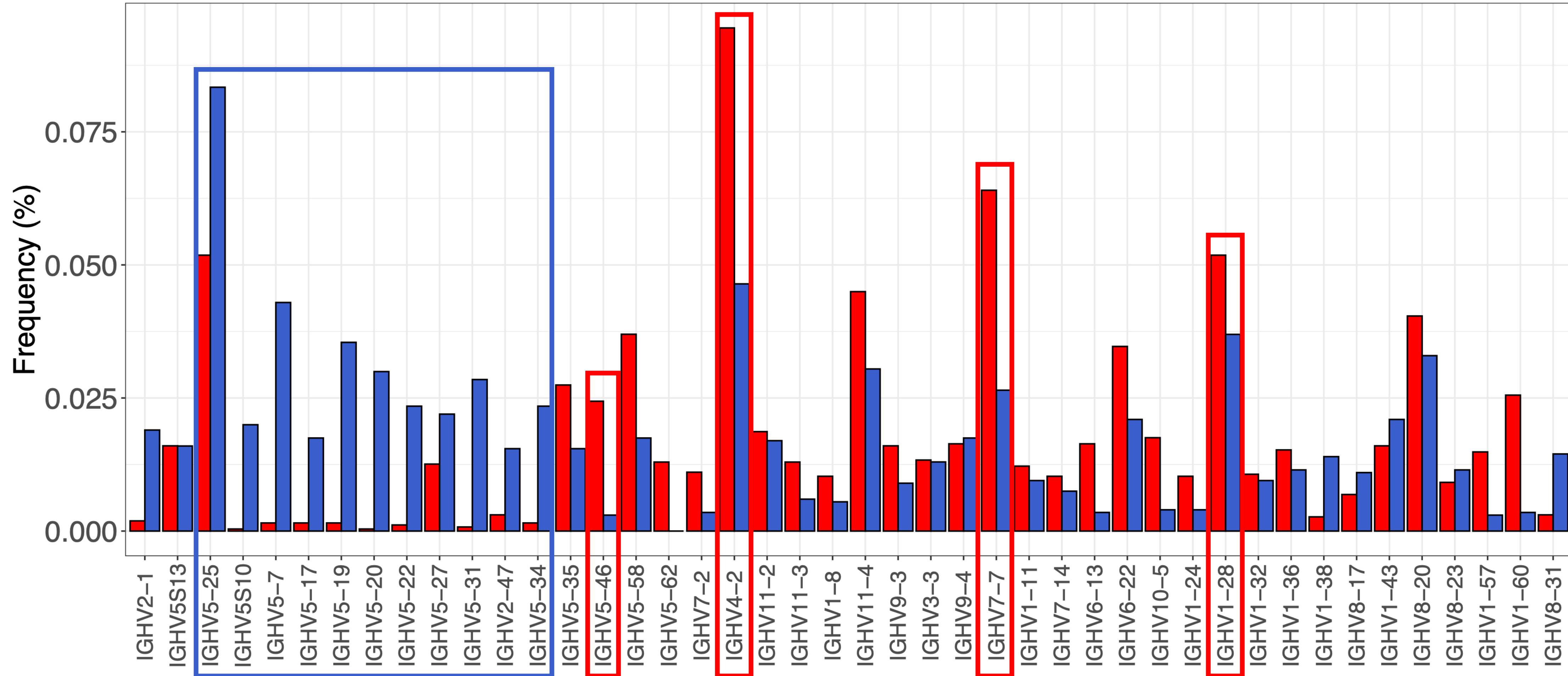
Min required V gene length: 9

Min required J gene length: 0

J gene mismatch penalty: -2

The screenshot shows the IgBLAST search interface. At the top, the NIH logo and 'National Library of Medicine National Center for Biotechnology Information' are displayed, along with a user profile icon 'pdoris'. Below the header, the title 'IgBLAST' and a subtitle 'A tool for immunoglobulin (IG) and T cell receptor (TR) V domain sequences' are shown. A navigation bar with links to Documentation, Introduction, Reference, FAQs, Stand-alone IgBLAST, Ig Germline Genes, and Other Resources follows. The main search area has two tabs: 'Analyze immunoglobulin (Ig) sequences' (selected) and 'Analyze T cell receptor (TR) sequences'. Below these tabs are sections for 'Enter Query Sequence' (with a text input field and 'Clear' button), 'Retrieve recent results' (button), 'Save search parameters' (button), and 'Reset page' (button). There is also a section for uploading local sequence files with a 'Choose File' button and a message 'No file chosen'. The 'Germline gene databases' section allows setting the organism to 'Rat' and selecting databases for V, D, and J genes. The 'Search Parameters' section contains various dropdown menus and checkboxes for blast parameters like mismatch penalties, alignment extensions, and gene overlaps.

Haplotype specific utilization of Variable (IGHV) genes is observed (all IGH-C classes combined)

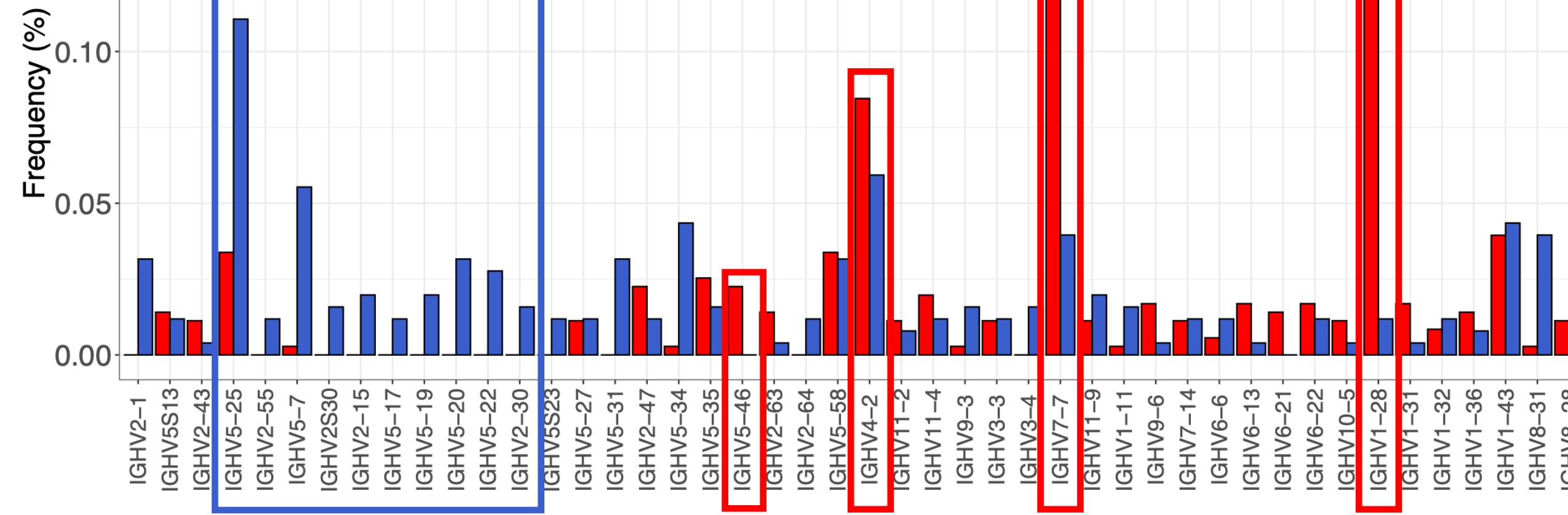


Expression from
SHR-A3 haplotype

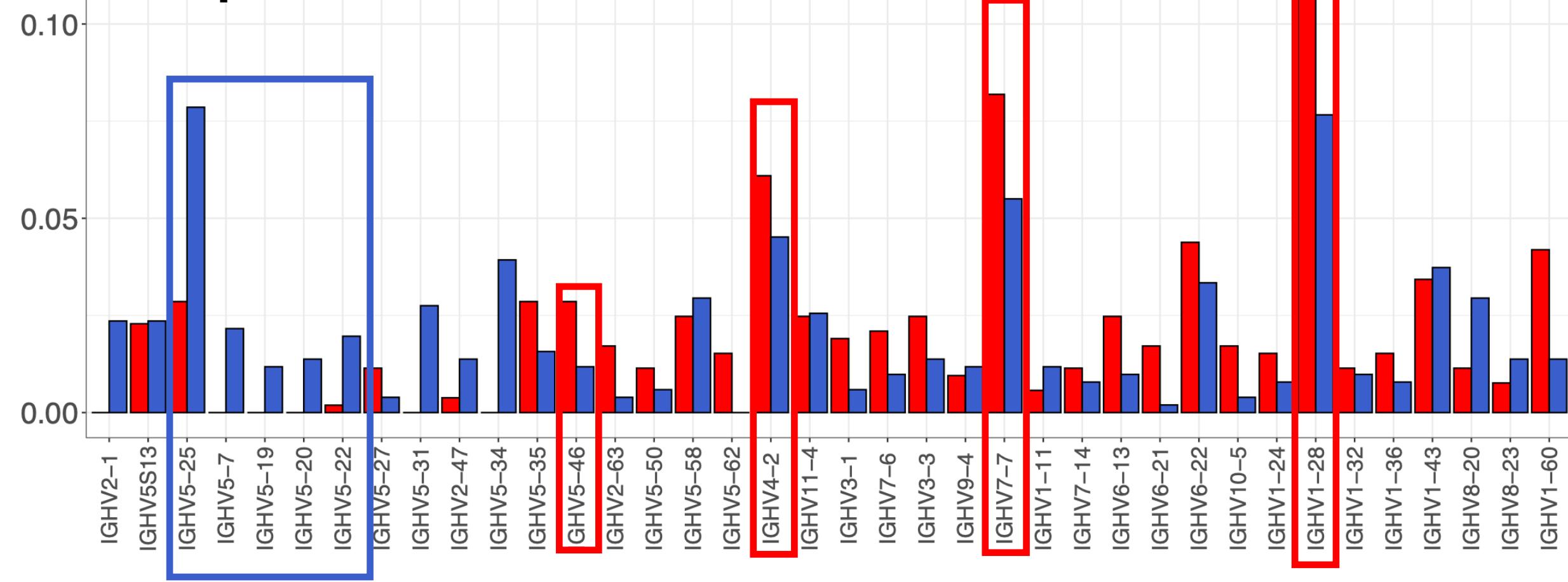
Expression from
SHR-B2 haplotype

Haplotype specific utilization of Variable (IGHV) genes is observed within single classes (IGHA) across multiple animal replicates (18 weeks of age)

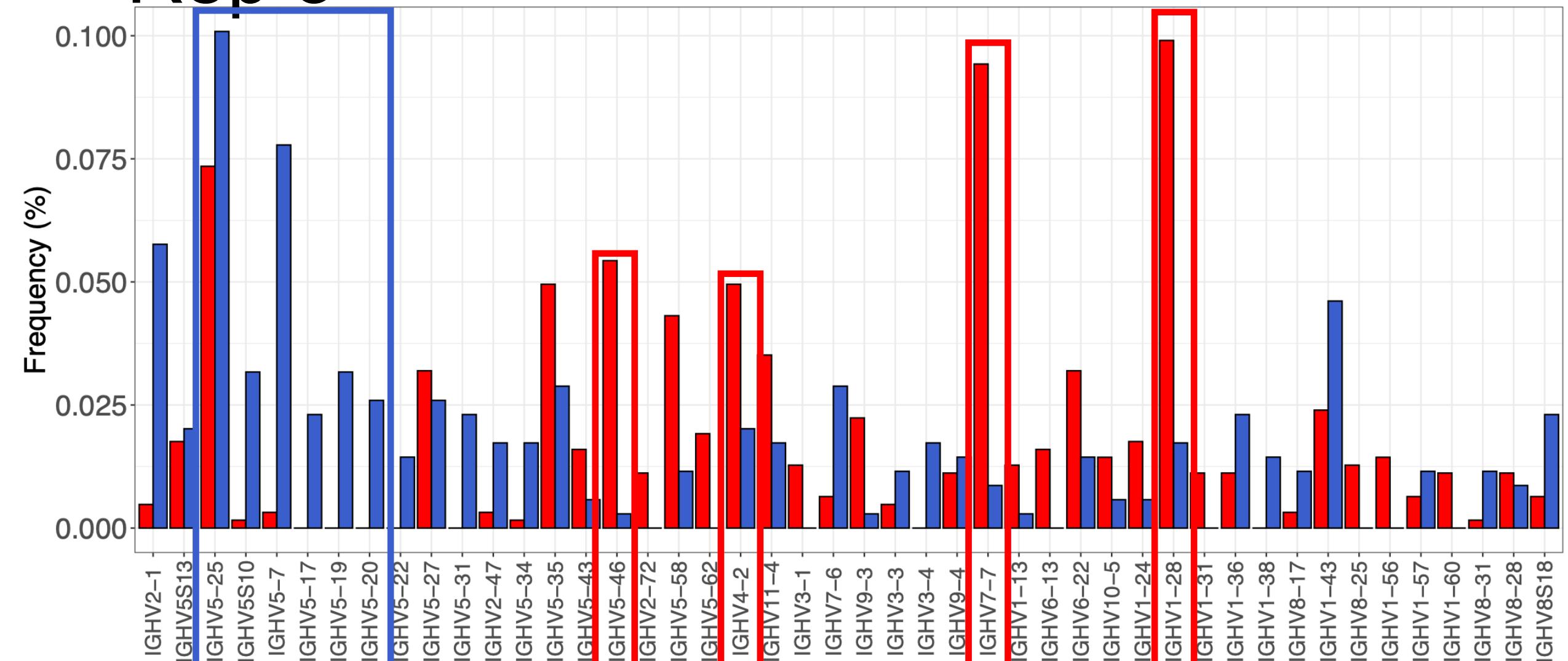
Rep 1



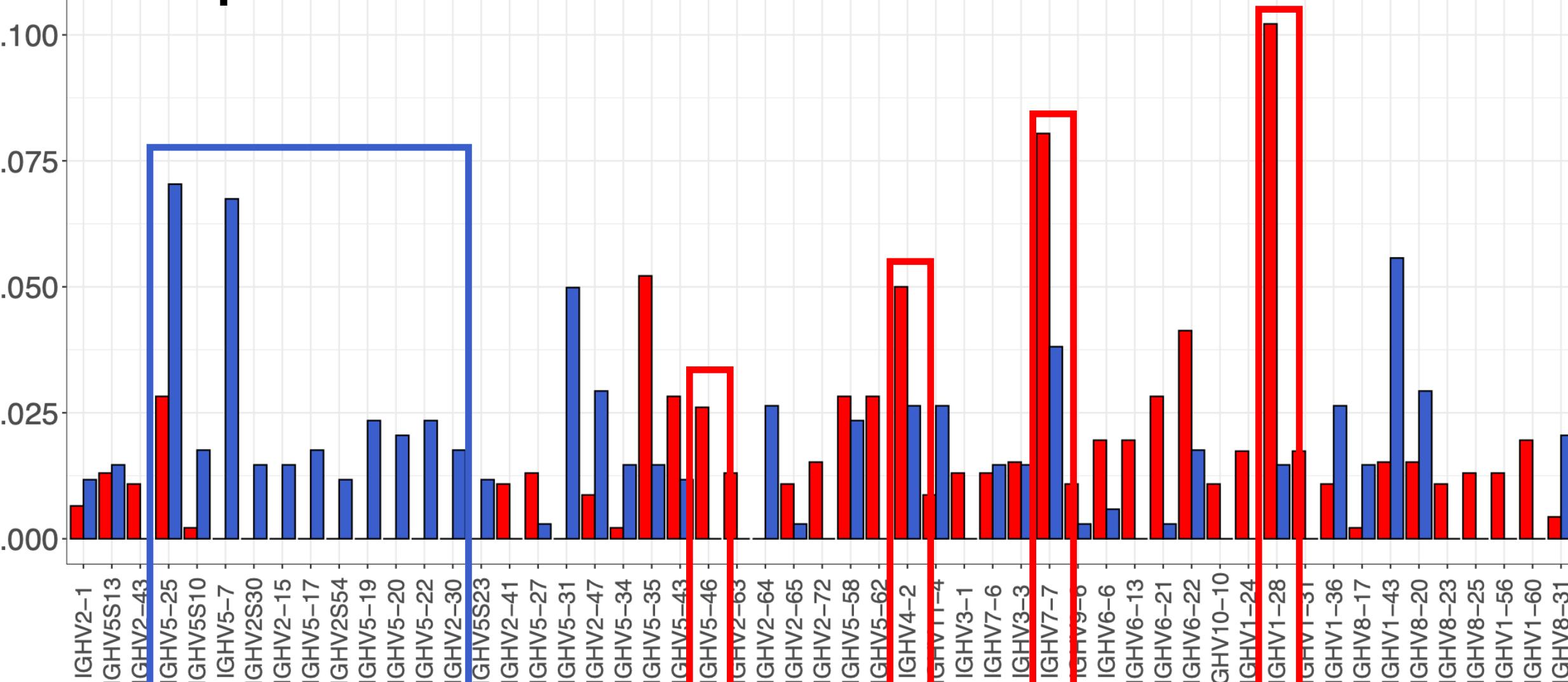
Rep 2



Rep 3



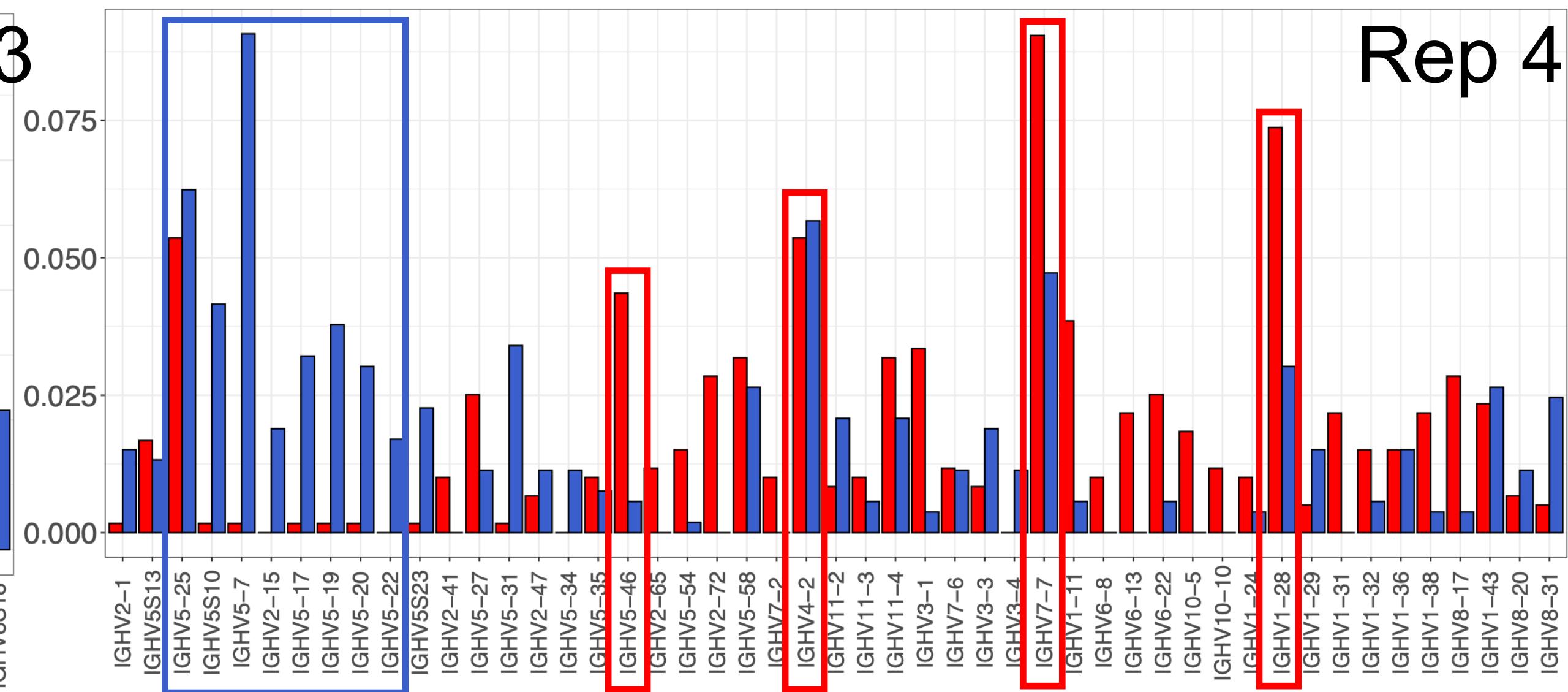
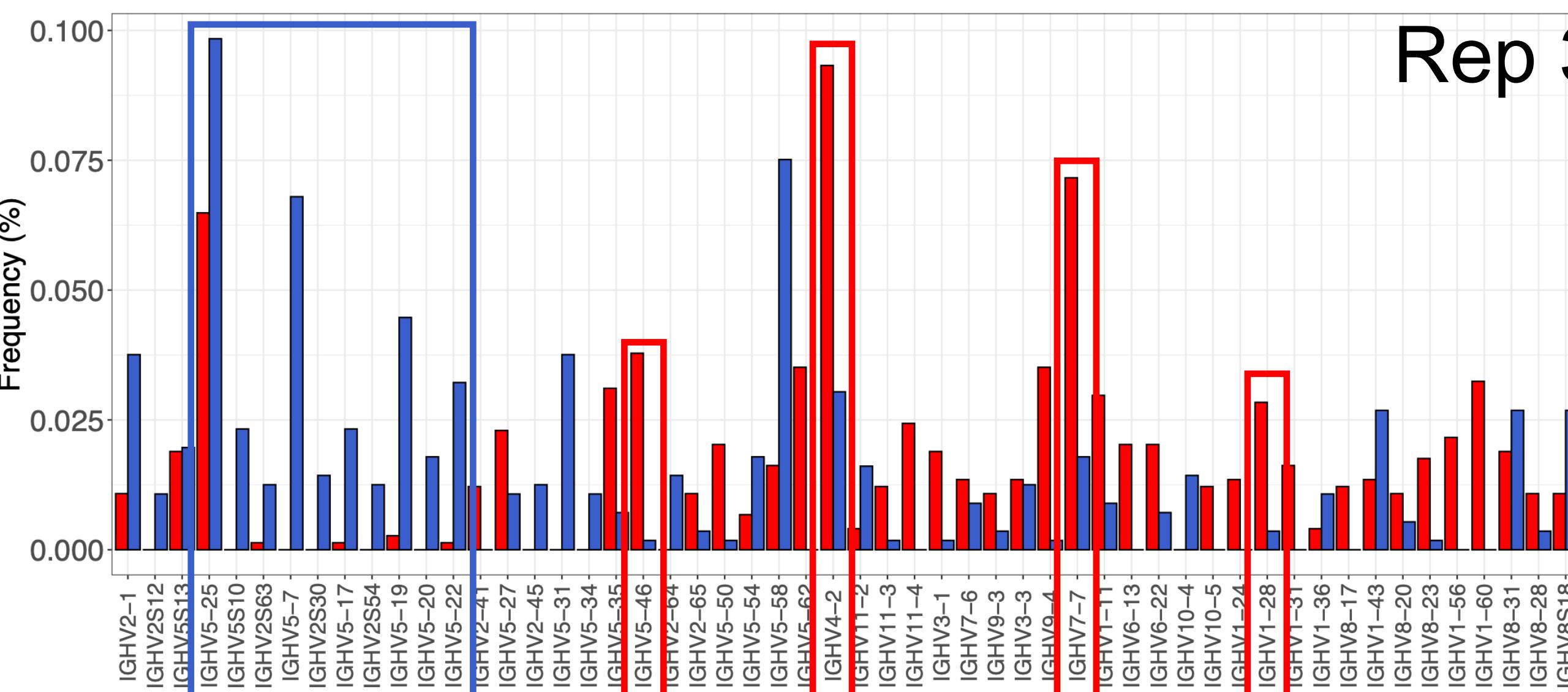
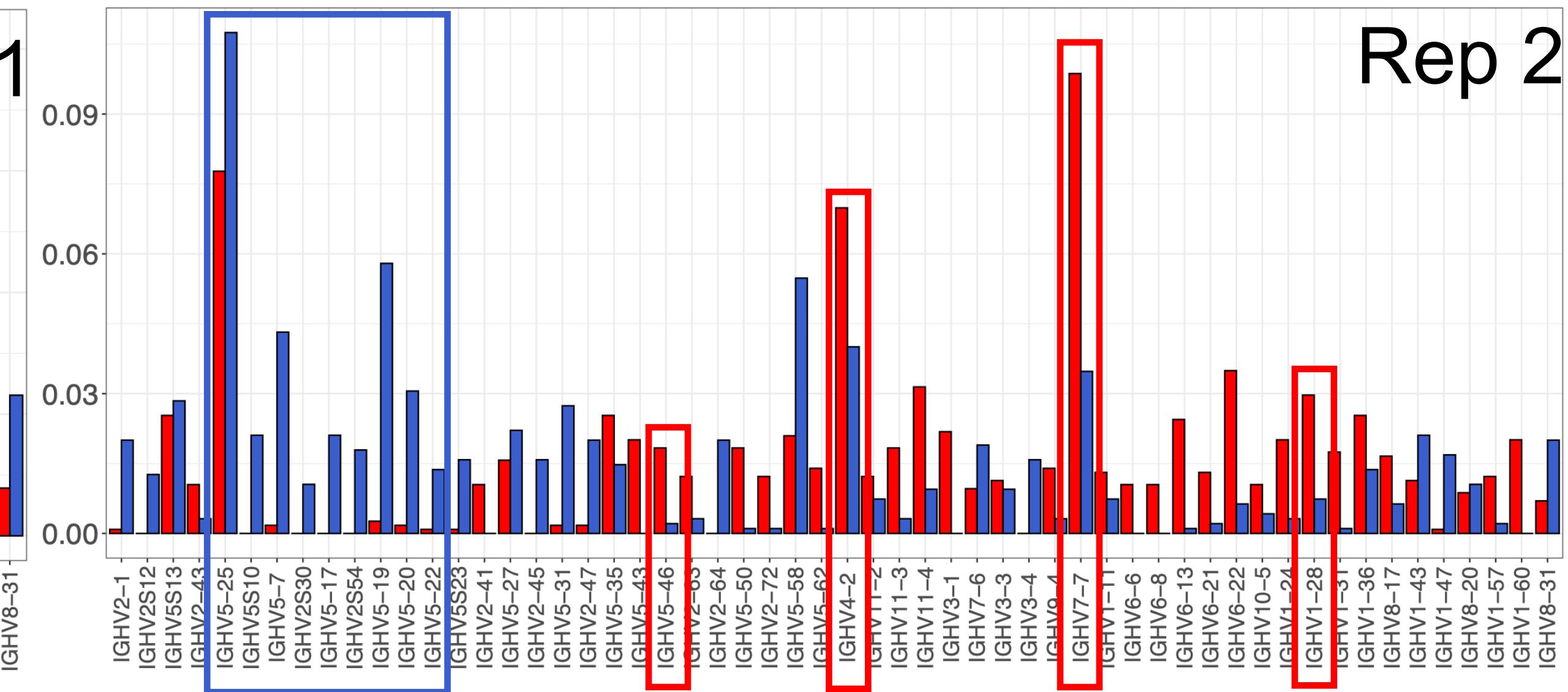
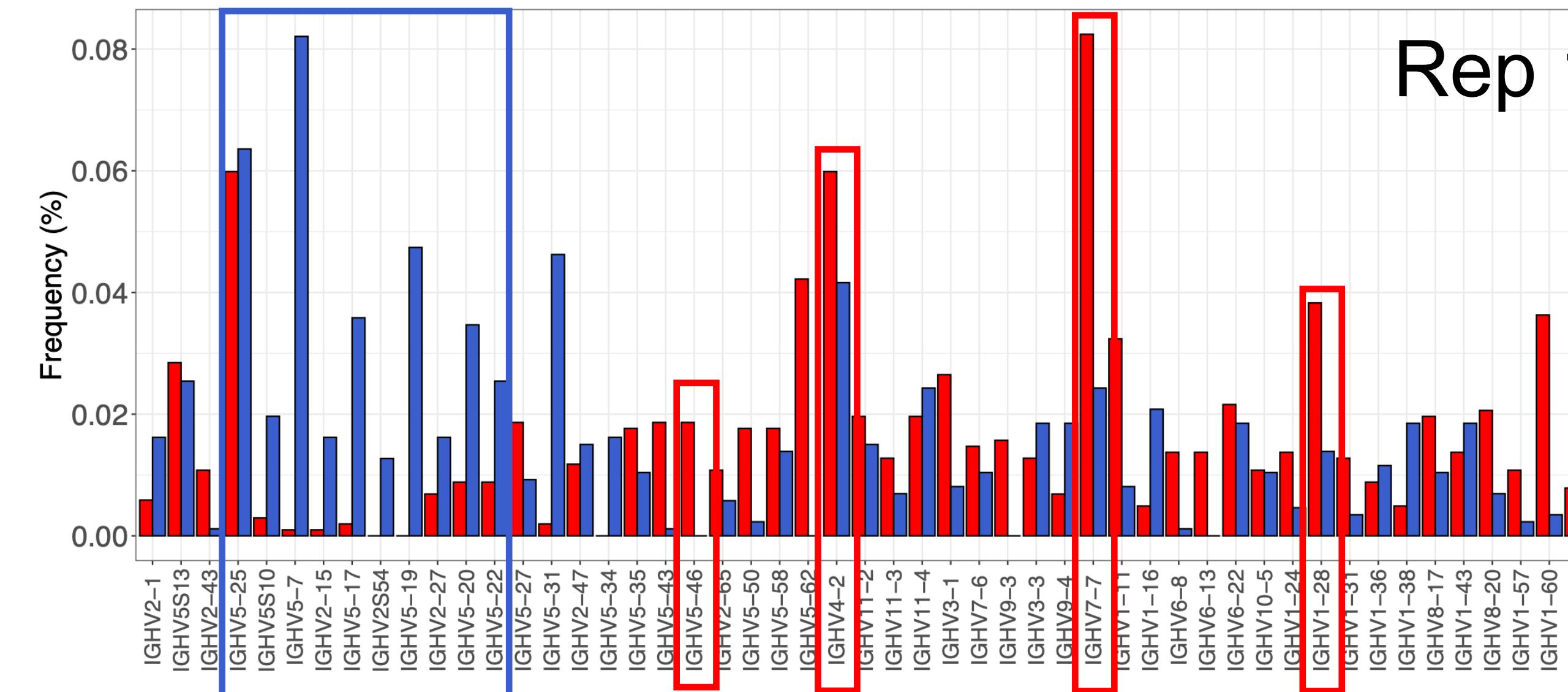
Rep 4



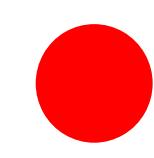
Expression from
SHR-A3 haplotype

Expression from
SHR-B2 haplotype

Haplotype specific utilization of Variable (IGHV) genes persists across time (40 weeks of age)



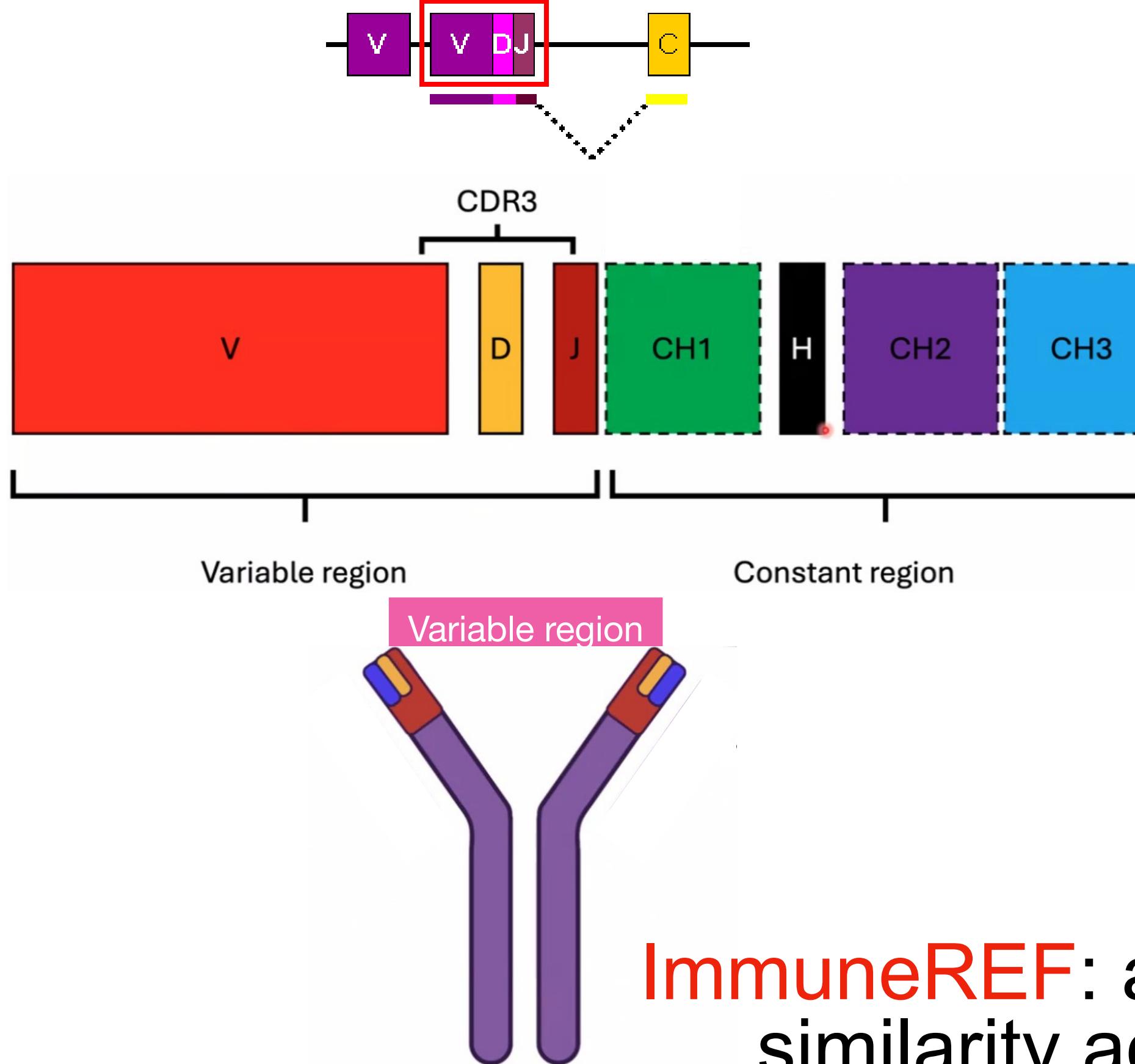
Expression from
SHR-A3 haplotype



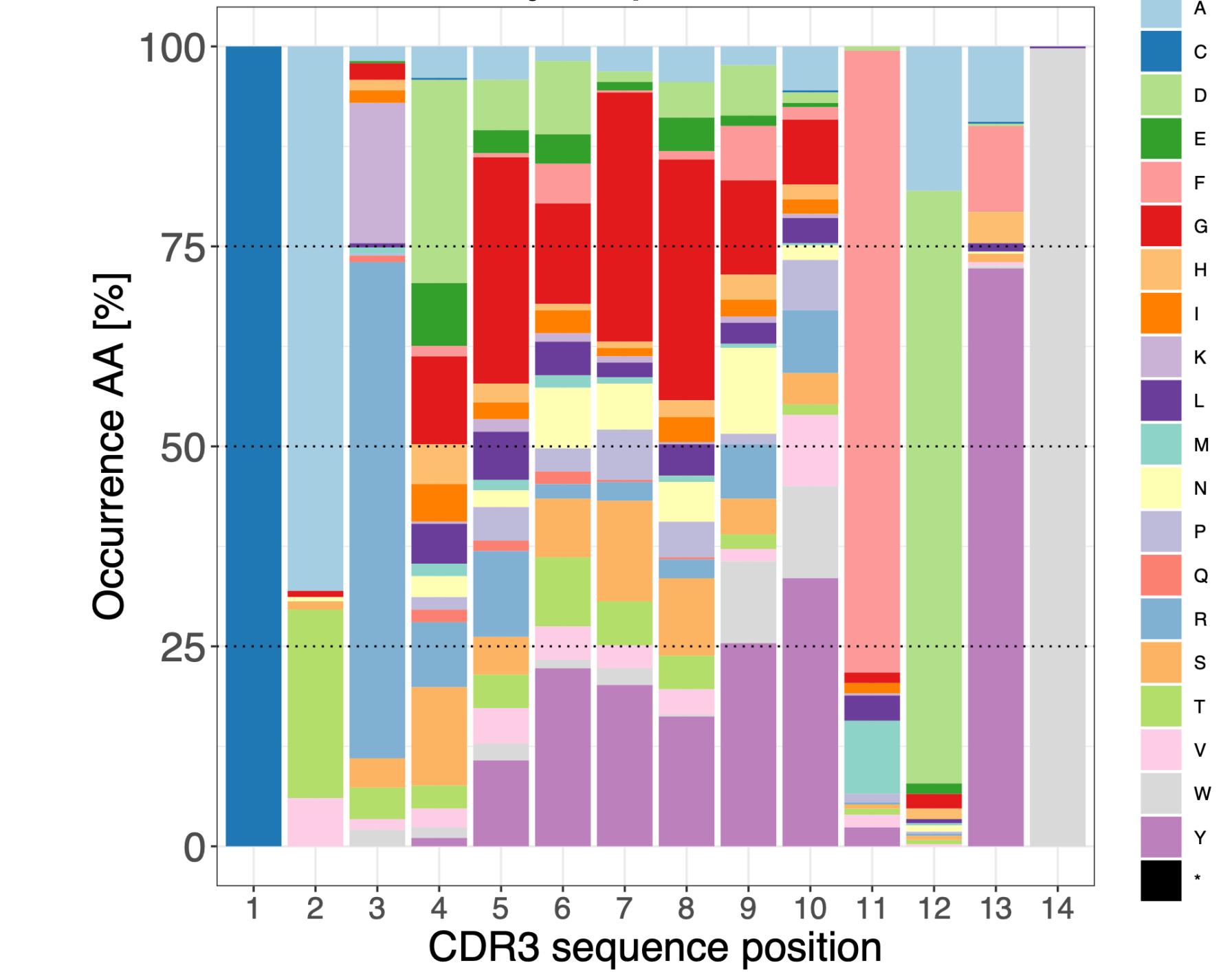
Expression from
SHR-B2 haplotype

Affinity maturation: somatic hypermutation in B cells targets residues in the CDR3 antigen recognition site to maximize diversity and minimize autoreactivity

More VDJ diversity by somatic hypermutation



Amino Acid diversity at positions 1 thru 14 of CDR3



ImmuneREF: a measure for quantifying immune repertoire similarity across multiple immune repertoire features

Cell Reports Methods

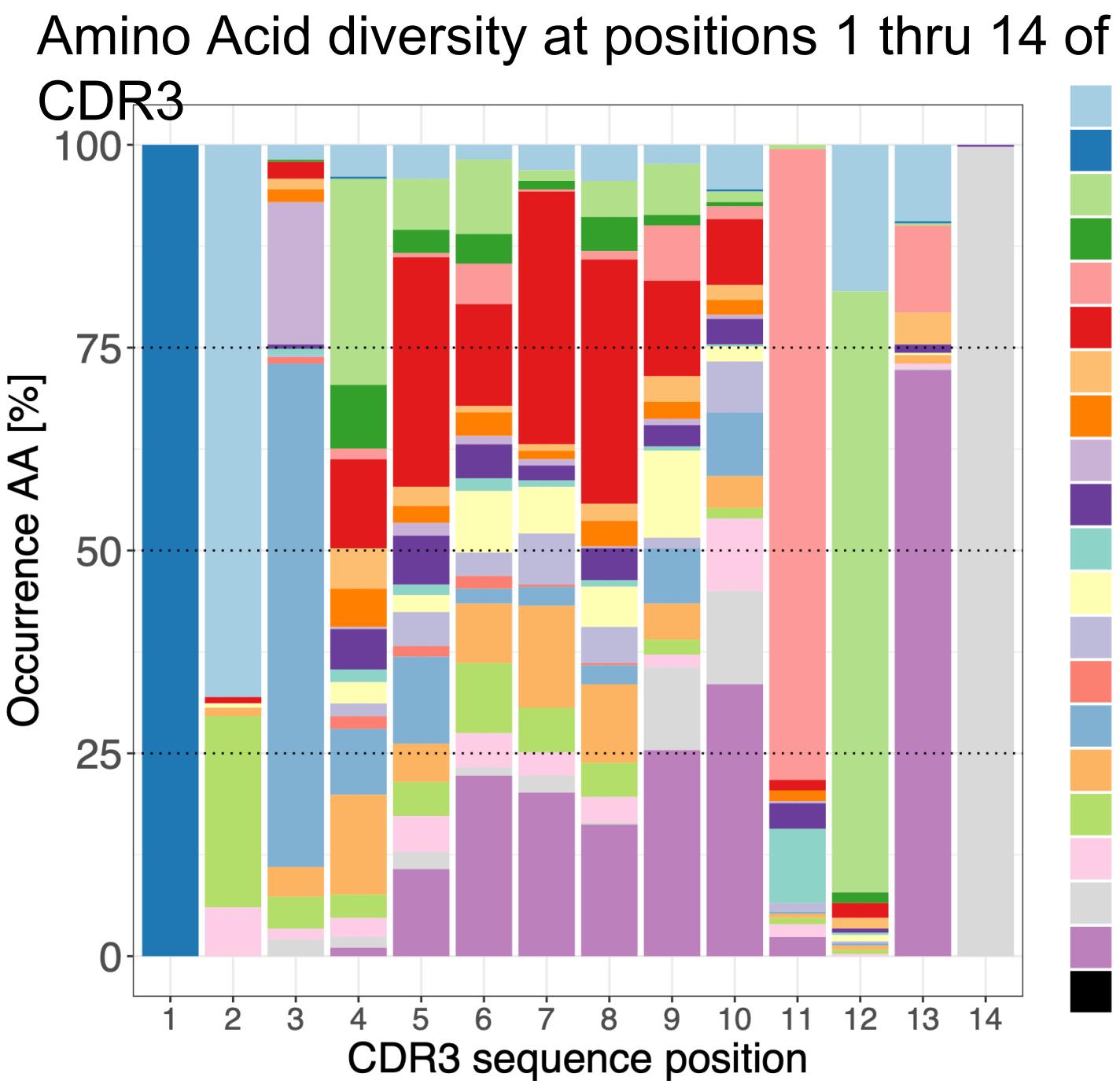
Cell Reports Methods 2, 100269, August 22, 2022

Reference-based comparison of adaptive immune receptor repertoires

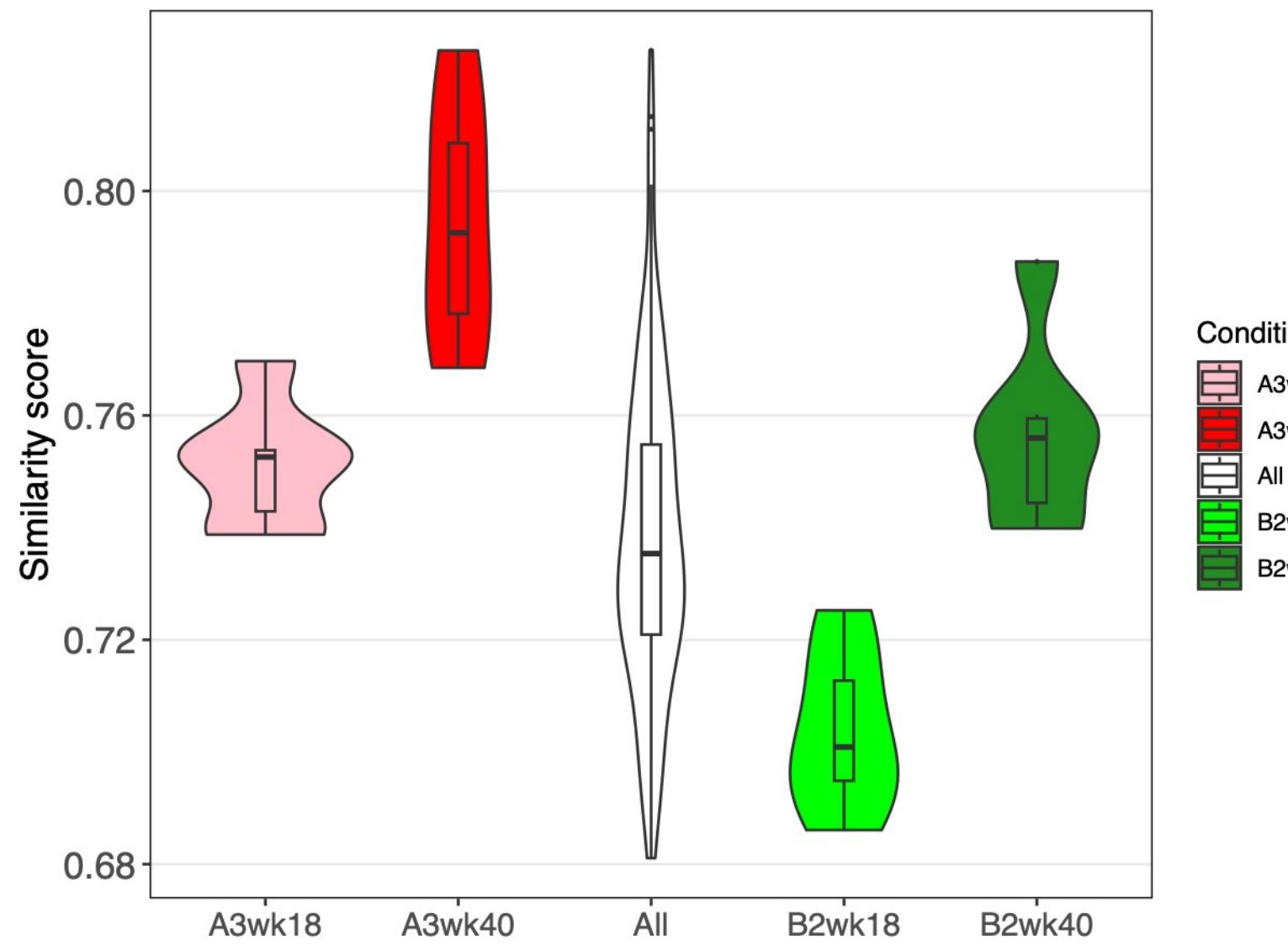
Cédric R. Weber,¹ Teresa Rubio,² Longlong Wang,^{3,4} Wei Zhang,^{3,5} Philippe A. Robert,⁶ Rahmad Akbar,⁶ Igor Snapkov,⁶ Jinghua Wu,³ Marieke L. Kuijjer,^{7,8,9} Sonia Tarazona,¹⁰ Ana Conesa,¹¹ Geir K. Sandve,¹² Xiao Liu,^{3,13,14} Sai T. Reddy,^{1,14} and Victor Greiff^{6,14,15,*}

ImmuneREF analysis

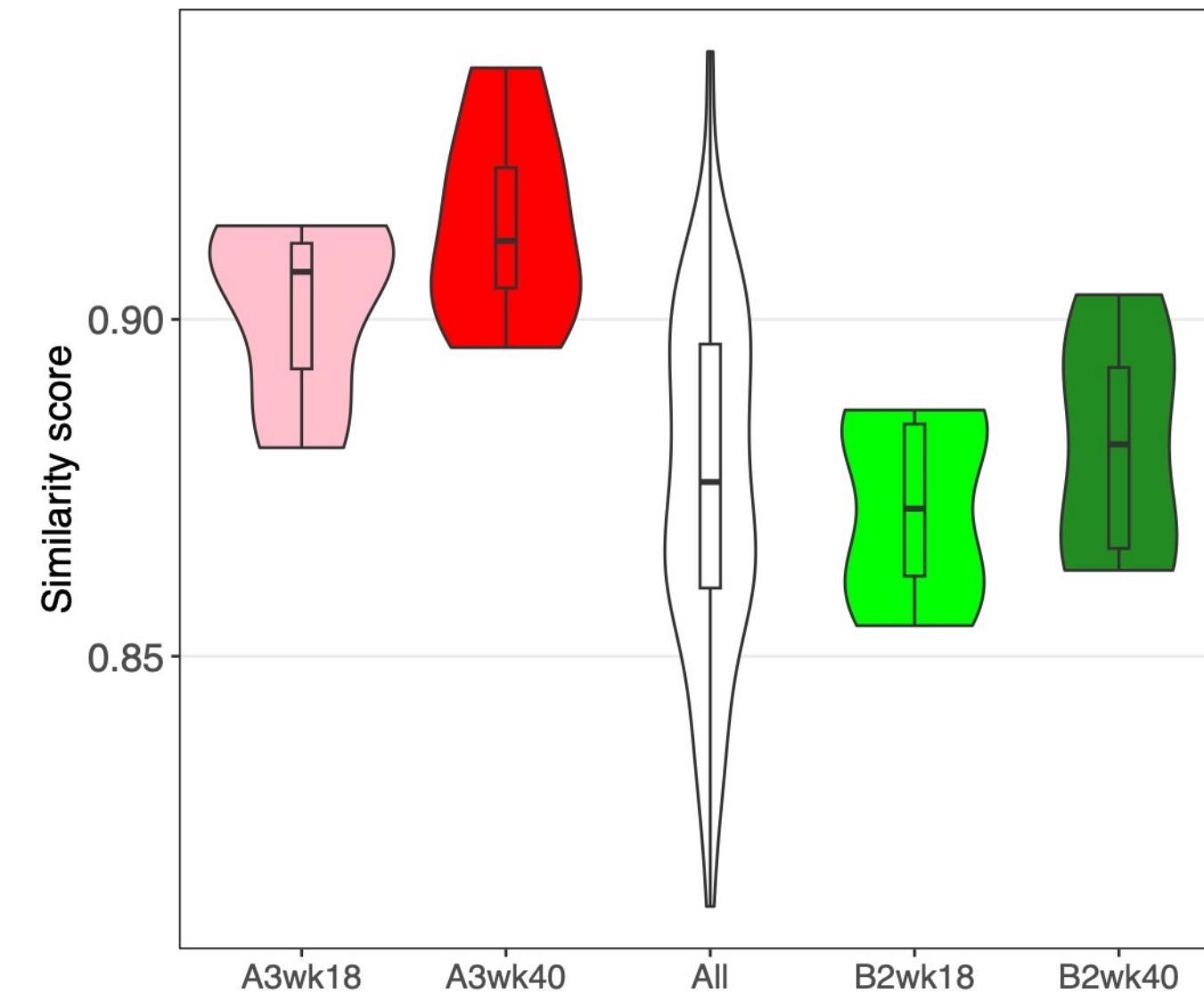
Studying similarity in the repertoire
as a population of CDR3 sequences



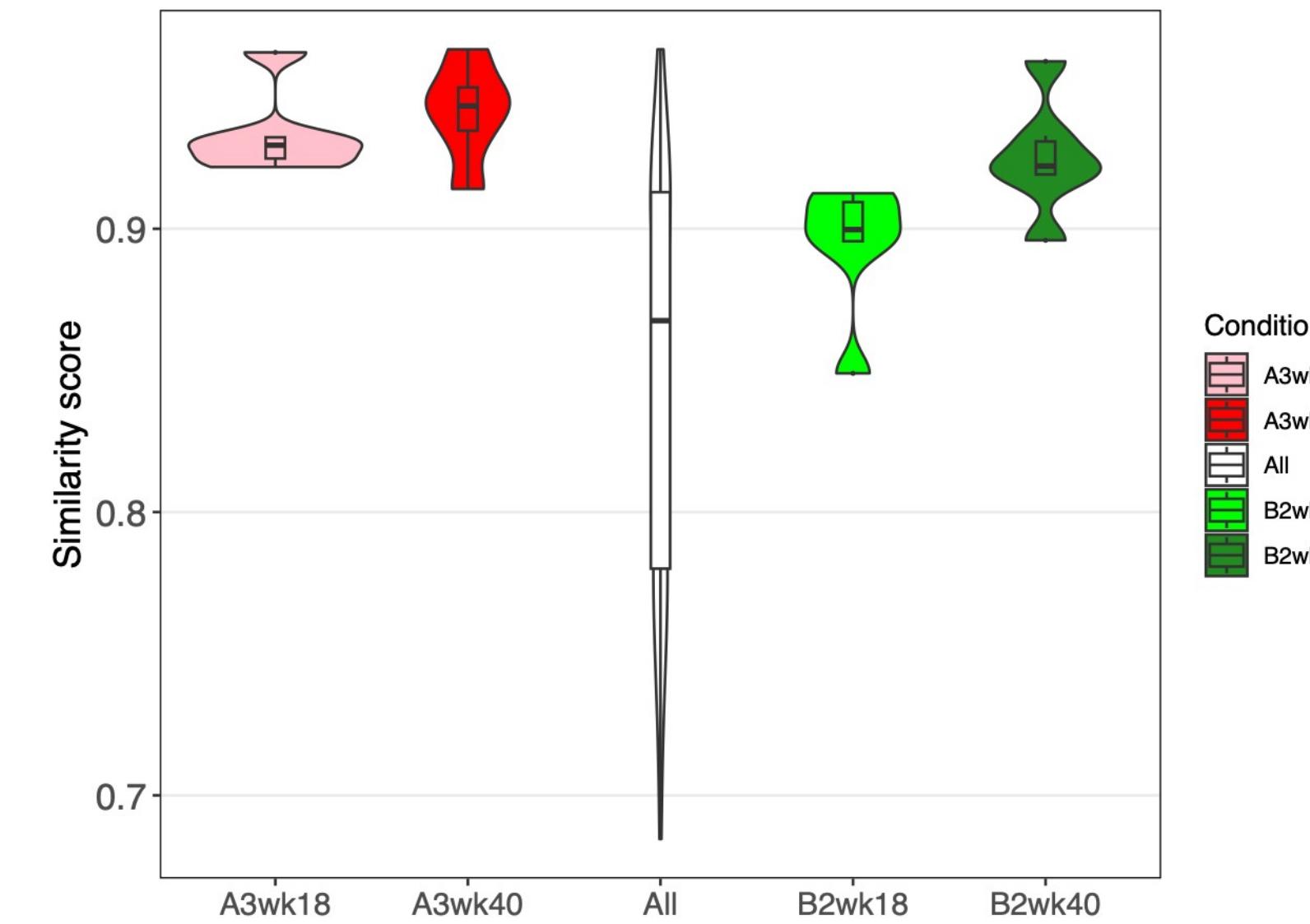
CDR3 amino acid frequency



CDR3 aa k-mer frequency



VDJ usage



Conclusions

- A congenic model containing two parental IGH haplotypes can be used to examine genetic effects on immunoglobulin gene expression
- PacBio Kinnex can provide quantitative analysis of IGH gene expression in lymphatic tissue
- Haplotype specific complete (VDJC) transcript reads can be obtained for multiple IGH isotypes (IgA, IgG1, IgG2a, etc)
- Quantitative expression effects arising in the haplotype of origin are observed for multiple IgH isotypes (IgA, IgG1, IgG2a, etc)
- Quantitative effects arising in the haplotype of origin are observed for Variable gene utilization
- CDR3 analysis reveals haplotype specific patterns in the antibody repertoire by strain and age

Thanks to:



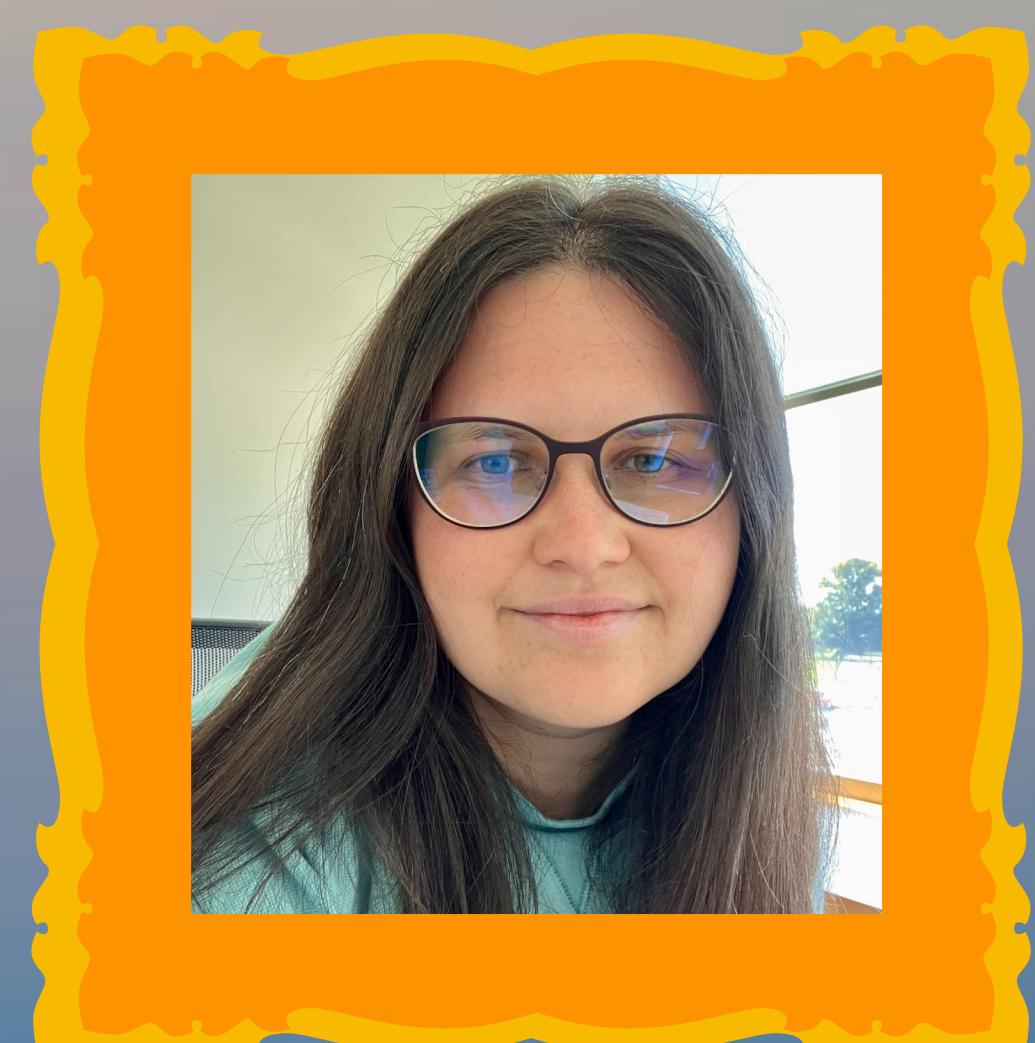
Melissa Smith
University of Louisville
School of Medicine



Ted Kalbfleisch
University of Kentucky



Yaming Zhu
UT Houston



Yana Safanova
Penn State