Identification and characterization of the T cell receptor (TCR) repertoire of the Cynomolgus macaque (*Macaca Fascicularis*)

Swati Jaiswal[1], Shayla Boyce[1], Sarah K. Nyquist[2,4,8], Tasneem Jivanjee[2], Samira Ibrahim[2], Joshua D. Bromley[2,3], G. James Gatter[2], Hannah P. Gideon[5], Kush V. Patel[5], Sharie Keanne C. Ganchua[5], Bonnie Berger[4], Sarah M. Fortune[6], JoAnne L. Flynn[5], Alex K. Shalek[2,6], Samuel M. Behar[1]*

1. Department of Microbiology and Physiological Systems, University of Massachusetts Medical School, Worcester, Massachusetts, USA.
2. Institute for Medical Engineering & Science, Massachusetts Institute of Technology, Cambridge, MA, Ragon Institute of MGH, MIT, and Harvard, Cambridge, MA, Broad Institute of MIT and Harvard, Cambridge, MA
3. Microbiology Graduate Program, Massachusetts Institute of Technology, Cambridge, MA
4. Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA
5. Department of Microbiology and Molecular Genetics and Center for Vaccine Research, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA
6. Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA, Ragon Institute of MGH, MIT and Harvard, Boston, MA, USA
7. Koch Institute for Integrative Cancer Research, MIT, Cambridge, MA, USA
8. Program in Computational and Systems Biology, Massachusetts Institute of Technology, Cambridge, MA, USA

Short Title: Cynomolgus TCR loci

Keywords: T cell receptor, Cynomolgus macaque, Locus map, NHP, Variable gene

*Correspondence:

Samuel M. Behar

E-mail address: samuel.behar@umassmed.edu

Abstract

Background: Non-human primates (NHP) are desirable as animal models of human disease because they share behavioral, physiological, and genomic traits with people. Hence, NHP recapitulate manifestations of disease not observed in other animal species. The *Macaca fascicularis* (i.e., Cynomolgus macaque) is an NHP species extensively used for biomedical research, but the TCR repertoire hasn't been characterized yet.

Result: We used the genomic sequences to design primers to identify the expressed TCR repertoire by single cell RNAseq. The data analysis from 22 unique samples were used to assign a functional status to each TCR genes. We identified and analyzed the TRA/D, TRB and TRG loci of the Cynomolgus macaque.

Conclusion: The genomic organization of the Cynomolgus macaque has great similarity with *Macaca mulatta* (i.e., Rhesus macaque) and they shared >90% sequence similarity with the human TCR repertoire. These data will facilitate the analysis of T cell immunity in Cynomolgus macaques.

Background

Experimental animal models are an essential tool in our pursuit of understanding human physiology. The mouse has been incredibly useful in elucidating the major concepts of immunology, including defining the genetic and molecular basis of immunoglobulin and TCR formation and diversity. As part of this effort, the murine TCR repertoire have been extensively characterized and its knowledge is being used to develop new approaches to facilitate antigen discovery and novel treatments for human disease. However, it is not surprising that many human diseases are inadequately modelled in mice. This has been repeatedly emphasized for cancer and is also true for many infectious diseases. Two important examples are acquired immunodeficiency syndrome (AIDS), which is caused by the Human Immunodeficiency Virus-1 (HIV-1), and COVID-19, which is caused by the SARS-CoV2 coronavirus [1-5]. Mice are naturally resistant to both infections. For HIV research, the field largely turned to nonhuman primates (NHP) as a better alternative because they could be infected with a highly related virus, Simian Immunodeficiency Virus (SIV). Consequently, the Rhesus macaque's TCR locus was among the first NHP TCR locus to be characterized [6]. Cynomolgus macaques have been increasingly used for biomedical research, especially in the fields of neurology, cardiology, and for drug development [7, 8]. Importantly, they are increasingly used for infectious disease research, including as a model for human HIV [9] and SARS-CoV2 infection [5]. Most NHP species, including rhesus macaques, whether in captivity or in the wild, rapidly succumb to *Mtb* infection [10, 11]. However, Flynn's group finds that following challenge with *Mtb*, 50% of infected Cynomolgus macaques develop a form of disease that resembles latent TB in people [12-15]. Indeed, the pathology observed among *Mtb*-infected Cynomolgus macaques *recapitulates* the entire spectrum of pathology of human TB granulomas [16]. Thus, the Cynomolgus macaque is providing insights into human disease not possible with other small animal models.

The tremendous capacity of T cells to recognize diverse antigens has a genetic basis that is inherent in the genomic organization of the T cell receptor (TCR) loci [17]. TCR repertoire diversity arises through genetic mechanisms that minimize the number of genetic elements encoded by the genome while maximizing the potential breadth of expressed TCRs. The germline configuration of TCR genes is not functional. Instead, the TCR loci encode families of variable (V), diversity (D), and joining (J) segments, which undergo rearrangement early during T cell development [17]. Recombination of V, D, and J segments leads to a gene fragment that encodes the V-region domain, which becomes the N-terminus of the TCR protein and determines its antigen specificity. Downstream of the V, D, and J genes are constant (C) region exons, which encode the C-terminus of all TCRs and couples the TCR to the Cluster of differentiation 3 (CD3) protein complex to mediate signal transduction into the cell. The primary diversity of TCRs arises from the nearly random assortment of V, D, and J gene segments, as well as additional diversity that occurs at the V-D and D-J junctions by imprecise recombination and the insertion of non-germline encoded nucleotides (N-regions). TCRs are heterodimers formed by TCRα and TCRβ chains, which are encoded by distinct loci (TRA and TRB, respectively) [18]. The TCRα is encoded when Vα and Jα gene segments recombine; the TCRβ is formed from the recombination of Vβ, D and Jβ gene segments. Additional diversity is created by the random pairing of the TCRα and TCRβ chains. Unlike immunoglobulin genes, somatic mutation does not occur in TCR genes. The potential TCR repertoire varies between animal species and is driven in large part by the number of functional members of V, D, or J segments. In humans, there is the potential to generate 10^{15} unique TCRs.

A second subset of T cells are known as Gamma-delta (γδ) T cells, express an alternative TCR, which is encoded by distinct gene segments found in the TRG and TRD loci. The γδ-TCR is structurally similar to the αβ-TCR. Like the TRA and TRB loci, the TRG and TRD loci contain sets of Vγ and Jγ, and Vδ, Dδ and Jδ gene segments, respectively. In general, there are fewer gene segments in the TRG and TRD loci [19]. γδ T cells remain enigmatic because

the antigens they recognize and the antigen presenting molecules that restrict their recognition of antigen are incompletely characterized. Nevertheless, they are identified in the circulation and in the tissues of all mammals, and play important roles in autoimmune disease, and in immunity to infection and cancer [20, 21].

Here we identified the TRA, TRB, TRG and TRD loci of the Cynomolgus macaque. Based initially on the homology with human TCR gene segments, and subsequently using the identified gene segments from Rhesus macaque and Cynomolgus macaque, we systematically identified all the V, D, J, and C gene segments belonging to all four T cell receptor loci. Finally, using the genomic sequences, we designed specific primers for the amplification of the Vα and Vβ regions, and determined which of the V gene segments are expressed in individual subjects. These data will allow the detailed analysis of the T cell responses in Cynomolgus macaques as well as comparative immunogenetics studies, comparing different species of Cynomolgus macaques, as well as the evolution of TCR genes among the primates.

Results

Identification of the *Macaca fascicularis* (macfas) TCR loci

Based on nucleic acid sequence homology with the human Cα, Cδ, Cβ, and Cγ gene segments, the TRA and TRD loci were identified on Chr.7, and TRB and TRG loci were identified on Chr.3 (Figure 1). Subsequently, each human V, D, J, and C gene segment was used to blast the macfas Chr.7 and 3, to identify homologous gene segments. Similarly, *Macaca mulatta* (macmul) gene segments were also used to identify homologous genes unique to the macaca genus. Using this approach, we were able to annotate and assemble a map of the macfas TRA, TRB, TRG, and TRD loci as described in detail below.

The macfas TRA locus

The structure of the macfas TRA locus is like the human locus in that it overlaps the TRD locus on Chromosome 7 (Figure 1A) [22]. We identified 62 TRAV genes in macfas, one more than the 61 human genes but less than the 67 macmul genes. The two human gene families TRAV7 and TRAV28, each contain a single member and are absent from the macfas and macmul TRA locus (Table 1, Figure 2). Conversely, the TRA loci of macfas and macmul have one additional gene in the TRAV24, TRAV25, and TRAV26 families. The greater number of macmul TRAV genes compared to macfas results from an expansion of the TRAV22 and TRAV23 families, from one member to three and four, respectively (Table 1). Of the 62 macfas TRAV genes, 12 are pseudogenes and 2 are ORFs (Table 1, Table S1). There might have been a gene duplication of the TRAV genes TRAV24, TRAV25, TRAV26, which differentiates the human TRAV locus from the macaque locus (Figure 2B). Second, there are additional members of the TRDV1. TRAV22, and TRAV23 families in NHP, compared to the human TRA/DV locus. We did not find macfas homologs of macmul TRDV1-1, TRAV22-2, TRAV22-3, TRAV-23-2, TRAV23-3, or TRAV23-4, despite searching the macfas genome using the macmul homologs. We believe the differences between the macfas and macmul genomic sequences could have

143 arisen from gaps in the known macfas sequence or problems with the genome assembly.

144 We identified 61 TRAJ genes, which is the same number as Rhesus macaque and
145 human TRAJ genes. There is a high degree of conservation between macfas and *Homo*
146 *sapiens* (homsap) TRAJ gene segments (Table S2). Finally, we compared the TRAC exons
147 from all three species. The macfas and macmul TRAC genes have identical amino acid
148 sequences (Figure S1).

149

150 **The macfas TRB locus**

151 The macfas TRB locus (Figure 1B) is similar in structure to the macmul TRB locus. We
152 identified 78 TRBV genes, compared to 77 annotated macmul TRBV genes (Table 1 & Table
153 S3). Both are expanded compared to the human species, for which there exists 68 distinct
154 genes. The overall TRBV family structure is similar, with some variation in the number of
155 members and the number of pseudogenes (n=17) and ORFs (n=2) (Table 1, Figure 3). The
156 organization of the TRBJ and TRBC genes is similar in all three species, characterized by a
157 duplication of the TRBJ and TRBC genes (Figure 1B). Comparing the macfas and macmul
158 TRBJ gene segments, four (including the TRBJ2.2P ORF) differ by a single nucleotide; the
159 other 10 genes are 100% conserved (Figure S2, Table S4). The TRBD1 and TRBD2 are also
160 100% conserved between macfas and macmul (Table S4). Similarly, there is a high degree of
161 conservation between macfas and homsap TRBJ gene segments (Figure S2). Finally, we
162 compared the TRBC exons from all three species. As noted, there are two TRBC genes,
163 TRBC1 and TRBC2, which are 97% identical. The macfas and macmul TRBC1 differ by only
164 two bp and the translated sequence is 100% identical; for TRBC2, there is a single aa difference
165 (Figure S1).

166

167 **The macfas TRG locus**

168 The macfas TRG locus is located on chromosome 3 (Figure 1C). We identified 12

TRGV genes of which 6 are predicted to be functional and an additional 4 are pseudogenes (Figure 4, Table 1, Table S5). These genes were compared to the homologous genes in human and rhesus (Figure 4). The same 12 genes were found in the macmul TRG locus. We named macfas TRGV4 based on its homology with homsap TRGV4*01 (92% homology, Figure 4). The ortholog in the macmul TRG locus is identical in sequence, but IMGT names it TRGV8, although it has only 88% homology to homsap TRGV8*01. In general, the macfas and macmul orthologs had between 0-2 mismatches (i.e., >99% homology), while the homology between macfas and homsap TRGV genes was 88-95%. The two NHP species lacked TRGV5, TRGV5P, TRGV7, and TRGV8, and macmul had two additional V genes, TRGVC and TRGVD. The human TRG locus has two clusters of J segments and C-region genes [22, 23]; IMGT/LIGM-DB: IMGT000011 (582960 bp), human (Homo sapiens) TRG locus), and the macmul locus has a similar structure (IMGT/LIGM-DB: IMGT000059 (197016 bp), Rhesus monkey (Macaca mulatta) TRG locus). While there are five macmul TRGJ gene segments, we detected only three in the macfas TRG locus. These three are more like the macmul TRGJ2-1, 2-2, and 2-3 gene segments (Figure 4B). Similarly, there is a single macfas TRGC region gene, and its amino acid sequence is 91.9% and 96.5% identical to macmul TRGC1 and TRGC2, respectively (Figure S1). There is 1, 0 and 2 mismatches between macfas and macmul TRGC2 exon 1, 2, and 3, respectively. The human TRGC2 exon 2 contains a duplicated sequence, which neither the macfas nor macmul exon 2 contains (Figure S2). Therefore, the macfas TRGC gene is the TRGC2 ortholog, and the genomic assembly of macfas is missing a region that spans TRGJ1-1 to TRGC1 (Figure 4C).

The macfas TRD locus

The macfas TRD locus is located on chromosome 7 and overlaps with the TRA locus (Figure 1A). Three canonical TRDV genes were identified as macfas homologs of homsap TRDV1, TRDV2, and TRDV3, with homologies between 91-97% (Figure 5, Table S6). A fourth

gene, TRDV4, was identified which was 100% homologous to macmul TRDV4. The macmul genome has a fifth gene, TRDV1-1, which is very homologous to TRDV1 (Figure 2, 5); no macfas orthologs were found for this gene. Three TRDD and four TRDJ macfas gene segments were identified, as in the homsap genome (Table S6). These genes are 100% identical to their macmul homologs. Similarly, the single macfas TRDC region has 100% DNA sequence identity and predicted amino acid sequence as the macmul TRDC (Figure S1). There is a two amino acid gap, which we suggest is a consequence of the artificial splicing between exons 2 and exon 3.

The expressed V gene repertoire used by Cynomolgus macaque T cells

To determine the functionality of the TRAV and TRBV gene segments we identified, the following criteria were used: (i) Defined L1 exon and L2-V exon, (ii) absence of nonsense or missense mutation, and (iii) encodes a cytosine (C) at position 21-23 followed by tryptophan (W) at position 31-33 of the 3' end. The terminal amino acids encoded by a functional TRAV gene is usually CAVR, CAL, or CAF. Similarly, the terminal amino acids encoded by a functional TRBV gene is usually CASSQ, CASSL, or CASSE. Based on these criteria, we initially assigned each V gene to be functional if it met these criteria. If the gene had an internal stop codon, or lacked the conserved C or W residue, it was deemed a pseudogene. Finally, if the gene appeared to be functional, but the L1 or L2 parts of the leader sequence could not be identified, or it lacked consensus splice site for intron A, we designated it an open reading frame (Supplemental tables 1, 3, and 5).

To determine the expressed TRAV and TRBV repertoire, TCRs from Cynomolgus macaques infected with *Mycobacterium tuberculosis* were analyzed. The distribution of the expressed TRAV and TRBV genes in uninvolved lung from these infected subjects was determined for 22 unique individuals (Figure 6). The distribution of TRAV and TRBV genes was also determined in BAL cells, before or after infection, involved (i.e., granulomas) lung tissue,

and single cell lung suspensions (Figure S3). The presence of stop codons in pseudogenes that were observed to be transcribed were confirmed (data not shown). One exception was detected. For TRBV6-4, the expected stop codon at position 85 (TAG) was CAG in the transcribed gene, and thus, encoded a functional glutamine (Q). This difference between the germline and the transcribed gene could be the result of a polymorphism or a sequencing error in the genomic reference sequence. Finally, the status of V genes designated as ORFs, was changed to ‘functional’ if the V gene was transcribed.

Discussion

The nucleic acid sequence of recombined V, D, and J gene segments encodes the protein structure of the TCR and contains immunological information about T cell responses. The complementarity determining region 3 (CDR3), defined as the V-D-J or V-J recombination site, is unique to each unique T cell clone, sometimes referred to as a clonotype. Analytical approaches are beginning to predict the antigen specificity based on the primary sequence of the TCR. In the absence of the antigen specificity, the TCR sequence can be used as a surrogate of antigen specificity. As T cells undergo clonal expansion after encountering antigens, TCR sequences are being used to track T cells, monitor immune responses, and identify new antigens for human tumors and pathogens [24-26]. Advances in T cell therapy are being driven by our ability to clone and recombinantly express TCRs, as exemplified by adoptive cell therapy (ACT) [27, 28]. Thus, defining the V, D, and J gene segments is an important step in the analysis of T cell immunity.

We identified and annotated the TRA, TRB, TRG, and TRD loci of the Cynomolgus macaque. There is generally more than 90% homology between the different V, D, and J gene segments in the Human, Rhesus and Cynomolgus macaque's TCR repertoire. As one might expect, the structure of the different TCR loci is highly conserved between Rhesus and Cynomolgus macaque. The differences we detected (e.g., Fig.2A) are more likely to be due to ascertainment bias arising from problems with genomic sequencing and assembly, than true evolutionary events. In support of this conclusion, among the expressed TCR repertoire, we found many macfas TCRs expressed in the lung matched to macmul reference sequences that were missing from the macfas genomic sequence. We also find that there is expansion of TCR beta locus of macfas and macmul compared to homsap. These differences, which are likely to have occurred by gene duplication [29, 30], may have occurred in response to changes in selective pressure during evolution of the TCR loci [31, 32].

Conclusions

We identified and annotated the TRA/D, TRB and TRG loci of the Cynomolgus macaque. The TRA and TRB genomic sequences were used to design primers, and as reference sequences, to amplify and identify TCR sequences expressed by single cells from the lungs of Cynomolgus macaques. By using these data to analyze the αβTCRs expressed by mature T cells, we were able to discern which V genes were functional based on their RNA expression. This allowed us to refine and validate our predictions based on the genomic sequences. Altogether, these data show the utility of these TCR reference sequences, and we expect that they will be useful for the study of T cell immunity in Cynomolgus macaques.

Methods

Source of genomic sequence. The genome of the Cynomolgus macaque (NCBI: taxid 9541), also known as the crab eating macaque, has been sequenced and we used the reference genome Macaca_fascicularis_5.0 available from NCBI (https://www.ncbi.nlm.nih.gov/genome/776) [33, 34]. The formal genus and species name is *Macaca fascicularis*, which we abbreviate as macfas. The Rhesus macaque (i.e., *Macaca mulatta*; macmul) TCR sequences were obtained initially from the literature [35] and later from IMGT (http://www.imgt.org) [36]. The Human (i.e., *Homo sapiens*; homsap) TCR sequences were obtained from IMGT. In cases where more than one allele was available, the first allele was used for sequence comparisons.

Annotation and analysis of Macfas TCR repertoire. To identify the location of the macfas TCR loci, the human TRAC, TRBC, TRGC, and TRDC were blasted against the macfas genome. Subsequently, all human gene segments were individually blasted against the macfas genome. As macfas gene segments were defined, they were also used to look for other

281 homologous genes. At the beginning of this study, the sequences of the macmul TCBV genes
282 were available and were used to look for homologous genes [35]. The names of the genes were
283 assigned based on the homology with the human genes, and the location in the genome. The
284 leader sequence (L1 & L2), TRV region, D region and J chain were identified for each gene. The
285 annotation was done by following standard IMGT rules (http://www.imgt.org). Clustal Omega
286 was used for multiple sequence alignments (https://www.ebi.ac.uk/Tools/msa/clustalo/) [37] and
287 visualized using Archaeopteryx for Figures 2–5 [38]. Sequences were entered and tracked in
288 SnapGene (version 5.0)
289
290 **Expressed TCR repertoire of Cynomolgus macaques.** Cells from bronchoalveolar lavage
291 (BAL), single cell suspensions of lung, or lung tissue, were obtained from Cynomolgus
292 macaques infected with *Mycobacterium tuberculosis* and single cell RNAseq libraries were
293 created [39]. Primers were synthesized that were specific for the different TRAV, TRBV, TRAC,
294 and TRBC gene segments based on the genomic sequences described herein and used to
295 enrich and amplify the TCR sequences from T cells in scRNA-Seq libraries generated using 3'
296 barcoded Seq-Well [40, 41]. Primers were not designed for pseudogenes that had internal stop
297 codons, or for some V genes that were not initially identified. The libraries were sequenced and
298 then aligned to the TCR reference sequences. The samples were analyzed for 48 TRAV and 73
299 TRBV genes. The V region and J region sequences were mapped using BOWTIE 2 as part of
300 the TCRGO algorithm (https://github.com/ShalekLab/tcrgo/tree/master/tcrgo)[41]. Briefly, reads
301 are aligned with the V and J regions in the reference TCR database, containing the sequences
302 annotated in this report (see Results, below). Each read from a Seq-Well library includes nucleic
303 acid tags that identify the cell of origin (cell barcode) and the transcript of origin (unique
304 molecular identifier, UMI). Reads with matching cell barcode and UMI are merged, and a
305 consensus V and J region mapping is determined based on sequence similarity identified
306 among the majority of reads. A consensus CDR3 sequence is identified from reads with shared

307 mappings.

308

309

List of abbreviations

Adoptive cell therapy (ACT)

Acquired immunodeficiency syndrome (AIDS)

Constant (C)

Cluster of differentiation 3 (CD3)

Complementary determining region 3 (CDR3)

Diversity (D)

Human Immunodeficiency Virus-1 (HIV-1)

Joining (J)

Gamma-delta (γδ)

Macaca fascicularis (macfas)

Macaca mulatta (macmul)

Mycobacterium tuberculosis (*Mtb*)

Non-human primates (NHP)

Simian Immunodeficiency Virus (SIV)

T cell receptor (TCR)

Variable (V)

Declarations

Ethics approval. All experiments, protocols, and care of animals were approved by the University of Pittsburgh School of Medicine Institutional Animal Care and Use Committee (IACUC). The Division of Laboratory Animal Resources and IACUC adheres to national guidelines established by the Animal Welfare Act (7 U.S. Code Sections 2131-2159) and the Guide for the Care and use of Laboratory Animals (Eighth Edition) as mandated by the U.S. Public Health Service Policy. The DLAR program is AAALAC accredited. Animals used in this study were housed in rooms with autonomously controlled temperature and provided enhanced enrichment procedures.

Consent for publication. Not applicable

Availability of data and materials. All data generated or analyzed during this study are included in this published article [and its supplementary information files]

Competing interests. A.K.S. reports compensation for consulting and/or SAB membership from Merck, Honeycomb Biotechnologies, Cellarity, Repertoire Immune Medicines, Ochre Bio, Third Rock Ventures, Hovione, Relation Therapeutics, FL82, and Dahlia Biosciences.

Funding. This project has been funded in part with Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under Contract No. 75N93019C00071, and additional support from The Bill and Melinda Gates Foundation.

Authors' contributions.

S.J. and S.M.B., conceptualized and wrote the manuscript; S.B., helps in preliminary data acquisition; S.K.N., contributed to methodology and performed sc RNA seq data analysis;

S.K.G., K.P., H.G. NHP Sample processing and facilitated transfer of the samples; T.J., S.I., J.B., and G.J.G., have contributed in T cell processing for scRNA sequencing; A.K.S. and B.B., provided resource and supervision for scRNA sequencing and data analysis; Funding Acquisition, J.F., S.M.F., S.M.B.; S.J. and S.M.B. have reviewed and edited the manuscript. All authors reviewed the manuscript.

Acknowledgements.

Roisin Floyd, Marc Wadsworth, and Travis Hughes have performed original sequence libraries for depletion experiment that were helpful in analyzing the expressed repertoire, Jake Rosenburg and Andy Tu helped in the development of TCR pipeline and interpretation of results.

List of Tables

Table 1: Comparison of genes (TRAV/ TRBV/ TRDV/ TRGV) in macfas, macmul and human

TRAV gene segments			
Subgroup	Macfas	Macmul	Human
TRAV1	2	2	2
TRAV2	1	1	1
TRAV3	1	1	1
TRAV4	1	1	1
TRAV5	1	1	1
TRAV6	1	1	1
TRAV7	0	0	1
TRAV8	7	7	8
TRAV9	2	2	2
TRAV10	1	1	1
TRAV11	3	3	2
TRAV12	3	3	3
TRAV13	2	2	2
TRAV14	2	2	2
TRAV15	1	1	1
TRAV16	1	1	1
TRAV17	1	1	1
TRAV18	1	1	1
TRAV19	1	1	1
TRAV20	1	1	1
TRAV21	1	1	1
TRAV22	1	3	1
TRAV23	1	4	1
TRAV24	2	2	1
TRAV25	2	2	1
TRAV26	3	3	2
TRAV27	1	1	1
TRAV28	0	0	1
TRAV29	1	1	1
TRAV30	1	1	1
TRAV31	1	1	1
TRAV32	1	1	1
TRAV33	1	1	1
TRAV34	1	1	1
TRAV35	1	1	1
TRAV36	1	1	1
TRAV37	1	1	1
TRAV38	2	2	2
TRAV39	1	1	1
TRAV40	1	1	1
TRAV41	1	1	1
TRAV46	1	1	1
TRAVA	1	1	1
TRAVB	1	1	1
TRAVC	1	1	1
Total	62	67	61

TRBV gene segments			
Subgroup	Macfas	Macmul	Human
TRBV1	3	3	1
TRBV2	3	3	1
TRBV3	4	4	2
TRBV4	3	3	3
TRBV5	10	10	8
TRBV6	10	10	9
TRBV7	11	11	9
TRBV8	2	2	2
TRBV9	1	1	1
TRBV10	3	3	3
TRBV11	3	3	3
TRBV12	4	4	5
TRBV13	1	1	1
TRBV14	1	1	1
TRBV15	1	1	1
TRBV16	1	1	1
TRBV17	1	0	1
TRBV18	1	1	1
TRBV19	1	1	1
TRBV20	1	1	1
TRBV21	1	1	1
TRBV22	1	1	1
TRBV23	1	1	1
TRBV24	1	1	1
TRBV25	1	1	1
TRBV26	1	1	1
TRBV27	1	1	1
TRBV28	1	1	1
TRBV29	1	1	1
TRBV30	1	1	1
TRBVA	1	1	1
TRBVB	1	1	1
TRBVC	1	1	1
Total	78	77	68

TRGV gene segments				
Subgroup	GENE	Macfas	Macmul	Human
TRGV1	TRGV1	1	ORF	ORF
	TRGV2	1	1	1
	TRGV3	1	1	1
	TRGV3P	0	NR	*
	TRGV4	1	0***	1
	TRGV5P	0	NR	P
	TRGV5	0	NR	1
	TRGV6	P	P	P
	TRGV7	0	NR	P
	TRGV8	0	1***	1
TRGV2	TRGV9	1	1	1
TRGV3	TRGV10	1	ORF	ORF
TRGV4	TRGV11	P	P	ORF
TRGVA	TRGVA	P	P	P
TRGVB	TRGVB	P	P	P
TRGVC	TRGVC	0	P	NR
TRGVD	TRGVD	0	P	NR
Total		12	12	14

TRDV gene segments				
Subgroup	GENE	Macfas	Macmul	Human
TRDV1	TRDV1	1	1	1
TRDV2	TRDV2	1	1	1
TRDV3	TRDV3	1	1	1
TRDV4	TRDV4	1	1	0
Total		12	12	14

P	pseudogene
ORF	open reading frame
NR	not reported
0	no homologous gene identified
***	nomenclature discrepancy

Notes:

1. The numerical value for every gene represents the number of allele present.
2. ORF: Open reading frame
3. NR: Not reported
4. P: Pseudogene
5. F: Functional
6. ***: Nomenclature discrepancy

Figure Legends

Figure 1: The Macfas TCR loci.

Structure of the TCR loci (A) TRA/TRD (B) TRB and (C) TRG loci. (A) The TRA and TRD loci are interspersed on Chr. 7. The genes above the x-axis belong to the TRA locus; those below the axis belong to the TRD locus. The boxed region is expanded to show greater detail. (B) The TRB locus is located on Chr. 3. The boxed region is expanded to show greater detail. (C) The TRG locus is located on Chr. 3. Each line represents a gene and the distance between them is proportional to their spacing on Chr.7 and Chr.3. The blue boxes represent the 3' region of the AMPH gene and exon 10 of STARD3NL, which are boundaries of the TRG locus. The black lines represent V gene, green lines are J chain, purple lines represent C region, and the red lines are representing D region.

Figure 2. TRAV families

Phylogenetic tree illustrating (A) functional genes (black), pseudogenes (red) and ORFs (blue) of the macfas TRAV locus. The genes clustered together belong to the same family. (B) Comparison of the TRAV/TRDV locus of homsap, macfas, and macmul. The genes that are exclusive to humans are highlighted in purple. Those TRAV genes found in macfas and macmul but not in homsap are in yellow, and the genes are present only in macmul but absent in macmul are in red. See text for the details.

Figure 3. TRBV families

Phylogenetic tree illustrating functional genes (black), pseudogenes (red) and ORFs (blue) of the macfas TRBV locus. The genes clustered together belongs to the same family. #, TRBV6-4 is a pseudogene in the genomic sequence, but the expressed gene is functional (see results, “Expressed V gene repertoire”).

Figure 4. TRGV and TRGJ gene segment homologies

(A) Phylogenetic tree illustrating functional genes (black), pseudogenes (red) and ORFs (blue) of the macfas TRGV locus. The number (i.e., "n=1") is the number of mismatches between the macfas and macmul genes. The % is the homology between the macfas and the homsap gene. Homologies between other genes of interest are indicated with a dotted line. *, homsap genes for which no macfas or macmul homologs were identified. ‡, nomenclature discrepancy. (B) Phylogenetic tree clustering macfas and macmul TRGJ genes. (C) Schematic of the genomic organization of the 3' region of the TRG locus. TRGV (red), TRGJ (blue) and TRGC (black).

Figure 5. TRDV and TRDJ gene segment homologies

(A) Phylogenetic tree showing the functional genes homsap, macfas and macmul TRDV genes. Comparisons are indicated with dotted lines and the percent homology is indicated followed by the number of sequence mismatches. Each TRDV gene family is color coded. (B) Alignment of macmul and macfas TRDJ showing the conserved amino acids (boxed).

Figure 6. The expressed TRAV and TRBV repertoire. Single cell analysis of lung mononuclear cells from Cynomolgus macaques reveals their functionally expressed TRAV and TRBV repertoire. Each dot represents a different animal. All samples are from uninvolved lung tissue from subjects infected with *M. tuberculosis*. The average percentage was calculated for the TRAV (A) and TRBV (B) and the distribution was individually normalized for each subject. Red bar represents the median. *, expression not analyzed.

427 **Supplemental information**

428

429 **Figure S1.** Constant region homology. Alignment of the amino acid sequence of the TCR
430 constant regions, derived from the in silico splicing of the human, macfas and macmul TRAC,
431 TRBC, TRGC, and TRDC exons. Dots represent identical homology. Amino acids are
432 represented by the 1-letter code. X, is undetermined.

433

434 **Figure S2.** TRBJ gene segment homology. Alignment of the nucleic acid sequences of the
435 human, macfas and macmul TRBJ genes. Dots represent identical homology.

436

437 **Figure S3.** Single cell analysis of lung mononuclear cells from Cynomolgus macaques. (A)
438 TRAV and (B) TRBV repertoire of T cells from bronchoalveolar lavage fluid, before or after
439 infection (BAL (PRE), BAL (POST)), uninvolved or involved lung tissue (UI and granuloma,
440 respectively), or single cell suspensions (SC). The mean value of the average percentage of the
441 UMI counts from each sample was plotted. Each point represents a different sample. *,
442 expression not analyzed.

443

444 **Table S1**: Macfas TRAV functional status, nucleotide sequence and the translated sequence

445

446 **Table S2.** Macfas TRAJ sequence.

447

448 **Table S3**: Macfas TRBV functional status, nucleotide sequence and the translated sequence.

449

450 **Table S4**: Macfas TRDV and TRDJ functional status, nucleotide sequence and the translated
451 sequence.

452

453 **Table S5:** Macfas TRGV/TRGJ functional status, nucleotide sequence and the translated

454 sequence.

455

References

1. Urano E, Okamura T, Ono C, Ueno S, Nagata S, Kamada H, Higuchi M, Furukawa M, Kamitani W, Matsuura Y *et al*: **COVID-19 cynomolgus macaque model reflecting human COVID-19 pathological conditions**. *Proc Natl Acad Sci U S A* 2021, **118**(43).

2. Shiina T, Suzuki S, Congy-Jolivet N, Aarnink A, Garchon HJ, Dereuddre-Bosquet N, Vaslin B, Tchitchek N, Desjardins D, Autran B *et al*: **Cynomolgus macaque IL37 polymorphism and control of SIV infection**. *Sci Rep* 2019, **9**(1):7981.

3. Rodgers MA, Ameel C, Ellis-Connell AL, Balgeman AJ, Maiello P, Barry GL, Friedrich TC, Klein E, O'Connor SL, Scanga CA: **Preexisting Simian Immunodeficiency Virus Infection Increases Susceptibility to Tuberculosis in Mauritian Cynomolgus Macaques**. *Infect Immun* 2018, **86**(12).

4. Mohns MS, Greene JM, Cain BT, Pham NH, Gostick E, Price DA, O'Connor DH: **Expansion of Simian Immunodeficiency Virus (SIV)-Specific CD8 T Cell Lines from SIV-Naive Mauritian Cynomolgus Macaques for Adoptive Transfer**. *J Virol* 2015, **89**(19):9748-9757.

5. Salguero FJ, White AD, Slack GS, Fotheringham SA, Bewley KR, Gooch KE, Longet S, Humphries HE, Watson RJ, Hunter L *et al*: **Comparison of rhesus and cynomolgus macaques as an infection model for COVID-19**. *Nat Commun* 2021, **12**(1):1260.

6. Greenaway HY, Ng B, Price DA, Douek DC, Davenport MP, Venturi V: **NKT and MAIT invariant TCRα sequences can be produced efficiently by VJ gene recombination**. *Immunobiology* 2013, **218**(2):213-224.

478 7. Carlsson HE, Schapiro SJ, Farah I, Hau J: **Use of primates in research: a global**
479 **overview**. *Am J Primatol* 2004, **63**(4):225-237.

480 8. Ebeling M, Kung E, See A, Broger C, Steiner G, Berrera M, Heckel T, Iniguez L, Albert
481 T, Schmucki R *et al*: **Genome-based analysis of the nonhuman primate Macaca**
482 **fascicularis as a model for drug safety assessment**. *Genome Res* 2011,
483 **21**(10):1746-1756.

484 9. Van Rompay KKA: **Tackling HIV and AIDS: contributions by non-human primate**
485 **models**. *Lab Anim (NY)* 2017, **46**(6):259-270.

486 10. Matz-Rensing K, Hartmann T, Wendel GM, Frick JS, Homolka S, Richter E, Munk MH,
487 Kaup FJ: **Outbreak of Tuberculosis in a Colony of Rhesus Monkeys (Macaca**
488 **mulatta) after Possible Indirect Contact with a Human TB Patient**. *J Comp Pathol*
489 2015, **153**(2-3):81-91.

490 11. Sapolsky RM, Share LJ: **A Pacific Culture among Wild Baboons: Its Emergence and**
491 **Transmission**. *PLOS Biology* 2004, **2**(4):e106.

492 12. Cadena AM, Hopkins FF, Maiello P, Carey AF, Wong EA, Martin CJ, Gideon HP,
493 DiFazio RM, Andersen P, Lin PL *et al*: **Concurrent infection with Mycobacterium**
494 **tuberculosis confers robust protection against secondary infection in macaques**.
495 *PLoS Pathog* 2018, **14**(10):e1007305.

496 13. Maiello P, DiFazio RM, Cadena AM, Rodgers MA, Lin PL, Scanga CA, Flynn JL:
497 **Rhesus Macaques Are More Susceptible to Progressive Tuberculosis than**
498 **Cynomolgus Macaques: a Quantitative Comparison**. *Infect Immun* 2018, **86**(2).

14. Flynn JL, Gideon HP, Mattila JT, Lin PL: **Immunology studies in non-human primate models of tuberculosis.** *Immunol Rev* 2015, **264**(1):60-73.

15. Lin PL, Ford CB, Coleman MT, Myers AJ, Gawande R, Ioerger T, Sacchettini J, Fortune SM, Flynn JL: **Sterilization of granulomas is common in active and latent tuberculosis despite within-host variability in bacterial killing**. *Nat Med* 2014, **20**(1):75-79.

16. Scanga CA, Flynn JL: **Modeling tuberculosis in nonhuman primates**. *Cold Spring Harb Perspect Med* 2014, **4**(12):a018564.

17. Qi Q, Liu Y, Cheng Y, Glanville J, Zhang D, Lee JY, Olshen RA, Weyand CM, Boyd SD, Goronzy JJ: **Diversity and clonal selection in the human T-cell repertoire**. *Proc Natl Acad Sci U S A* 2014, **111**(36):13139-13144.

18. Gouaillard C, Huchenq-Champagne A, Arnaud J, Chen Cl CL, Rubin B: **Evolution of T cell receptor (TCR) alpha beta heterodimer assembly with the CD3 complex**. *Eur J Immunol* 2001, **31**(12):3798-3805.

19. Murphy K, Weaver C: **Janeway's Immunobiology**; 2017.

20. Kim DS, Lee KY, Yang WI, Han SJ, Hwang EH: **Gamma/delta T lymphocytes in the BCG granulomatous lesions**. *Yonsei Med J* 1996, **37**(5):319-324.

21. Zhao Y, Niu C, Cui J: **Gamma-delta (gammadelta) T cells: friend or foe in cancer development?** *J Transl Med* 2018, **16**(1):3.

22. Lefranc MP, Lefranc G: **The T Cell Receptor FactsBook**. London: Academic Press; 2001.

520 23. Lefranc MP, Chuchana P, Dariavach P, Nguyen C, Huck S, Brockly F, Jordan B, Lefranc
521 G: **Molecular mapping of the human T cell receptor gamma (TRG) genes and**
522 **linkage of the variable and constant regions**. *Eur J Immunol* 1989, **19**(6):989-994.

523 24. Huang H, Wang C, Rubelt F, Scriba TJ, Davis MM: **Analyzing the Mycobacterium**
524 **tuberculosis immune response by T-cell receptor clustering with GLIPH2 and**
525 **genome-wide antigen screening**. *Nat Biotechnol* 2020, **38**(10):1194-1202.

526 25. Glanville J, Huang H, Nau A, Hatton O, Wagar LE, Rubelt F, Ji X, Han A, Krams SM,
527 Pettus C *et al*: **Identifying specificity groups in the T cell receptor repertoire**. *Nature*
528 2017.

529 26. Munson DJ, Egelston CA, Chiotti KE, Parra ZE, Bruno TC, Moore BL, Nakano TA,
530 Simons DL, Jimenez G, Yim JH *et al*: **Identification of shared TCR sequences from T**
531 **cells in human breast cancer using emulsion RT-PCR**. *Proc Natl Acad Sci U S A*
532 2016, **113**(29):8272-8277.

533 27. Feldman SA, Assadipour Y, Kriley I, Goff SL, Rosenberg SA: **Adoptive Cell Therapy--**
534 **Tumor-Infiltrating Lymphocytes, T-Cell Receptors, and Chimeric Antigen**
535 **Receptors**. *Semin Oncol* 2015, **42**(4):626-639.

536 28. Tang L, Zheng Y, Melo MB, Mabardi L, Castano AP, Xie YQ, Li N, Kudchodkar SB,
537 Wong HC, Jeng EK *et al*: **Enhancing T cell therapy through TCR-signaling-**
538 **responsive nanoparticle drug delivery**. *Nat Biotechnol* 2018, **36**(8):707-716.

539 29. Ohno S: **Evolution by Gene Duplication**. Heidelberg, Germany: Springer-Verlag; 1970.

540 30. Lynch M, Conery JS: **The evolutionary fate and consequences of duplicate genes**.
541 *Science* 2000, **290**(5494):1151-1155.

542 31. Demuth JP, De Bie T, Stajich JE, Cristianini N, Hahn MW: **The evolution of**
543 **mammalian gene families**. *PLoS One* 2006, **1**:e85.

544 32. Olson MV: **When less is more: gene loss as an engine of evolutionary change**. *Am*
545 *J Hum Genet* 1999, **64**(1):18-23.

546 33. Yan G, Zhang G, Fang X, Zhang Y, Li C, Ling F, Cooper DN, Li Q, Li Y, van Gool AJ *et*
547 *al*: **Genome sequencing and comparison of two nonhuman primate animal models,**
548 **the cynomolgus and Chinese rhesus macaques**. *Nat Biotechnol* 2011, **29**(11):1019-
549 1023.

550 34. Schoch CL, Ciufo S, Domrachev M, Hotton CL, Kannan S, Khovanskaya R, Leipe D,
551 McVeigh R, O'Neill K, Robbertse B *et al*: **NCBI Taxonomy: a comprehensive update**
552 **on curation, resources and tools**. *Database (Oxford)* 2020, **2020**.

553 35. Greenaway HY, Kurniawan M, Price DA, Douek DC, Davenport MP, Venturi V:
554 **Extraction and characterization of the rhesus macaque T-cell receptor β-**
555 **chain genes**. *Immunol Cell Biol* 2009, **87**(7):546-553.

556 36. Lefranc MP, Giudicelli V, Duroux P, Jabado-Michaloud J, Folch G, Aouinti S, Carillon E,
557 Duvergey H, Houles A, Paysan-Lafosse T *et al*: **IMGT(R), the international**
558 **ImMunoGeneTics information system(R) 25 years on**. *Nucleic Acids Res* 2015,
559 **43**(Database issue):D413-422.

560 37. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN,
561 Potter SC, Finn RD *et al*: **The EMBL-EBI search and sequence analysis tools APIs in**
562 **2019**. *Nucleic Acids Res* 2019, **47**(W1):W636-W641.

38. Han MV, Zmasek CM: **phyloXML: XML for evolutionary biology and comparative genomics**. *BMC Bioinformatics* 2009, **10**:356.

39. Gideon HP, Hughes TK, Wadsworth MH, 2nd, Tu AA, Gierahn TM, Hopkins FH, Wei J-R, Kummerlowe C, Grant NL, Nargan K *et al*: **Multimodal profiling of lung granulomas reveals cellular correlates of tuberculosis control**. 2020. https://www.biorxiv.org/content/10.1101/2020.10.24.352492v1

40. Tu AA, Gierahn TM, Monian B, Morgan DM, Mehta NK, Ruiter B, Shreffler WG, Shalek AK, Love JC: **TCR sequencing paired with massively parallel 3' RNA-seq reveals clonotypic T cell signatures**. *Nat Immunol* 2019, **20**(12):1692-1699.

41. Jivanjee T, Ibrahim S, Nyquist SK, Gatter GJ, Bromley JD, Jaiswal S, Berger B, Behar SM, Love JC, Shalek AK: **Enriching and Characterizing T-Cell Repertoires from 3' Barcoded Single-Cell Whole Transcriptome Amplification Products**. 2022. https://doi.org/10.48550/arXiv.2203.11266

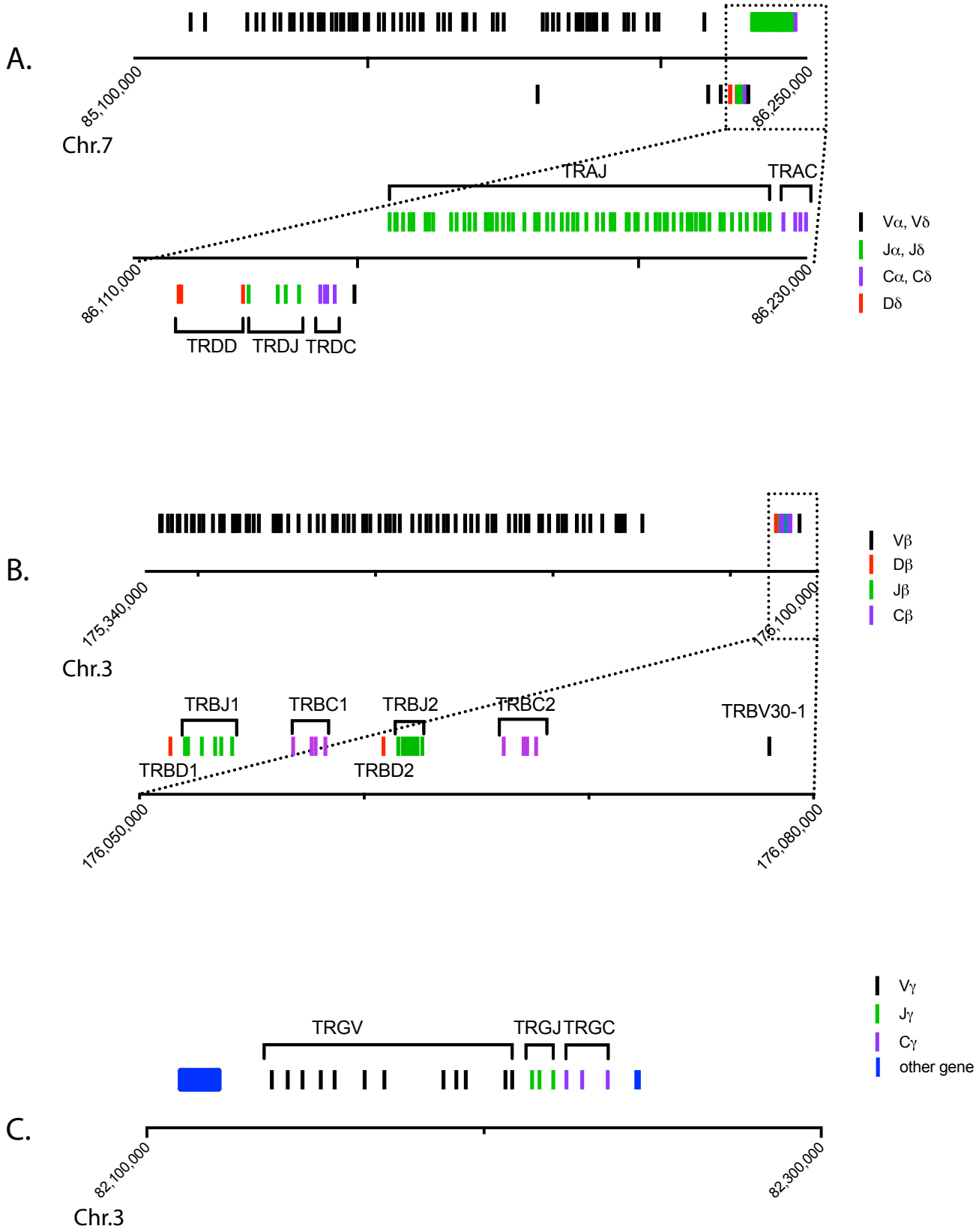
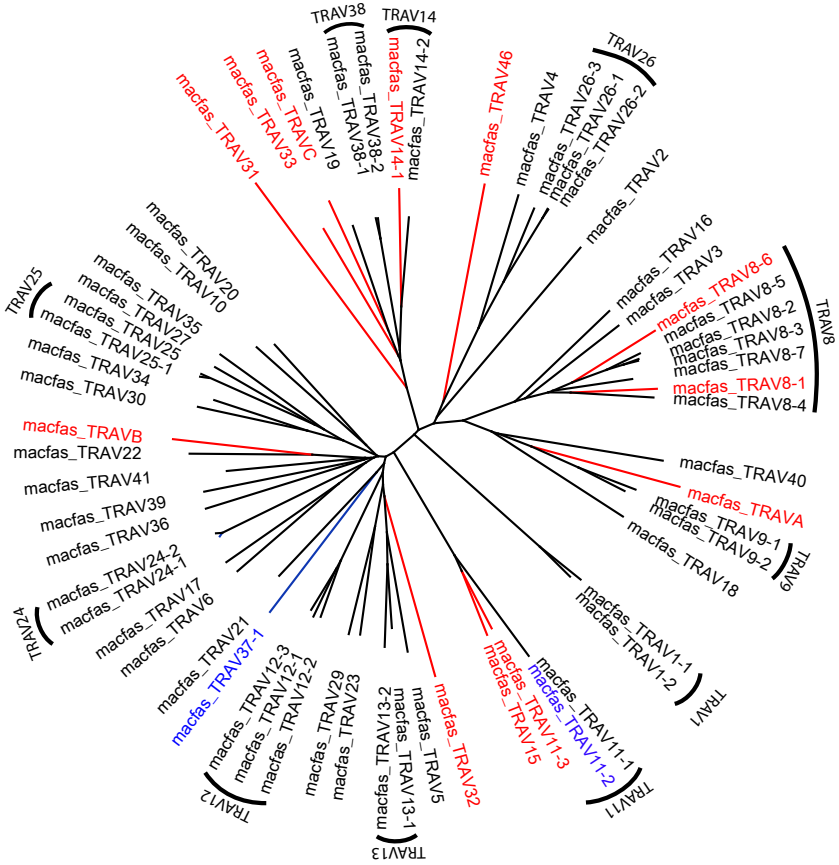


Figure 1

A.


TRAV38
TRAV14
macfas_TRAV38-2
macfas_TRAV38-1
macfas_TRAV14-2
macfas_TRAV14-1
macfas_TRAV31
macfas_TRAVC
macfas_TRAV33
macfas_TRAV19
macfas_TRAV46
macfas_TRAV4
TRAV26
macfas_TRAV26-3
macfas_TRAV26-1
macfas_TRAV26-2
macfas_TRAV2
macfas_TRAV16
macfas_TRAV3
macfas_TRAV8-6
macfas_TRAV8-5
macfas_TRAV8-2
macfas_TRAV8-3
macfas_TRAV8-7
macfas_TRAV8-1
macfas_TRAV8-4
TRAV8
macfas_TRAV40
macfas_TRAVA
macfas_TRAV9-1
macfas_TRAV9-2
TRAV9
macfas_TRAV18
macfas_TRAV1-1
macfas_TRAV1-2
TRAV1
macfas_TRAV11-1
TRAV11
macfas_TRAV11-2
macfas_TRAV11-3
macfas_TRAV15
macfas_TRAV32
TRAV13
macfas_TRAV13-1
macfas_TRAV13-2
macfas_TRAV5
macfas_TRAV29
macfas_TRAV23
TRAV12
macfas_TRAV12-3
macfas_TRAV12-1
macfas_TRAV12-2
macfas_TRAV21
macfas_TRAV37-1
macfas_TRAV17
macfas_TRAV6
TRAV24
macfas_TRAV24-2
macfas_TRAV24-1
macfas_TRAVB
macfas_TRAV22
macfas_TRAV41
macfas_TRAV39
macfas_TRAV36
TRAV25
macfas_TRAV20
macfas_TRAV10
macfas_TRAV35
macfas_TRAV27
macfas_TRAV25
macfas_TRAV25-1
macfas_TRAV34
macfas_TRAV30

B.

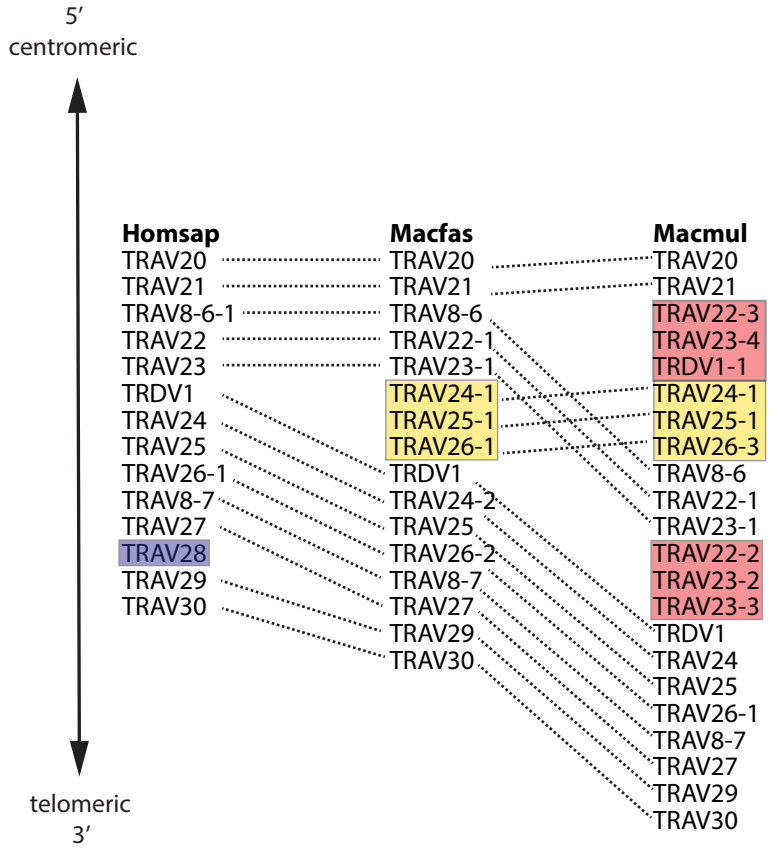

5′
centromeric
telomeric
3′
Homsap
TRAV20
TRAV21
TRAV8-6-1
TRAV22
TRAV23
TRDV1
TRAV24
TRAV25
TRAV26-1
TRAV8-7
TRAV27
TRAV28
TRAV29
TRAV30
Macfas
TRAV20
TRAV21
TRAV8-6
TRAV22-1
TRAV23-1
TRAV24-1
TRAV25-1
TRAV26-1
TRDV1
TRAV24-2
TRAV25
TRAV26-2
TRAV8-7
TRAV27
TRAV29
TRAV30
Macmul
TRAV20
TRAV21
TRAV22-3
TRAV23-4
TRDV1-1
TRAV24-1
TRAV25-1
TRAV26-3
TRAV8-6
TRAV22-1
TRAV23-1
TRAV22-2
TRAV23-2
TRAV23-3
TRDV1
TRAV24
TRAV25
TRAV26-1
TRAV8-7
TRAV27
TRAV29
TRAV30

Figure 2

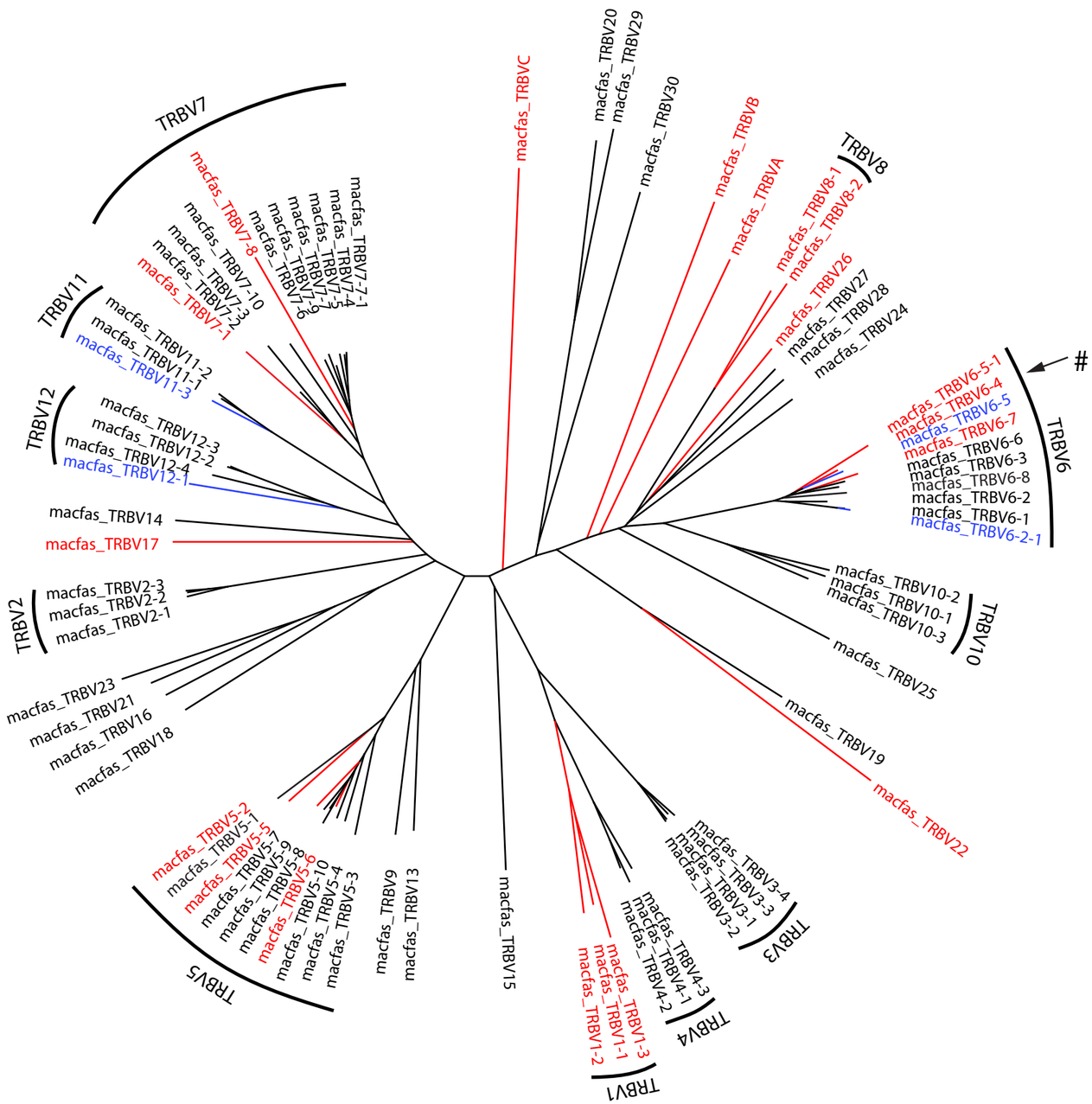


Figure 3

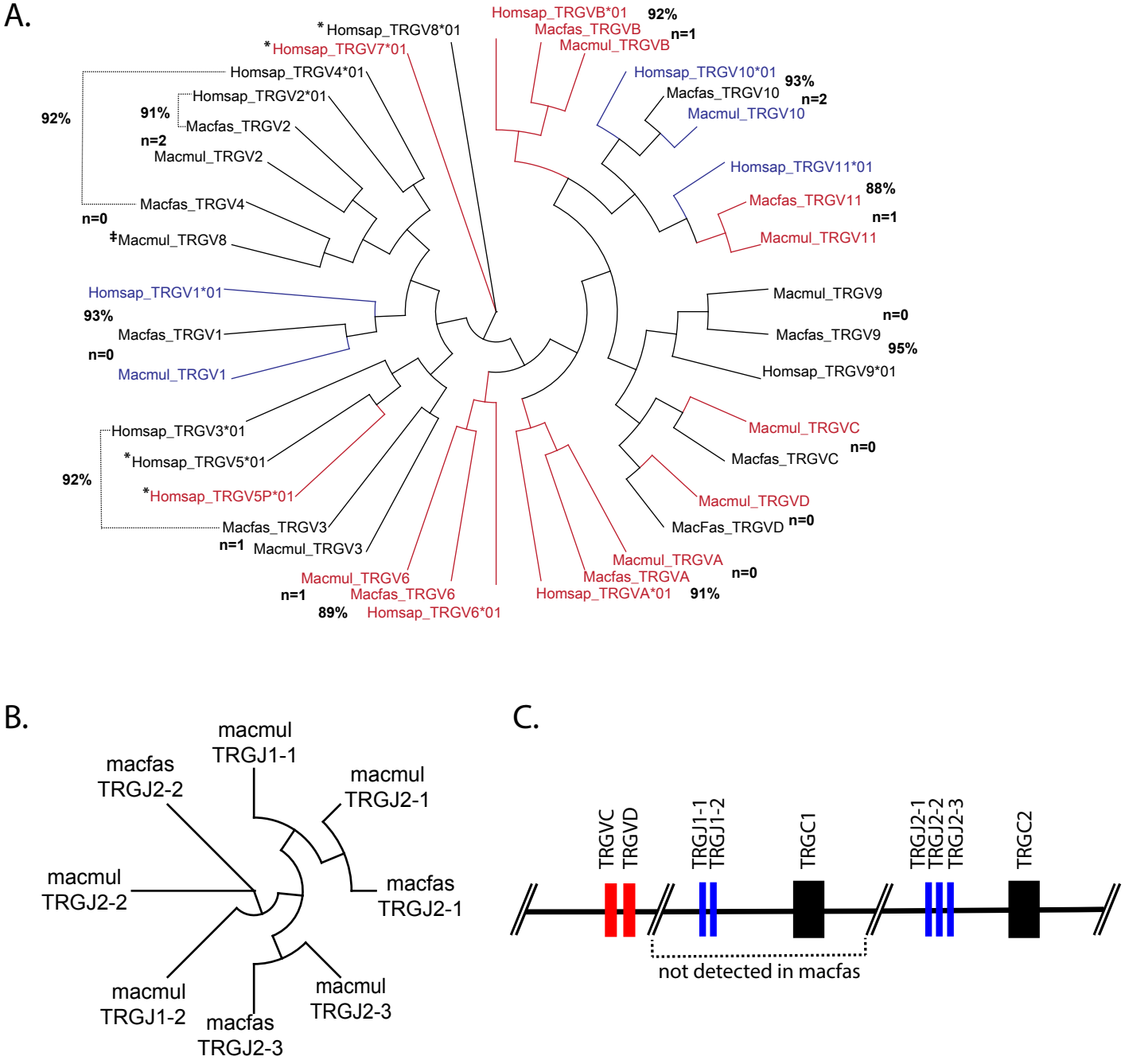
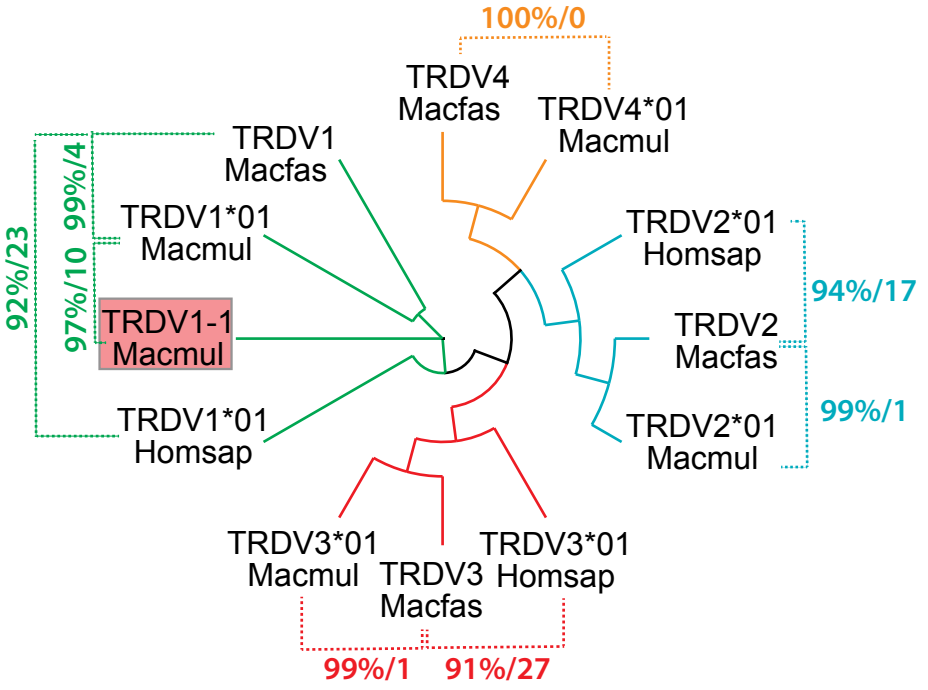


Figure 4

A.



B.

```
                                                  PheGly     Gly                 ValGlu
macfas TRDJ1 --------ACACTGATAAACTCATCTTTGGAAAAGGAACCCGTGTGACTGTGGAACCAA
macmul TRDJ1 --------.....................................................
macfas TRDJ2 -----CTTT...A.CAC.....T.........G.....T.AAC.C.TC.........G
macmul TRDJ2 -----CTTT...A.CAC.....T.........G.....T.AAC.C.TC.........G
macfas TRDJ3 CTCCTGGG....CCGAC.GA.GT.T......CT..C...AAAC.CTTC.....G..CC
macmul TRDJ3 CTCCTGGG....CCGAC.GA.GT.T......CT..C...AAAC.CTTC.....G..CC
macfas TRDJ4 -----------.CAGACC...G.....C.........TA.C..GAG..A...AG.C
macmul TRDJ4 -----------.CAGACC...G.....C.........TA.C..GAG..A...AG.C
```

Figure 5

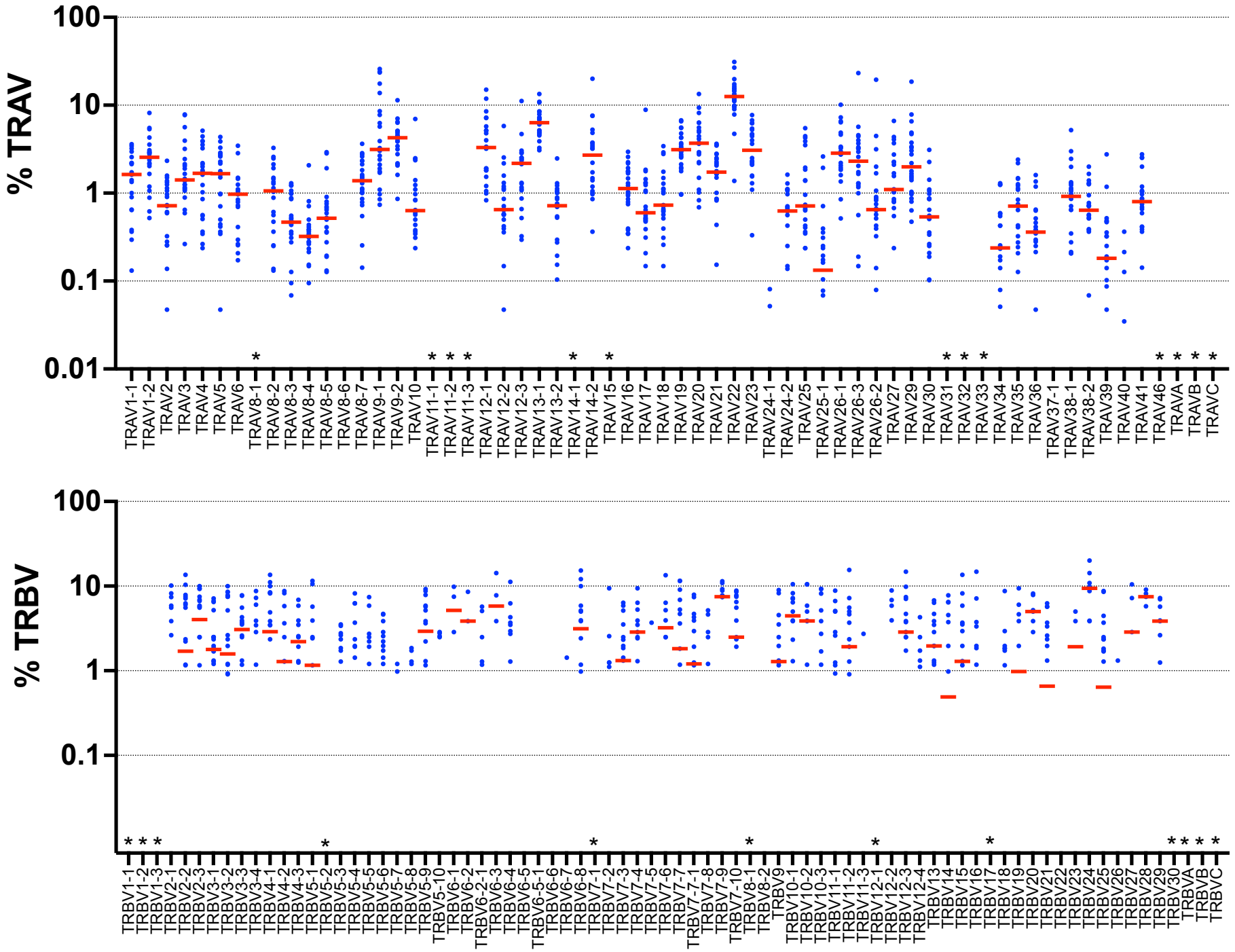

% TRAV
100
10
1
0.1
0.01
TRAV1-1
TRAV1-2
TRAV2
TRAV3
TRAV4
TRAV5
TRAV6
TRAV8-1
TRAV8-2
TRAV8-3
TRAV8-4
TRAV8-5
TRAV8-6
TRAV8-7
TRAV9-1
TRAV9-2
TRAV10
TRAV11-1
TRAV11-2
TRAV11-3
TRAV12-1
TRAV12-2
TRAV12-3
TRAV13-1
TRAV13-2
TRAV14-1
TRAV14-2
TRAV15
TRAV16
TRAV17
TRAV18
TRAV19
TRAV20
TRAV21
TRAV22
TRAV23
TRAV24-1
TRAV24-2
TRAV25
TRAV25-1
TRAV26-1
TRAV26-3
TRAV26-2
TRAV27
TRAV29
TRAV30
TRAV31
TRAV32
TRAV33
TRAV34
TRAV35
TRAV36
TRAV37-1
TRAV38-1
TRAV38-2
TRAV39
TRAV40
TRAV41
TRAV46
TRAVA
TRAVB
TRAVC
% TRBV
100
10
1
0.1
TRBV1-1
TRBV1-2
TRBV1-3
TRBV2-1
TRBV2-2
TRBV2-3
TRBV3-1
TRBV3-2
TRBV3-3
TRBV3-4
TRBV4-1
TRBV4-2
TRBV4-3
TRBV5-1
TRBV5-2
TRBV5-3
TRBV5-4
TRBV5-5
TRBV5-6
TRBV5-7
TRBV5-8
TRBV5-9
TRBV5-10
TRBV6-1
TRBV6-2
TRBV6-2-1
TRBV6-3
TRBV6-4
TRBV6-5
TRBV6-5-1
TRBV6-6
TRBV6-7
TRBV6-8
TRBV7-1
TRBV7-2
TRBV7-3
TRBV7-4
TRBV7-5
TRBV7-6
TRBV7-7
TRBV7-7-1
TRBV7-8
TRBV7-9
TRBV7-10
TRBV8-1
TRBV8-2
TRBV9
TRBV10-1
TRBV10-2
TRBV10-3
TRBV11-1
TRBV11-2
TRBV11-3
TRBV12-1
TRBV12-2
TRBV12-3
TRBV12-4
TRBV13
TRBV14
TRBV15
TRBV16
TRBV17
TRBV18
TRBV19
TRBV20
TRBV21
TRBV22
TRBV23
TRBV24
TRBV25
TRBV26
TRBV27
TRBV28
TRBV29
TRBV30
TRBVA
TRBVB
TRBVC

Figure 6