Immunoglobulin heavy constant gamma gene evolution is modulated by both the divergent and birth-and-death evolutionary models

Diego Garzón-Ospina^{1,2,3*}

e-mail: degazon@gmail.com

ORCID: 0000-0003-3829-6719

Sindy P. Buitrago^{1,2,3*}

e-mail: biopao.k18@gmail.com

ORCID: 0000-0002-9872-5661

*Corresponding author

e-mail: degazon@gmail.com; biopao.k18@gmail.com

¹: PGAME - Population Genetics And Molecular Evolution, Fundación Scient, Tunja, Boyacá, Colombia.

²: GEO, School of Biological Sciences, Universidad Pedagógica y Tecnológica de Colombia - UPTC, Tunja, Boyacá, Colombia.

³: GEBIMOL, School of Biological Sciences, Universidad Pedagógica y Tecnológica de Colombia - UPTC, Tunja, Boyacá, Colombia.

Abstract

Immunoglobulin G (IgG) is one of the five antibody classes produced in mammals as part of the humoral responses. This high-affinity antibody produced late in a primary immune response is responsible for protecting the organisms from infection. This protein’s heavy chain constant region is encoded by the Ig gamma gene (*Igγ*). In Humans, IgG has evolved into four subclasses with specialized effector functions. However, in *Platyrrhini*, IgG has been reported to be encoded by a single-copy gene. Here, we analyzed data from 38 primate genome sequences to identify *Igγ* genes and describe the evolution of this immunoglobulin in this group. *Igγ* belongs to a multigene family that evolves by the birth-death evolutionary model in primates. Whereas *Strepsirrhini* and *Platyrrhini* have a single-copy gene, in *Catarrhini* species it has expanded to having several paralogs in their genomes; some have been deleted or have become pseudogenes. Furthermore, episodic positive selection might promote a specialized species-specific IgG effector function. A hypothesis for the *Igγ* evolution has been proposed, suggesting that IgG has evolved to reach an optimal number of copies per genome to adapt their humoral immune responses to different environmental conditions.

Keywords: IgG, Primates, Immune system, Multigene family, Birth-death evolutionary model, Divergent evolutionary model

Graphical abstract

1. Introduction

Antibodies or immunoglobulins (Ig) are the primary molecules mediating humoral immunity. These molecules are found in all gnathostomes (jawed vertebrates), evolving to block or neutralize pathogens (bacteria, viruses, and parasites). They are formed by four polypeptide chains, two immunoglobulin heavy (H) chains and two light (L). The N-terminal or Fab (Fragment antigen-binding) region is extremely diverse and responsible for recognizing and interacting with the antigen, while the C-terminal or Fc (fragment crystallizable region, the “antibody tail”) region is responsible for triggering the effector cell responses. In mammals, from the most primitive (egg-laying mammals) to primates, there are five genes: *mu* (μ), *delta* (δ), *gamma* (γ), *epsilon* (ϵ), and *alpha* (α), each encoding a different Ig class (IgM, IgD, IgG, IgE, and IgA, respectively) (Senger et al., 2015). They carry out diverse effector functions, such as complement binding, phagocytic cell binding, opsonization, and transport through the placenta epithelium (Schroeder and Cavacini, 2010).

IgG plays an essential role in mammalian adaptive immunity. Unlike other Ig, class G is involved in activating the decisive effector responses that allow antigen/pathogen neutralization or elimination. Moreover, this is the most abundant Ig in the serum, with a half-life of 23 days (Birdsall, 2015; Vidarsson et al., 2014) and it is the only isotype transmitted through the human placenta to the fetus through the placental transport receptor (FcRn). There are four IgG subclasses in humans, each with unique features and functions, although they share more than 90% amino acid sequence similarity (Nimmerjahn and Ravetch, 2005; Vidarsson et al., 2014). For instance, IgG1 and IgG3 can efficiently trigger the classical complement pathway; however, IgG2 and IgG4 do it less efficiently. IgG1 and IgG3 are also the predominant subclasses involved in responding to protein antigens, while IgG2 participates in responding to polysaccharide antigens. Placental transport is also higher for IgG1 and IgG3 but lower for IgG2 and IgG4. IgG3 has a powerful pro-inflammatory effect, given that it is the first IgG to appear after viral infections, followed by IgG1. On the other hand, higher levels of IgG4 have been associated with asymptomatic infections caused by helminths and filariasis (Mishra et al., 2019; Napodano et al., 2020; Vidarsson et al., 2014). Likewise, IgG abundance in serum varies among subclasses; IgG1 represents up to 70% of the total IgG, followed by IgG2 (up to 30%), and IgG3 (~8%); IgG4 is the least frequent subclass (less than 5%) (Schur, 1987).

These IgG subclasses are encoded by paralogous γ genes resulting from tandem duplication events which has been differentially expanded in mammal species (Esteves and Binaghi, 1972; Sun et al.,

2012); Platypus, a primitive mammal, shows two γ paralogs genes. Cattle displays three while humans, mice, and rats show four copies. At least six distinct γ duplicates have been reported in pigs. To date, the horse is the mammal with the largest copy-number (7 genes) (Sun et al., 2012). The number of γ genes also appears to be different in primates. Macaques have four paralogs (Jacobsen et al., 2011; Ramesh et al., 2017) while in humans there are five (Takahashi et al., 1982), in contrast to new world monkeys where IgG is encoded by a single-copy gene (Brusco et al., 1998; Garzon-Ospina and Buitrago, 2020; Olivieri and Gambon Deza, 2018). Furthermore, in spite macaques and humans have four functional paralogs there does not appear to be a structural and functional correlation between subclasses (Crowley and Ackerman, 2019; Mishra et al., 2019; Nguyen et al., 2014; Ramesh et al., 2017; Warncke et al., 2012). This suggests that *Ig γ* multigene family evolutionary history could differ among primates.

Multigene families can evolve by divergent evolution (duplicates diverged gradually acquiring new functions), concerted evolution (duplicates evolve in a concerted manner rather than independently), or by the birth-and-death model (duplicates may be maintained for a long time, whereas others are deleted or become nonfunctional) (Nei and Rooney, 2005). To the best of our knowledge it is not clear how *Ig γ* has expanded or contracted in the primate order. With the purpose of assessing the evolutionary history of this gene in these mammals, we screened 38 primates genomes available in the GenBank database and performed phylogenetic and molecular evolutionary analyses. This data can help understand the differences among the primates' immune system and its evolution to cope with infections. Likewise, it could have implications for biomedical trials using these animals.

2. Materials and Methods

2.1. Datasets and *Ig γ* gene identification

The genome sequences from 38 primate species (Supplementary material 1) available in GenBank (Benson et al., 2015) were analyzed to identify the DNA regions encoding the immunoglobulin heavy constant gamma gene (*Ig γ*). The genomic region containing the constant *Ig γ* loci is circumscribed by *Ig δ* and *Ig ϵ* heavy chain genes. Therefore, for those species with a complete ensemble genome, the chromosome region encompassing both *Ig δ* and *Ig ϵ* genes were analyzed to identify putative *Ig γ* loci. For the remaining species, a Blast search was performed to locate DNA contigs containing *Ig γ* genes. To determine the exon-intron gene structure, the putative *Ig γ* genes were aligned with human IgG CDS subclasses (GenBank accession numbers: AJ294730.1, AJ294731.1, AK097307.1, and

AJ294733.1) using the MUSCLE method (Edgar, 2004). Alignments were manually edited in GeneDoc software (K. and H., 1997), identifying the donor and acceptor-splice sites (GT...AG). The coding DNA sequences found were used to deduce amino acid sequences using the GeneRunner software. These sequences were then screened to distinguish immunoglobulin domains through the Pfam webserver (El-Gebali et al., 2019).

2.2. The primate *Igγ* family's phylogenetic relationships

Igγ gene sequences and their deduced-amino-acid-sequences were aligned using the MUSCLE method. The best DNA and protein substitution models were selected using Akaike information criterion through the JModelTest (Posada, 2008) and ProtTest (Darriba et al., 2011) methods, respectively. The GTR+I+G and JTT+G+F models were used to infer phylogenetic DNA and amino acid trees, respectively, using the Bayesian method (BY). A Metropolis-coupled Markov Chain Monte Carlo (MCMC) algorithm was used for this analysis using MrBayes (Huelsenbeck et al., 2001). It was run for 10 million generations. Sump and Sumt commands were used for tabulating posterior probabilities and building a consensus tree. Both analyses were carried out using the CIPRES Science Gateway (Miller MA, 2010; Miller et al., 2015).

Alternatively, we used the multigene family model called DLTRS (Duplications, losses, transfers, rates and sequence evolution) (Sjostrand et al., 2014). This method infers a gene tree by evolving down on a given species tree through duplication, loss, and transfer events according to a birth-death-like process (Sjostrand et al., 2012; Sjostrand et al., 2014). The species tree was inferred with a concatenate fragment of cytochrome C oxidase subunit I and MT-CO2 pseudogene 9. DLTRS was then run using the MCMC algorithm for 10 million generations.

Once the introns were removed, a multiple DNA alignment for *Igγ* was carried out considering the amino acid sequence information using the MUSCLE method in the TranslatorX webserver (Abascal et al., 2010). This alignment was then used to assess natural selection signatures using Datamonkey (Weaver et al., 2018). The Branch-site Unrestricted Statistical Test for Episodic Diversification (BUSTED) method (Murrell et al., 2015) was used to identify whether the *Igγ* gene had experienced positive selection. This gene-wide test assesses positive selection at least in one site on at least one branch. Then, the adaptive Branch-Site Random Effects Likelihood (aBSREL) method (Smith et al., 2015) was performed to identify those *Igγ* branches (lineages) under episodic diversifying selection. In addition, natural selection signatures at individual codon sites were detected by using codon-based

tests (FUBAR (Murrell et al., 2013), MEME (Murrell et al., 2012), and SLAC (Kosakovsky Pond and Frost, 2005)). These tests infer the nonsynonymous (d_N or K_N) and synonymous (d_S or K_S) substitution rates for each site basis for a given coding alignment and corresponding phylogeny (Weaver et al., 2018). A sliding window for omega rates ($\omega = K_N/K_S$) was draw using the SLAC data to assess the effect of natural selection throughout the *Ig γ* gene sequence. Because recombination can bias K_N/K_S estimation (Anisimova et al., 2003; Arenas M, 2014; Arenas and Posada, 2010), the GARD method (Kosakovsky Pond et al., 2006) was performed before running codon-based tests. Finally, the RELAX method (Wertheim et al., 2015) was performed. It allows partitioning a codon-based phylogeny into two subsets of branches to assess whether selective strength was relaxed or intensified in one of these subsets (test branch) relative to the other (reference branch).

2.3. Estimation of *Ig γ* paralogs divergence times

The 4-fold degenerate sites from the *Ig γ* coding alignment were used to estimate the divergence times among *Ig γ* paralogs using the Bayesian MCMC approach available in BEAST software (Drummond and Rambaut, 2007; Suchard et al., 2018). BEAUti (Drummond and Rambaut, 2007; Suchard et al., 2018) was used to generate a BEAST XML file. Divergence times were estimated for individual species using the evolutionary model GTR+G with the relaxed clock option and simulating a Yule process. Four calibration dates (*Cebus albifrons*/*Cebus capucinus*: 1.89 +/- 0.1; *Eulemur macaco*/*Eulemur fulvus*: 7.66 +/- 0.14; *Pithecia pithecia*/*Plecturocebus donacophilus*: 18 +/- 0.6 and *Indri indri*/*Propithecus coquereli*: 29 +/- 6.6) were used based on Timetree web server data (<http://www.timetree.org/> (Kumar et al., 2017)). The MCMC length chain was run for 50 million generations. Tracer v1.5 (Rambaut et al., 2018) was used to check when the MCMC chains had reached a stationary distribution by visual inspection of plotted posterior estimates and by checking that ESS (Effective Sample Size) was greater than 1000 for divergence times.

3. Results

3.1. *Ig γ* gene identification in primate genomes

The sequences containing the immunoglobulin heavy constant gamma locus from 38 primate species were screened to describe the number of putative genes encoding IgG in these species. *Strepsirrhini* and *Haplorhini* suborders displayed differences in the *Ig γ* copy number (Fig. 1). A single *Ig γ* gene was found in all *Lemuriformes* (*Strepsirrhini*). As mentioned previously, in most *Platyrrhini* species, IgG seems to be encoded by a single-copy gene. *Alpa* was the only species in this parvorder containing

two *Ig γ* copies (paralogs) (Garzon-Ospina and Buitrago, 2020). On the other hand, the *Catarrhini* species had a high copy-number diversity. Some *Cercopithecoidea* species showed just one *Ig γ* gene; other species had three or four duplicates; this feature also was detected in *Hominoidea*. *Gogo* was the species with the lowest *Ig γ* copy-number, displaying two paralogues, followed by *Hosa* with five, *Poab* with six, and *Patr* with eight duplicates. All *Ig γ* copies were named in ascending order according to their position in the contigs from left to right from *Ig δ* gene. The *Hosa* *Ig γ* subclasses were numbered as is typically done, *Ig γ 3*, *Ig γ 1*, *Ig γ Ψ* , *Ig γ 2*, and *Ig γ 4*, respectively.

The *Ig γ* exon-intron structure was stable in *Lemuriformes*, *Platyrrhini*, and *Cercopithecoidea* lineages displaying four exons (Fig. 1 and Supplementary material 2). Although *Caja* had been predicted as single-exon gene (Garzon-Ospina and Buitrago, 2020), a re-analysis suggested that two exons might encode IgG in this species; further cDNA analysis will be required to confirm this. According to the Pfam search, the putative IgG proteins in all these species showed the typical three Ig-domains. In *Hominoidea*, some paralogues have expanded the second exon producing a large hinge encoding region. On the other hand, the *PatrIg γ 2* gene seems to have lost the first exon (Fig. 1); therefore, this gene would be encoding a two Ig-domains protein. Finally, *PoabIg γ 3* and *PoabIg γ 4* had several premature stop codons.

3.2. The primate *Ig γ* family's phylogenetic relationships

Phylogenetic relationships for the *Ig γ* family were inferred by Bayesian (BY) and DLTRS methods using both DNA and deduced-amino-acid-sequences (Fig. 2-3 and Supplementary material 3-4). A multiple alignment for 67 *Ig γ* gene sequences was used for inferring a BY phylogenetic tree using the GTR+I+G model. This phylogeny showed posterior probabilities higher than > 0.95 for most of the nodes, thus, supporting the branching pattern. Four major clades were observed resembling species relationships. The first *Ig γ* clade included sequences from *Strepsirrhini* species. The *Platyrrhini* parvorder formed the second clade, where *Alpa* duplicates were clustered in the same branch. The third clade grouping the genes of the *Cercopithecoidea* superfamily. Two different subclades were observed in this group. The subclade 1 clustered the single gene found for *Colobinae* (*Nala*, *Trfr*, *Rhro*, and *Pyne*) and the *Ig γ 1* from *Cercopithecinae* (*Erpa*, *Thge*, *Masp*, *Ceat*, *Mane*, *Mafu*, and *Mamu*). The second subclade was formed by the *Ig γ 2*, *Ig γ 3*, and *Ig γ 4* identified in *Cercopithecinae* species. However, the *Ig γ 2* and *Ig γ 3* genes, from *Pyne* and *Ceat*, showed a species-specific clustering (Fig. 2). The latter mayor clade included paralog genes found in the *Hominoidea* superfamily (Fig. 2). Most of the *Patr* paralogs were clustered together with *Hosa* *Ig γ* genes, whereas *Poab* paralogues

were found forming a specie-specific subclade. The same topology was displayed by amino acid sequences BY phylogenetic tree inferred with the JTT+G+F model (Supplementary material 3).

In addition, the DLTRS model, which reconciles the gene tree to the species tree (Sjostrand et al., 2012; Sjostrand et al., 2014), was also performed. The trees inferred by DLTRS (for both DNA and deduced-amino-acid-sequences) also had four major clades (Fig. 3 and Supplementary material 4), showing similar branching patterns to the BY topologies. Thus, both methods displayed the same phylogenetic relationships (Fig. 2-3 and Supplementary material 3-4).

3.3. Natural selection signatures in *Igγ*

Several methods were applied to assess how natural selection has occurred in the *Igγ* family. The BUSTED method found evidence of gene-wide episodic diversifying selection in *Igγ* phylogeny (p-value = $0.000 \leq 0.05$). The aBSREL method was performed to identify those *Igγ* branches (lineages) under episodic diversifying selection; 15 branches displayed positive selection signatures (Fig. 4). Thirteen percent of the sites from the branch leading *Haplorrhini* displayed a $\omega = 8.12$. Likewise, the *Platyrrhini* lineage showed episodic selection ($\omega = 18.1$ (7.5% of the sites)). Moreover, several branches showed evidence of episodic diversifying selection within this parvorder.

Codon-based tests identified 77 sites under negative selection ($\omega < 0$) and 44 under positive selection ($\omega > 0$) (Fig. 5 and Supplementary material 5). The negative sites selected were found throughout the *Igγ* gene. Conversely, most of the positive sites selected were located on the CH1 and CH2 domains. Furthermore, the sites homolog to those involved in binding to complement and/or effector cell receptors previously reported for humans displayed a high divergence among species; some of them had signals of positive selection (Fig. 5 and Supplementary material 6).

Finally, the RELAX method was performed to determine whether selective strength was relaxed or intensified in the branch subsets. Because *Igγ* was not expanded in *Strepsirrhini*, this branch was used as the reference. Both *Platyrrhini* and *Hominoidea* (*Gogo/Hosa/Patr*) showed that intensified selection appears to modulate *Igγ* paralog evolution. In contrast, *Poab* displayed a statistically significant k value lower than 1, indicating a relaxed selection (Supplementary material 7), likely, because two of the six *Poab Igγ* were pseudogenes.

3.4. Estimation of *Igγ* paralogs divergence times

Traditionally, four-fold degenerate sites have been considered to follow the neutral expectation because they are virtually free of selective constraint. Therefore, these sites were used for paralogues divergence time estimation (Table 1 and 2). The paralogues divergence times were consistent with the split times for mayor groups. For instance, according to data on *Hosa* and *Patr* the *Igy* paralogues divergence took place around 18.45 – 20.29 MYA, which is posterior to the *Cercopithecoidea* and *Hominoidea* split (27.9 – 31.3 MYA (Kumar et al., 2017)) but before *Hominoidea* divergence (14.7 – 16.7 MYA (Kumar et al., 2017)). Likewise, *Mamu* and *Mane Igys* divergence occurred 21.3 – 22.2 MYA, which is the time estimated for the *Cercopithecoidea* split (16.4 – 22.2 MYA (Kumar et al., 2017)). It is worth noting that some species showed low divergence times between some *Igy* paralogues. Besides *Alpa*, which was the only *Platyrrhini* with a paralog gene (duplication could have taken place around 2.9 MYA), *Ceat*, *Pyne*, *Patr*, and *Poab* also had paralogues that diverged relatively recently (2 – 6 MYA) (Table 1 and 2).

4. Discussion

Gene duplication is the raw material for evolution (Lynch and Conery, 2000; Stephens, 1951; Zhang, 2003). After duplication, one of the genes (a paralog) can undergo the relaxation of selective strength, allowing the accumulation of mutations that can drive gene inactivation (pseudogenization) or result in acquiring a new function (neofunctionalization) conferring an adaptive advantage (Kondrashov and Kondrashov, 2006; Kondrashov et al., 2002; Krakauer and Nowak, 1999; Zhang, 2003). A multigene family will be formed if duplication is repeated several times in the same DNA region.

Multigene families can evolve by divergent evolution, concerted evolution, or by the birth-and-death model (Nei and Rooney, 2005). The latter appears be common in vertebrates adaptive immune system (Hughes and Nei, 1990; Nei et al., 1997; Ota and Nei, 1994; Su and Nei, 2001). In fact, the main molecules mediating humoral immune responses have emerged by gene duplication. Some of them evolve by a birth-and-death process (Nei et al., 1997; Ota and Nei, 1994; Senger et al., 2015). It has previously been shown that immunoglobulin heavy constant gamma gene (*Igy*) number varies greatly among *Platyrrhines* and *Catarrhini* (Brusco et al., 1998; Garzon-Ospina and Buitrago, 2020; Olivieri and Gambon Deza, 2018), suggesting that IgG evolution was different among primates.

The *Igy* phylogenetic trees inferred showed a branching pattern that resembles the species' phylogeny, suggesting that this gene has evolved in a lineage and/or species-specific manner. The 12 *Lemuriformes*

and 10 *Platyrrhines* species displayed a single *Ig γ* gene in their genomes, suggesting that this region has remained stable for more than 20 million years in these primates. *Alpa* was the only one having a paralog gene that appears to have emerged recently.

On the other hand, the evolutionary history of this gene was highly complex in *Catarrhini*. Previous studies have shown that the IgG subclasses effector function appears to be species-specific among humans and macaques (Crowley and Ackerman, 2019; Nguyen et al., 2014; Ramesh et al., 2017; Warncke et al., 2012). For instance, the interaction of IgG2 and IgG4 to FcR receptors is more effective in cynomolgus monkey than in humans (Warncke et al., 2012). However, this would not necessarily be explained by changes in the critical amino acids of the IgG2 and IgG4 subclasses (Warncke et al., 2012), rather because these duplications are lineage-specific and therefore not orthologous. In addition to these lineage-specific duplications, species-specific duplications have also occurred within the *Catarrhini*. In *Cercopithecoidea*, the *Ig γ* number varies. Some species had a single-copy gene, while others had three or four paralogs. This behavior might be because the DNA region containing the immunoglobulin γ locus has not been fully sequenced yet and thus, some duplicates could be not found in this study. For instance, *Mafu*, *Mamu*, and *Mane* are closely related species with a split around 3.99 – 7.07 million years ago (Kumar et al., 2017). However, only one gene was found in *Mafu*, whereas four different genes were observed in *Mamu* and *Mane* genomes.

Although not all genes were found in *Mafu*, the data obtained for others *Cercopithecoidea* suggest an *Ig γ* differential expansion. *Cercopithecinae* (*Erpa*, *Thge*, *Masp*, *Ceat*, *Mane*, *Mafu*, and *Mamu*) and *Colobinae* (*Nala*, *Trfr*, *Rhro*, and *Pyne*) are two subfamilies of this group. While *Mamu* and *Mane*, belonging to *Cercopithecinae*, had four genes, *Nala*, *Trfr*, and *Rhro*, from the *Colobinae*, displayed a single one. It has been previously scanned some primate genomes in order to identify exon sequences encoding immunoglobulin CH domains (Olivieri and Gambon Deza, 2018). The authors found that in *Nala* and *Rhro* genomes, there was only one *Ig γ* gene, which is consistent with our findings. Thus, it can be suggested that *Ig γ* is a single-copy gene in these species. Therefore, *Ig γ* duplication took place after *Cercopithecinae* and *Colobinae* split. On the other hand, *Ceat* and *Pyne* had three genes; *Ig γ 2* and *Ig γ 3*, within these species, were more related to each other than to genes from other species. Because they appear to have diverged relatively recently (2-6 MYA), it can be suggested that *Ceat* and *Pyne* *Ig γ* paralogs have emerged by the species-specific duplication process (they are in-paralogs).

In *Hominidae*, species-specific duplications and gene loss were more noticeable. *Gogo* was the species with the lowest *Ig* gene number with two duplicates, while *Patr* showed eight paralogs. Most of them seem to have emerged before the main ape groups' divergence (out-paralogs). However, few of them had orthologs in the ape species assessed here. Hence, several duplicates could have been lost in some *Hominidae*-species evolutionary history. For instance, four of the six *Poab Ig* genes appear not to have orthologs in other *Hominidae* species because they form a monophyletic clade. This could be due to concerted evolution (a process that homogenizes gene duplicates) or species-specific duplication that arose after *Poab* diverged from other apes. Divergence times support the latter scenario. *PoabIg5* and *PoabIg3* split 5.8 MYA while *PoabIg2* and *PoabIg4* diverged 1.6 MYA. After gene duplication, *PoabIg3* and *PoabIg4* became pseudogenes (several premature stop codons were found), ruling out concerted evolution. Meanwhile, *PoabIg6* and *PoabIg1* could be *Hominidae* gene orthologs (or even pseudoorthologs) to *PatrIg8/HosaIg4* and *PatrIg1/GogoIg1*, respectively, due to their similarities in exon-intron structures. The remaining ancestral genes (*PatrIg2/HosaIg3*, *PatrIg5-6/HosaIg1-Ψ* and *PatrIg3-4*) were lost in *Poab*.

Gogo also appears to have lost most *Ig* orthologs because only two genes were located in its genome. Previous reports could support this hypothesis. (Olivieri and Gambon Deza, 2018) found 11 IgG encoding-exons in *Igh* locus. The first one was located just after the *Igδ* gene; the next three exons formed the first functional gene. Another exon was found next to the *Igε* paralog, while the last six exons (belonging to two different genes) were upstream of the *Igε-Igα* genes. However, one gene seems to have been inserted inside another, leaving only one of them as a functional gene (Olivieri and Gambon Deza, 2018). Therefore, most of the genes that emerged early in *Hominidae* evolution were lost or inactivated in this species.

Conversely, several *Ig* orthologs were shared between *Hosa* and *Patr*. However, in *Hosa*, some ancestral genes were also lost. In this species, no gene showed the same exon-intron structure as *PatrIg1/GogoIg1/PoabIg1*. Therefore, this gene seems to have been maintained in most ape species but lost in *Hosa*. Likewise, the *PatrIg4-3* lineage was also lost in *Hosa*.

Regarding *Patr*, the *PatrIg2* gene (which could have given origin to *PatrIg4-3* lineage) was put together with the *HosaIg3* in the phylogenetic trees. Both of them displayed a similar exon-intron structure, suggesting that they are orthologues. However, *PatrIg2* lost the first 842pb. It is unclear

whether this is a pseudogene in this species, or it encode for an IgG with just two immunoglobulin domains. Finally, *PatrIgy4* and *PatrIgy3* were species-specific duplicates, splitting around 2 MYA.

4.1. Natural selection could have favored IgG paralogue neofunctionalization

The outcomes before gene duplication might be subfunctionalization or neofunctionalization. The latter could be facilitated by functional constraints relaxation, followed by positive selection (Lynch and Conery, 2000; Wertheim et al., 2015). However, the number of sites involved in a new function is usually small (Gu, 2003; Newcomb et al., 1997; Perutz, 1984; Zhang et al., 1998) and, therefore, difficult to trace. Moreover, the positive selection acting at these sites could be transient or episodic, making its detection more problematic. However, if a gene/protein has experienced positive selection for a new or modified function, it could be expected that the lineage subject to positive selection shows a higher evolution rate than other phylogenetic lineages (Fay and Wu, 2003; Messier and Stewart, 1997). Thus, several tests have been developed to assess whether positive selection has occurred on a proportion of lineages (Smith et al., 2015). Furthermore, some codon-based methods are useful to identify sites under episodic selection signatures (Murrell et al., 2012).

Here, several tests displayed that the *Igy* family has experienced positive selection. Several internal branches displayed signatures of positive selection, which could indicate ancestral diversification in the *Igy* family. Moreover, several positive selected sites were found by codon-based methods, most of them located in the CH2 domain. Critical amino acid residues, binding to complement and effector cell receptors (Fc receptors), have been previously determined in *Hosa*. They have been located at the CH2 and CH3 domains (Idusogie et al., 2001; Vidarsson et al., 2014; Woof and Burton, 2004). Several of these binding sites were not conserved in the *Igy* primate alignment. Furthermore, some of them showed signatures of episodic positive selection. Diversity and polymorphism (allotypes) have proven to reduce subclass binding or affinity in *Hosa* IgG, as well as placental transport and protein half-life (Vidarsson et al., 2014). Thus, a change in amino acids, favored by positive selection, could have provoked a functional divergence in the *Igy* paralogs. IgG respond to different antigen types in humans, having a marked skewing towards one of the four IgG subclasses. Furthermore, these subclasses have different affinities to Fc-receptor or complement binding; thus, they can trigger different effector mechanisms (Irani et al., 2015; Jacobsen et al., 2011; Vidarsson et al., 2014). Accordingly, an *Igy* expansion, followed by positive selection, could have occasioned a high specialized IgG-response in *Catarrhines*.

Likewise, unequal crossover is an important mechanism for gene expansion (or contraction) (Zhang, 2003). Moreover, selective pressures for an expanded (or contracted) number of genes would be favorably selected in a population (Hood et al., 1980). Consequently, a gene might have an optimal number of copies per genome, depending on the environmental conditions (Garzon-Ospina et al., 2010; Kondrashov and Kondrashov, 2006). Thus, the *Ig γ* gene family in primate species could be evolving by a recurrent process of lineage-specific (or species-specific) gene duplication and gene loss (birth-and-death model). According to the data reported here, the *Ig γ* family might have expanded (or contracted) until reaching an optimal number of copies in each species' genome in order to generate novel strategies for pathogens recognition, neutralization or elimination as competitive ELISA assays previously suggest (Asada et al., 2002). Consequently, IgG-based immune responses might be species-specific, and thus protective immune response observed in vaccine trials using primate biomedical models could not be fully extrapolated to humans as lately has been suggested (Crowley and Ackerman, 2019; Garzon-Ospina and Buitrago, 2020; Nguyen et al., 2014; Ramesh et al., 2017; Warncke et al., 2012).

4.2. A hypothesis for *Ig γ* family evolution

A model for *Ig γ* family evolution in primates is proposed here based on the data obtained (Fig. 6). Primate ancestors had a single *Ig γ* gene. About 70.9 – 76.9 million years ago, *Strepsirrhini* and *Haplorrhini* split. In the former, *Ig γ* locus has been unchanging without duplication events. Then, 41.0 – 45.7 MYA, the *Platyrrhini*/*Catarrhini* divergence took place. Duplication events had not occurred in most *Platyrrhini* species; however, around 2.9 MYA, an *Ig γ* paralog arose in *Alpa*. Meanwhile, after *Cercopithecoidea*'s and *Hominidae*'s divergence, several lineage and species-specific duplicates emerged. *Cercopithecinae* *Ig γ 1* arose 21.6 MYA, then, *Ig γ 2* emerged (21.1 MYA), and, finally, *Ig γ 3* diverge from the ancestral *Ig γ 4* gene around 18.4 million years ago. It is not clear whether some genes were lost in *Erpa*, *Thge*, *Masp*, or *Ceat*. However, in the latter species, a new duplication took place 5.6 MYA. Likewise, two paralogs emerged recently in *Pyne*. Lineage *PyneIg γ 3-Ig γ 2* split from *PyneIg γ 1* around 6.6 MYA and *Ig γ 3* and *Ig γ 2* diverged 5.4 MYA. In other *Colobinae* species (*Nala*, *Trfr*, and *Rhro*) IgG was encoded by a single-copy gene.

Regarding the *Hominidae* species, the *Ig γ* was expanded around 20 MYA. The ancestral gene (*PatrIg γ 8/Hosalg γ 4/Gogolg γ 2/PoabIg γ 6*) gave rise to *PatrIg γ 7/Hosalg γ 2* (19.2 MYA); then, *PatrIg γ 2/Hosalg γ 3* arose (19.8 MYA). In this gene, the exon two underwent intragenic duplication. Moreover, the *PatrIg γ 2* gene lost 842pb towards the 5'-end. Later, the *PatrIg γ 5-6/Hosalg γ 1- Ψ* lineage

emerged around 18.9 MYA, followed by the rise of the *PatrIgy4-3* lineage about 11.3 million years ago. However, several of these genes were lost before the *Hominidae* split. For instance, in *Poab*, the orthologs to *PatrIgy2*, *PatrIgy5-6*, and *PatrIgy4-3* were not found at the genome. Nevertheless, in this species, new duplications took place a few million years ago. *PoabIgy3* diverged from *PoabIgy5* around 5.9 MYA; then, *PoabIgy2* and *PoabIgy4* diverged 1.7 MYA. Both the *PoabIgy3* and *PoabIgy4* have become pseudogenes. *Gogo* also lost most of the ancestral duplicates, maintaining just two *Igy* genes. *Hosa* lost the ortholog to *PatrIgy1*, as well as the ortholog to *PatrIgy4-3*. Moreover, *Hosa* orthologue to *PatrIgy6* was inactivated and is currently a pseudogene. Finally, the last duplication process in this superfamily took place around 2 MYA in *Patr*, giving rise to *PatrIgy4* and *PatrIgy3*.

Abbreviations

Alpa: *Alouatta palliata*, **Aona:** *Aotus nancymae*, **Atge:** *Ateles geoffroyi*, **Caja:** *Callithrix jacchus*, **Ceal:** *Cebus albifrons*, **Ceca:** *Cebus capucinus*, **Ceat:** *Cercocebus atys*, **Erpa:** *Erythrocebus patas*, **Eufl:** *Eulemur flavifrons*, **Eufu:** *Eulemur fulvus*, **Euma:** *Eulemur macaco*, **Gogo:** *Gorilla gorilla*, **Hosa:** *Homo sapiens*, **Inin:** *Indri indri*, **Mafu:** *Macaca fuscata*, **Mamu:** *Macaca mulata*, **Mane:** *Macaca nemestrina*, **Masp:** *Mandrillus sphinx*, **Migr:** *Microcebus griseorufus*, **Mimi:** *Microcebus mittermeieri*, **Mimu:** *Microcebus murinus*, **Mira:** *Microcebus ravelobensis*, **Mita:** *Microcebus tavaratra*, **Miza:** *Mirza zaza*, **Nala:** *Nasalis larvatus*, **Patr:** *Pan troglodytes*, **Pipi:** *Pithecia pithecia*, **Pldo:** *Plecturocebus donacophilus*, **Poab:** *Pongo abelii*, **Prsi:** *Prolemur simus*, **Prco:** *Propithecus coquereli*, **Pyne:** *Pygathrix nemaeus*, **Rhro:** *Rhinopithecus roxellana*, **Saim:** *Saguinus imperator*, **Sabo:** *Saimiri boliviensis*, **Saap:** *Sapajus apella*, **Thge:** *Theropithecus gelada*, **Trfr:** *Trachypithecus francoisi*.

Data statement

The datasets supporting this article's results are available in Mendeley Data: Garzón-Ospina, Diego; Buitrago, Sindy P (2020), "Immunoglobulin heavy constant gamma gene evolution is modulated by both the divergent and birth-and-death evolutionary models", Mendeley Data, V1, doi: 10.17632/r6vh2rfnts.1 (DOI is reserved but not active yet)

Acknowledgments

The authors wish to thank Gypsy Bonny Español for reviewing the manuscript.

439

440 **Funding information**

441 This work was supported by the Fundación para la Promoción de la Investigación y la Tecnología,
442 through the cooperation agreement #202111. This foundation was not involved in conceptualization,
443 propose development and the decision to submit the article for publication.

444

445 **Conflict of interest**

446 The authors declare that they have no conflict of interest.

447

448 **References**

- 449 Abascal, F., Zardoya, R., Telford, M.J., 2010. TranslatorX: multiple alignment of nucleotide
450 sequences guided by amino acid translations. *Nucleic Acids Res* 38, W7-13.
- 451 Anisimova, M., Nielsen, R., Yang, Z., 2003. Effect of recombination on the accuracy of the likelihood
452 method for detecting positive selection at amino acid sites. *Genetics* 164, 1229-1236.
- 453 Arenas M, P.D., 2014. The influence of recombination on the estimation of selection from coding
454 sequence alignments. In: FaresMA (Ed.), *Natural Selection: Methods and Applications*. CRC
455 Press/Taylor & Francis, Boca Raton, pp. 112-125.
- 456 Arenas, M., Posada, D., 2010. Coalescent simulation of intracodon recombination. *Genetics* 184, 429-
457 437.
- 458 Asada, Y., Kawamoto, Y., Shotaki, T., Terao, K., 2002. Molecular evolution of IgG subclass among
459 nonhuman primates: Implication of differences in antigenic determinants among apes.
460 *Primates* 43, 343-349.
- 461 Benson, D.A., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers, E.W., 2015. GenBank.
462 *Nucleic Acids Res* 43, D30-35.
- 463 Birdsall, H.H., 2015. 5 - Adaptive Immunity: Antibodies and Immunodeficiencies. In: Bennett, J.E.,
464 Dolin, R., Blaser, M.J. (Eds.), *Mandell, Douglas, and Bennett's Principles and Practice of*
465 *Infectious Diseases (Eighth Edition)*. Content Repository Only!, Philadelphia, pp. 34-49.e32.
- 466 Brusco, A., Carbonara, A.O., Crovella, S., Bottaro, A., 1998. Primate immunoglobulin heavy chain
467 constant gamma genes: an hypothesis of their evolution. *Human Evolution* 13, 49-56.
- 468 Crowley, A.R., Ackerman, M.E., 2019. Mind the Gap: How Interspecies Variability in IgG and Its
469 Receptors May Complicate Comparisons of Human and Non-human Primate Effector
470 Function. *Front Immunol* 10, 697.
- 471 Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2011. ProtTest 3: fast selection of best-fit models
472 of protein evolution. *Bioinformatics* 27, 1164-1165.
- 473 Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees.
474 *BMC evolutionary biology* 7, 214.
- 475 Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput.
476 *Nucleic Acids Res* 32, 1792-1797.
- 477 El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson,
478 L.J., Salazar, G.A., Smart, A., Sonnhammer, E.L.L., Hirsh, L., Paladin, L., Piovesan, D.,
479 Tosatto, S.C.E., Finn, R.D., 2019. The Pfam protein families database in 2019. *Nucleic Acids*
480 *Res* 47, D427-D432.

Esteves, M.B., Binaghi, R.A., 1972. Antigenic similarities among mammalian immunoglobulins. *Immunology* 23, 137-145.

Fay, J.C., Wu, C.I., 2003. Sequence divergence, functional constraint, and selection in protein evolution. *Annu Rev Genomics Hum Genet* 4, 213-235.

Garzon-Ospina, D., Buitrago, S.P., 2020. Igh locus structure and evolution in Platyrrhines: new insights from a genomic perspective. *Immunogenetics* 72, 165-179.

Garzon-Ospina, D., Cadavid, L.F., Patarroyo, M.A., 2010. Differential expansion of the merozoite surface protein (msp)-7 gene family in Plasmodium species under a birth-and-death model of evolution. *Mol Phylogenet Evol* 55, 399-408.

Gu, X., 2003. Functional divergence in protein (family) sequence evolution. *Genetica* 118, 133-141.

Hood, J.M., Huang, H.V., Hood, L., 1980. A computer simulation of evolutionary forces controlling the size of a multigene family. *J Mol Evol* 15, 181-196.

Huelsenbeck, J.P., Ronquist, F., Nielsen, R., Bollback, J.P., 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 294, 2310-2314.

Hughes, A.L., Nei, M., 1990. Evolutionary relationships of class II major-histocompatibility-complex genes in mammals. *Mol Biol Evol* 7, 491-514.

Idusogie, E.E., Wong, P.Y., Presta, L.G., Gazzano-Santoro, H., Totpal, K., Ultsch, M., Mulkerrin, M.G., 2001. Engineered antibodies with increased activity to recruit complement. *J Immunol* 166, 2571-2575.

Irani, V., Guy, A.J., Andrew, D., Beeson, J.G., Ramsland, P.A., Richards, J.S., 2015. Molecular properties of human IgG subclasses and their implications for designing therapeutic monoclonal antibodies against infectious diseases. *Mol Immunol* 67, 171-182.

Jacobsen, F.W., Padaki, R., Morris, A.E., Aldrich, T.L., Armitage, R.J., Allen, M.J., Lavalley, J.C., Arora, T., 2011. Molecular and functional characterization of cynomolgus monkey IgG subclasses. *J Immunol* 186, 341-349.

K., N., H., N., 1997. GeneDoc: A tool for editing and annotating multiple sequence alignments.

Kondrashov, F.A., Kondrashov, A.S., 2006. Role of selection in fixation of gene duplications. *J Theor Biol* 239, 141-151.

Kondrashov, F.A., Rogozin, I.B., Wolf, Y.I., Koonin, E.V., 2002. Selection in the evolution of gene duplications. *Genome Biol* 3, RESEARCH0008.

Kosakovsky Pond, S.L., Frost, S.D., 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol* 22, 1208-1222.

Kosakovsky Pond, S.L., Posada, D., Gravenor, M.B., Woelk, C.H., Frost, S.D., 2006. Automated phylogenetic detection of recombination using a genetic algorithm. *Mol Biol Evol* 23, 1891-1901.

Krakauer, D.C., Nowak, M.A., 1999. Evolutionary preservation of redundant duplicated genes. *Semin Cell Dev Biol* 10, 555-559.

Kumar, S., Stecher, G., Suleski, M., Hedges, S.B., 2017. TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Mol Biol Evol* 34, 1812-1819.

Lynch, M., Conery, J.S., 2000. The evolutionary fate and consequences of duplicate genes. *Science* 290, 1151-1155.

Messier, W., Stewart, C.B., 1997. Episodic adaptive evolution of primate lysozymes. *Nature* 385, 151-154.

Miller MA, P.W., Schwartz T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop (GCE)*. New Orleans, LA. p. 1-8.

Miller, M.A., Schwartz, T., Pickett, B.E., He, S., Klem, E.B., Scheuermann, R.H., Passarotti, M., Kaufman, S., O'Leary, M.A., 2015. A RESTful API for Access to Phylogenetic Tools via the CIPRES Science Gateway. *Evolutionary bioinformatics online* 11, 43-48.

Mishra, R., Panda, S.K., Sahoo, P.K., Mishra, S., Satapathy, A.K., 2019. Self-reactive IgG4 antibodies are associated with blocking of pathology in human lymphatic filariasis. *Cell Immunol* 341, 103927.

Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Kosakovsky Pond, S.L., Scheffler, K., 2013. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol* 30, 1196-1205.

Murrell, B., Weaver, S., Smith, M.D., Wertheim, J.O., Murrell, S., Aylward, A., Eren, K., Pollner, T., Martin, D.P., Smith, D.M., Scheffler, K., Kosakovsky Pond, S.L., 2015. Gene-wide identification of episodic selection. *Mol Biol Evol* 32, 1365-1371.

Murrell, B., Wertheim, J.O., Moola, S., Weighill, T., Scheffler, K., Kosakovsky Pond, S.L., 2012. Detecting individual sites subject to episodic diversifying selection. *PLoS Genet* 8, e1002764.

Napodano, C., Marino, M., Stefanile, A., Pocino, K., Scatena, R., Gulli, F., Rapaccini, G.L., Delli Noci, S., Capozio, G., Rigante, D., Basile, U., 2020. Immunological Role of IgG Subclasses. *Immunol Invest*, 1-18.

Nei, M., Gu, X., Sitnikova, T., 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proc Natl Acad Sci U S A* 94, 7799-7806.

Nei, M., Rooney, A.P., 2005. Concerted and birth-and-death evolution of multigene families. *Annu Rev Genet* 39, 121-152.

Newcomb, R.D., Campbell, P.M., Ollis, D.L., Cheah, E., Russell, R.J., Oakeshott, J.G., 1997. A single amino acid substitution converts a carboxylesterase to an organophosphorus hydrolase and confers insecticide resistance on a blowfly. *Proc Natl Acad Sci U S A* 94, 7464-7468.

Nguyen, D.C., Sanghvi, R., Scinicariello, F., Pulit-Penaloza, J., Hill, N., Attanasio, R., 2014. Cynomolgus and pigtail macaque IgG subclasses: characterization of IGHG genes and computational analysis of IgG/Fc receptor binding affinity. *Immunogenetics* 66, 361-377.

Nimmerjahn, F., Ravetch, J.V., 2005. Divergent immunoglobulin g subclass activity through selective Fc receptor binding. *Science* 310, 1510-1512.

Olivieri, D.N., Gambon Deza, F., 2018. Immunoglobulin genes in Primates. *Mol Immunol* 101, 353-363.

Ota, T., Nei, M., 1994. Divergent evolution and evolution by the birth-and-death process in the immunoglobulin VH gene family. *Mol Biol Evol* 11, 469-482.

Perutz, M.F., 1984. Species adaptation in a protein molecule. *Adv Protein Chem* 36, 213-244.

Posada, D., 2008. jModelTest: phylogenetic model averaging. *Mol Biol Evol* 25, 1253-1256.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst Biol* 67, 901-904.

Ramesh, A., Darko, S., Hua, A., Overman, G., Ransier, A., Francica, J.R., Trama, A., Tomaras, G.D., Haynes, B.F., Douek, D.C., Kepler, T.B., 2017. Structure and Diversity of the Rhesus Macaque Immunoglobulin Loci through Multiple De Novo Genome Assemblies. *Front Immunol* 8, 1407.

Schroeder, H.W., Jr., Cavacini, L., 2010. Structure and function of immunoglobulins. *J Allergy Clin Immunol* 125, S41-52.

Schur, P.H., 1987. IgG subclasses--a review. *Ann Allergy* 58, 89-96, 99.

Senger, K., Hackney, J., Payandeh, J., Zarrin, A.A., 2015. Antibody Isotype Switching in Vertebrates. *Results Probl Cell Differ* 57, 295-324.

Sjostrand, J., Sennblad, B., Arvestad, L., Lagergren, J., 2012. DLRS: gene tree evolution in light of a species tree. *Bioinformatics* 28, 2994-2995.

Sjostrand, J., Tofigh, A., Daubin, V., Arvestad, L., Sennblad, B., Lagergren, J., 2014. A Bayesian method for analyzing lateral gene transfer. *Syst Biol* 63, 409-420.

- Smith, M.D., Wertheim, J.O., Weaver, S., Murrell, B., Scheffler, K., Kosakovsky Pond, S.L., 2015. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol Biol Evol* 32, 1342-1353.
- Stephens, S.G., 1951. Possible significances of duplication in evolution. *Adv Genet* 4, 247-265.
- Su, C., Nei, M., 2001. Evolutionary Dynamics of the T-Cell Receptor VB Gene Family as Inferred from the Human and Mouse Genomic Sequences. *Molecular Biology and Evolution* 18, 503-513.
- Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., Rambaut, A., 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol* 4, vey016.
- Sun, Y., Liu, Z., Ren, L., Wei, Z., Wang, P., Li, N., Zhao, Y., 2012. Immunoglobulin genes and diversity: what we have learned from domestic animals. *J Anim Sci Biotechnol* 3, 18.
- Takahashi, N., Ueda, S., Obata, M., Nikaido, T., Nakai, S., Honjo, T., 1982. Structure of human immunoglobulin gamma genes: implications for evolution of a gene family. *Cell* 29, 671-679.
- Vidarsson, G., Dekkers, G., Rispens, T., 2014. IgG subclasses and allotypes: from structure to effector functions. *Front Immunol* 5, 520.
- Warncke, M., Calzascia, T., Coulot, M., Balke, N., Touil, R., Kolbinger, F., Heusser, C., 2012. Different adaptations of IgG effector function in human and nonhuman primates and implications for therapeutic antibody treatment. *J Immunol* 188, 4405-4411.
- Weaver, S., Shank, S.D., Spielman, S.J., Li, M., Muse, S.V., Kosakovsky Pond, S.L., 2018. Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary processes. *Mol Biol Evol*.
- Wertheim, J.O., Murrell, B., Smith, M.D., Kosakovsky Pond, S.L., Scheffler, K., 2015. RELAX: detecting relaxed selection in a phylogenetic framework. *Mol Biol Evol* 32, 820-832.
- Woof, J.M., Burton, D.R., 2004. Human antibody-Fc receptor interactions illuminated by crystal structures. *Nat Rev Immunol* 4, 89-99.
- Zhang, J., 2003. Evolution by gene duplication: an update. *Trends in Ecology & Evolution* 18, 292-298.
- Zhang, J., Rosenberg, H.F., Nei, M., 1998. Positive Darwinian selection after gene duplication in primate ribonuclease genes. *Proc Natl Acad Sci U S A* 95, 3708-3713.

607 **Tables**

608

609 **Table 1. Paralogues divergence times in *Hominidae***

<i>Hosa</i>		<i>Patr</i>		<i>Gogo</i>		<i>Poab</i>	
<i>Igys</i>	Divergence times	<i>Igys</i>	Divergence times	<i>Igys</i>	Divergence times	<i>Igys</i>	Divergence times
		4/3	2.05 (0.01)			2/ Ψ4	1.67 (0.01)
Ψ/1**	14.48 (0.05)	6/5**	9.80 (0.03)			1/2-Ψ4	5.79 (0.02)
Ψ/3-1	16.36 (0.05)	4-3-2/1	11.53 (0.05)			5/ Ψ3	5.89 (0.02)
4-2/3-1-Ψ**	17.02 (0.05)	8-7/6-5**	18.96 (0.05)				
		8/1*	19.18 (0.06)	2-1*	7.97 (0.02)	6/1*	11.20
4-2/3-1§	17.26 (0.05)	8-7/4-3-2§	19.21 (0.06)				
4/2*‡	12.87 (0.05)	8/7*‡	19.24 (0.06)			6/5‡	11.59
<i>Hosa Igγ</i>	18.45 (0.05)	<i>Patr Igγ</i>	20.29 (0.06)			<i>Poab Igγ</i>	12.54 (0.03)
	<i>Hosa/Poab</i> ^a		14.7 - 16.7				
	<i>Hosa/Gogo</i> ^a		8.38 - 9.58				
	<i>Hosa/Patr</i> ^a		5.10 - 11.8				

610 * and **: Putative orthologs

611 *, **, §, ‡ Putative pseudoorthologs, they had similar location at the genome as well as exon-intron structures.

612 ^a: Divergence times based on timetree web server data (<http://www.timetree.org/> (Kumar et al., 2017)).

613

614

615

616

617

618 **Table 2. Paralogues divergence times in *Cercopithecoidea***

<i>Mamu</i>		<i>Mane</i>		<i>Ceat</i>		<i>Erpa</i>		<i>Pyne</i>	
<i>Igys</i>	Divergence times	<i>Igys</i>	Divergence times	<i>Igys</i>	Divergence times	<i>Igys</i>	Divergence times	<i>Igys</i>	Divergence times
4/3	18.47 (0.06)	4/3	16.84 (0.06)	3/2	5.63 (0.02)	2/1	12.33 (0.02)	3/2	5.49 (0.01)
4-3/2	21.10 (0.06)	4-3/2	21.47 (0.06)	3-2/1	13.33 (0.03)			3-2/1	6.65 (0.02)
4-3/1	21.65 (0.06)	4-3/1	19.99 (0.05)						
<i>Mamu Igγ</i>	21.31 (0.06)	<i>Mane Igγ</i>	22.29 (0.07)	<i>Ceat Igγ</i>	13.60 (0.06)			<i>Pyne Igγ</i>	7.21 (0.02)
<i>Catarrhini</i> split ^a			27.9 – 31.3						
<i>Cercopithecoidea</i> split ^a			16.4 – 22.2						
<i>Mamu/Pyne</i> ^a									
<i>Mamu/Erpa</i> ^a			10.7 – 16.6						
<i>Mamu/Ceat</i> ^a			9.30 – 14.4						
<i>Mamu/Mane</i> ^a			3.99 – 7.07						

619 ^a: Divergence times based on timetree web server data (<http://www.timetree.org/> (Kumar et al., 2017)).

Figure legends

Figure 1. Schematic representation of the chromosomal *Ig γ* loci in 38 primate genomes. The δ and ε genes flanking the γ chromosome locus are represented by purple and fuchsia boxes, respectively. Light purple and fuchsia boxes represent the expected δ and ε genes since these genomes are not fully sequenced or properly ensembled. γ genes are depicted in shades of green while pseudogenes are displayed in black. Structure, size and number of exon-intron for each *Ig γ* is represented to scale. However, the length of intergenic regions is not informative.

Figure 2. Bayesian (BY) tree inferred for the 67 sequences from the multigene *Ig γ* family. Four major clades were identified clustering DNA sequences in agreement with primate phylogenetic relationships. The branching pattern shows posterior probabilities higher than > 0.95 (numbers on branches) supporting the topology. Purple clade cluster the *Strepsirrhini* *Ig γ* gene sequences; red clade put together *Platyrrhini* genes; yellow clade group *Cercopithecoidea* paralogues and the clade clustering *Ig γ* duplicates from *Hominoidea* are depicted in blue.

Figure 3. γ gene family phylogeny inferred by the DLTRS evolutionary model. Both BY and DLTRS trees showed similar topologies. A. Species tree used for generating the γ gene tree. B. γ gene tree inferred by evolving down the species tree. Numbers on branches are posterior probability values. Purple clade cluster the *Strepsirrhini* *Ig γ* gene sequences; red clade put together *Platyrrhini* genes; yellow clade group *Cercopithecoidea* paralogues and the clade clustering *Ig γ* duplicates from *Hominoidea* are depicted in blue.

Figure 4. Identification of branches under positive selection. BY tree was analyzed by the aBSREL method. The shade of each color on branches indicates selection strength; red indicates positive selection ($\omega > 1$) and green negative selection ($\omega < 1$). The size of each colored segment represents the percentage of selected sites in the corresponding ω class. Branches have been classified as undergoing episodic diversifying selection, by the p-value corrected for multiple testing, using the Holm–Bonferroni method at $p < 0.05$.

Figure 5. Sliding window for the ω rate. Primate *Ig γ* evolutionary rate (ω : K_N/K_S) was draw based on SLAC data. A diagram of IgG protein is given below the sliding window. Positive selected sites (red) and negative selected sites (green) are shown by vertical lines within Ig domains (CH1, CH2

and CH3) or hinge region (grey box). Location of sites involved in binding to complement and/or effector cell receptors reported for human are shown as color boxes below protein scheme; IgG-Fc receptors binding sites are displayed in blue, orange box represent the sites involved in complement component 1q binding while neonatal Fc receptor binding sites are showed in dark blue.

Figure 6. Evolutionary model for the *Igγ* multigene family. Primate ancestor had a single *Igγ* gene. 70.9 - 76.9 million years ago *Strepsirrhini* and *Haplorrhini* split. In *Lemuriformes* as well as *Platyrrhini* *Igγ* has not undergo duplication. *Alpa* is the only *Platyrrhini* showing a paralog gene. After *Cercopithecoidea* and *Hominidae* divergence (27.9 – 31.3 MYA) several duplicates have emerged. Three duplication events took place in *Cercopithecinae* subfamily whereas in most *Colobinae* species *IgG* is encoded by a single-copy gene. Recent specie-specific duplication increased *Igγ* copy number in some *Cercopithecoidea* species. *Hominidae* paralogs arose before apes' divergence however, several of them has been lost or become in pseudogenes. *Hosa* and *Patr* shared several orthologs while *PoabIgγ* duplicates have arose by specie-specific duplication. 1. *PoabIgγ2* and *PoabIgγ4* divergence. 2. *PoabIgγ1* and *PoabIgγ2-4* divergence. 3. *PoabIgγ5* and *PoabIgγ3* divergence 4. Loss of ortholog to *PatrIgγ2*. 5. Loss of ortholog to *PatrIgγ5-6* lineage. 6. Loss of ortholog to *PatrIgγ3-4* lineage. 7. *Gogo* lost the ortholog to *PatrIgγ7/HosaIgγ2*. 8. *PatrIgγ6* ortholog pseudogenization in *Hosa*. 9. Loss of ortholog to *PatrIgγ1*. 10. *PatrIgγ4* and *PatrIgγ3* divergence. 11. *PatrIgγ6* (*HosaIgγΨ*) and *PatrIgγ5* (*HosaIgγ1*) divergence. 12. Origin of *PatrIgγ3-4* lineage. 13. Origin of *PatrIgγ5-6* (*HosaIgγ1-Ψ*) lineage. 14. Origin of *PatrIgγ1* (*PoabIgγ1*, *GogoIgγ1*). 15. Origin of *PatrIgγ2* (*HosaIgγ3*). 16. Origin of *PatrIgγ7* (*HosaIgγ2*, *PoabIgγ5*, *GogoIgγ2*). 17. *CeatIgγ2* and *CeatIgγ3* divergence. 18. *MamulIgγ3* (*ManeIgγ3*) and *MamulIgγ4* (*ManeIgγ4*) divergence. 19. Origin of *MamulIgγ2* (*ManeIgγ2*). 20. Origin of *MamulIgγ1* (*ManeIgγ1*). 21. *PyneIgγ2* and *PyneIgγ3* divergence. 22. *PyneIgγ2-3* and *PyneIgγ1* divergence. 23. *AlpaIgγ1* and *AlpaIgγ2* divergence. Orange circles represent speciation events; red triangles symbolize duplication events; grey hexagon denote gene lost; unfilled hexagon means loss of the first *PatrIgγ2* exon and black diamond indicate pseudogenization. ?: Loss of genes or absence due to incomplete genomes.

Supplementary files

Supplementary material 1. GenBank accession numbers from 38 primate species used here. *Igγ* genes were searching using genome sequences available in GenBank database.

Supplementary material 2. Data regarding *Igγ* exon-intron structure. Exon and intron base pair length for each *Igγ* gene are shown

Supplementary material 3. Bayesian (BY) tree inferred for the 67 amino acid sequences from the multigene IgG family. Four major clades were identified clustering amino acid sequences in agreement with primate phylogenetic relationships. The branching pattern shows posterior probabilities higher than > 0.95 (numbers on branches) supporting the topology. Purple clade cluster the *Strepsirrhini* IgG protein sequences; red clade put together *Platyrrhini* genes; yellow clade group *Cercopithecoidea* paralogues and the clade clustering *Igγ* duplicates from *Hominoidea* are depicted in blue.

Supplementary material 4. IgG protein family phylogeny inferred by the DLTRS evolutionary model. Both BY and DLTRS trees showed similar topologies using DNA or proteins sequences. A. Species tree used for generating the IgG tree. B. IgG tree inferred by evolving down the species tree. Numbers on branches are posterior probability values. Purple clade cluster the *Strepsirrhini* IgG protein sequences; red clade put together *Platyrrhini* genes; yellow clade group *Cercopithecoidea* paralogues and the clade clustering *Igγ* duplicates from *Hominoidea* are depicted in blue.

Supplementary material 5. Negative and positive selected sites identified by codon-based methods. Codon-based methods identified sites under natural selection. Location of these sites within Ig-domains (CH1-CH3) or hinge regions are indicated.

Supplementary material 6. Alignment of deduced IgG amino acid sequences. The human IgG critical amino acid residues binding to complement, and effector cell receptors are shown. They are located at the CH2 and CH3 domains. FcγRn (IgG-Fc receptors) binding sites are highlighted in light blue, while yellow sites are those involved in Cq1 (complement component 1q) binding. Yellow asterisks correspond to sites involved in binding to both FcγRn and Cq1. FcRn (neonatal Fc receptor)

binding sites are highlighted in dark blue. CH (Ig) domains and hinge region were determined using the EU numbering tool. Positive selected sites within CH2 and CH3 domains are indicated.

Supplementary material 7. Relaxed or intensified natural selection between reference and tested branches. Selection intensity (relaxation or intensification) found by the RELAX method is shown for pairwise-lineage comparisons.

751 **Figure 1**

Figure 4

765 **Figure 5**

766

767

Figure 6

CRediT author statement

Diego Garzón-Ospina: Conceptualization, Investigation, Formal analysis, Methodology, Visualization, Writing- Original draft preparation. **Sindy P. Buitrago:** Investigation, Methodology, Visualization, Formal analysis, Writing- Reviewing and Editing

Highlights

- The evolution of the immunoglobulin gamma has been assessed in primates
- Immunoglobulin gamma has been differentially expanded in primates
- Several lineage/specie-specific paralogues were found in primates
- Immunoglobulin gamma in primates evolves by divergent and birth-and-death models

Immunoglobulin heavy constant gamma gene evolution is modulated by both the divergent and birth-and-death evolutionary models

Supplementary material 1. GenBank accession numbers from 38 primate species used here

		Species	Code	Accession number
1	Lemur	<i>Microcebus murinus</i>	<i>Mimu</i>	ABDC03036542.1
2		<i>Microcebus tavaratra</i>	<i>Mita</i>	VSME01204543.1
3		<i>Microcebus griseorufus</i>	<i>Migr</i>	VSMI01045886.1
4		<i>Microcebus mittermeieri</i>	<i>Mimi</i>	VSMH01406796.1
5		<i>Microcebus ravelobensis</i>	<i>Mira</i>	VSMG01235548.1
6		<i>Mirza zaza</i>	<i>Miza</i>	VSMD01117104.1
7		<i>Propithecus coquereli</i>	<i>Prco</i>	JZKE01178794.1
8		<i>Indri indri</i>	<i>Inin</i>	RJWJ010004486.1
9		<i>Eulemur flavifrons</i>	<i>Eufi</i>	LGHW01004402.1
10		<i>Eulemur macaco</i>	<i>Euma</i>	LGHX01004347.1
11		<i>Eulemur fulvus</i>	<i>Eufu</i>	PVJU010070514.1
12		<i>Prolemur simus</i>	<i>Prsi</i>	MPIZ01000217.1
13	Platyrrhine	<i>Callithrix jacchus</i>	<i>Caja</i>	NC_013905.1; NC_013911.1
14		<i>Saguinus imperator</i>	<i>Saim</i>	PVHO010036867.1; PVHO010174353.1; PVHO010059366.1; PVHO010062799.1; PVHO010145012.1; PVHO010052763.1; PVHO010064649.1; PVHO010032362.1
15		<i>Cebus capucinus</i>	<i>Ceca</i>	NW_003943660.1
16		<i>Cebus albifrons</i>	<i>Ceal</i>	PVKJ010080388.1; PVKJ010058919.1; PVKJ010095968.1; PVKJ010062743.1; PVKJ010097001.1; PVKJ010055809.1
17		<i>Sapajus apella</i>	<i>Saap</i>	WRPQ000000000.1
18		<i>Saimiri boliviensis</i>	<i>Sabo</i>	NW_003943660.1
19		<i>Aotus nancymae</i>	<i>Aona</i>	NW_018503652.1
20		<i>Ateles geoffroyi</i>	<i>Atge</i>	PVHS01058872.1; PVHS01029498.1; PVHS01055800.1; PVHS01050565.1; PVHS01079306.1; PVHS01143702.1; PVHS01046096.1
21		<i>Alouatta palliata</i>	<i>Alpa</i>	PVKV010051123.1; PVKV010021875.1; PVKV010060016.1; PVKV010120891.1; PVKV010063013.1; PVKV010096668.1; PVKV010036590.1; PVKV010060166.1; PVKV010015137.1

22		<i>Pithecia pithecia</i>	<i>Pipi</i>	PVIP01013496.1; PVIP01045506.1; PVIP01025469.1; PVIP01044121.1; PVIP01022314.1
23		<i>Plecturocebus donacophilus</i>	<i>Pldo</i>	PVKP010051085.1; PVKP010061223.1; PVKP010077881.1; PVKP010046511.1; PVKP010065619.1; PVKP010022991.1
24	<i>Cercopithecidae</i>	<i>Macaca fuscata</i>	<i>Mafu</i>	BFBW00000000.1
25		<i>Macaca mulata</i>	<i>Mamu</i>	NC_041760.1
26		<i>Macaca nemestrina</i>	<i>Mane</i>	JZLF00000000.1
27		<i>Mandrillus sphinx</i>	<i>Masp</i>	SRPC00000000.1
28		<i>Cercocebus atys</i>	<i>Ceat</i>	JZLG00000000.1
29		<i>Theropithecus gelada</i>	<i>Thge</i>	NC_037675.1
				NC_037674.1
30		<i>Erythrocebus patas</i>	<i>Erpa</i>	PVJV000000000.1
31		<i>Pygathrix nemaeus</i>	<i>Pyne</i>	PVHW000000000.1
32		<i>Nasalis larvatus</i>	<i>Nala</i>	CM003015.1
33		<i>Rhinopithecus roxellana</i>	<i>Rhro</i>	NC_044555.1
34		<i>Trachypithecus francoisi</i>	<i>Trfr</i>	VVIV00000000.1
35	<i>Hominidae</i>	<i>Pongo abelii</i>	<i>Poab</i>	NC_036917.1
36		<i>Gorilla gorilla</i>	<i>Gogo</i>	NC_044616.1
37		<i>Homo sapiens</i>	<i>Hosa</i>	NC_000014.9
38		<i>Pan troglodytes</i>	<i>Patr</i>	NC_036893.1

Igy genes were searching using genome sequences available in GenBank database.

		Species	# per genome	Available under aCC-BY-NC-ND 4.0 International license.	Pseudogenes	Exon 1	Intron 1	Exon 2	Intron 2	Exon 3	Intron 3	Exon 4	Intron 4	Exon 5	Intron 5	Exon 6	Total gene	IgG Domains
1	Lemur	<i>Microcebus murinus</i>	1	-	295	223	30	112	330	100	320						1410	3
2		<i>Microcebus tavaratra</i>	1	-	295	223	30	112	330	100	320						1410	3
3		<i>Microcebus griseorufus</i>	1	-	295	223	30	112	330	100	320						1410	3
4		<i>Microcebus mittermeieri</i>	1	-	295	223	30	112	330	100	320						1410	3
5		<i>Microcebus ravelobensis</i>	1	-	295	223	30	112	330	100	320						1410	3
6		<i>Mirza zaza</i>	1	-	295	227	33	112	330	100	320						1417	3
7		<i>Propithecus coquereli</i>	1	-	295	259	36	124	330	125	320						1489	3
8		<i>Indri indri</i>	1	-	295	257	41	122	330	118	320						1483	3
9		<i>Eulemur flavifrons</i>	1	-	295	234	39	124	330	119	320						1461	3
10		<i>Eulemur macaco</i>	1	-	295	234	39	124	330	119	320						1461	3
11		<i>Eulemur fulvus</i>	1	-	295	234	39	124	330	119	320						1461	3
12		<i>Prolemur simus</i>	1	-	295	231	39	124	330	109	320						1448	3
13	Platyrrhine	<i>Callithrix jacchus</i>	1	-	652	66	320										1038	3
14		<i>Saguinus imperator</i>	1	-	286	414	30	117	330	91	323						1591	3
15		<i>Cebus capucinus</i>	1	-	283	372	33	117	330	92	323						1550	3
16		<i>Cebus albifrons</i>	1	-	283	372	33	117	330	92	323						1550	3
17		<i>Sapajus apella</i>	1	-	286	370	33	117	330	92	320						1548	3
18		<i>Saimiri boliviensis</i>	1	-	296	375	23	117	330	92	323						1556	3
19		<i>Aotus nancymae</i>	1	-	286	373	30	117	330	90	323						1549	3
20		<i>Ateles geoffroyi</i>	1	-	283	391	33	115	330	92	323						1567	3
21		<i>Alouatta palliata</i>	1	-	298	370	33	109	330	92	272						1504	3
22			2	-	289	370	33	118	324	92	272						1498	3
23		<i>Pithecia pithecia</i>	1	-	289	389	33	118	330	92	323						1574	3
24		<i>Plecturocebus donacophilus</i>	1	-	280	388	36	118	330	92	323						1567	3
25	Cercopithecidae	<i>Macaca fuscata</i>	1	-	292	393	30	118	330	97	320						1580	3
26			2	-	292	391	54	118	330	97	320						1602	3
27		<i>Macaca mulata</i>	3	-	292	388	36	118	327	97	320						1578	3
28			4	-	292	389	51	118	330	97	320						1597	3
29			4	-	292	393	30	118	330	97	320						1580	3
30			1	-	292	391	54	118	330	97	320						1602	3
31		<i>Macaca nemestrina</i>	2	-	292	388	27	118	327	97	320						1569	3
32			3	-	292	369	51	118	330	97	166	35	118				1576	3
33			4	-	292	389	30	118	330	97	320						1576	3
34		<i>Mandrillus sphinx</i>	1	-	292	388	54	118	330	97	320						1599	3
35			1	-	292	389	54	118	330	97	320						1600	3
36		<i>Cercocebus atys</i>	2	-	292	389	36	118	330	97	320						1582	3
37			3	-	292	523	-	-	330	97	320						1562	3
38	Hominidae	<i>Theropithecus gelada</i>	1	-	292	369	51	118	330	97	320						1577	3
39			1	-	292	391	54	118	330	97	320						1602	3
40		<i>Erythrocebus patas</i>	2	-	292	354	36	118	330	97	320						1547	3
41			1	-	292	370	54	118	330	97	320						1581	3
42			2	-	292	389	36	118	330	97	320						1582	3
43			3	-	292	389	36	118	330	97	320						1582	3
44		<i>Pygathrix nemaeus</i>	1	-	292	370	54	118	330	97	320						1581	3
45			2	-	292	389	36	118	330	97	320						1582	3
46			3	-	292	389	36	118	330	97	320						1582	3
47		<i>Nasalis larvatus</i>	1	-	292	479	36	118	330	97	320						1672	3
48		<i>Rhinopithecus roxellana</i>	1	-	292	370	51	118	330	97	320						1578	3
49		<i>Trachypithecus francoisi</i>	1	-	292	370	54	118	330	97	320						1581	3
50	Hominidae		1	-	292	392	51	143	45	118	330	97	320				1788	3
51			2	-	292	392	51	695	330	97	320						2177	3
52		<i>Pongo abelii</i>	3	-	281	298	12	97	39	118	331	97	319				1592	-
53			4	-	290	388	50	677	327	97	319						2148	-
54			5	-	292	392	45	118	330	97	320						1594	3
55			6	-	292	392	45	118	330	97	320						1594	3
56		<i>Gorilla gorilla</i>	1	-	292	390	51	143	45	118	330	97	320				1786	3
57			2	-	292	392	42	161	330	97	320						1634	3
58			3	-	292	392	51	143	45	331	45	118	330	97	323		1747	3
59			4	-	292	391	45	118	330	97	323						1596	-
60		<i>Homo sapiens</i>	1	-	292	389	51	118	330	97	321						1598	3
61			2	-	292	392	36	118	227	97	323						1485	3
62			4	-	292	392	36	118	330	97	323						1588	3
63	Hominidae		1	-	292	392	51	143	45	118	330	97	320				1788	3
64			2	-	81	143	46	142	45	139	45	117	330	97	320		1088	2
65			3	-	292	392	51	894	45	329	330	97	320				2750	3
66			4	-	292	392	51	519	45	306	330	97	320				2352	3
67			5	-	292	391	45	118	330	97	320						1593	3
68			6	-	292	389	51	118	330	97	320						1597	3
69			7	-	292	392	36	118	327	97	320						1582	3
70			8	-	292	392	36	118	330	97	320						1585	3

Exon and intron base pair length for each *Ig* gene are shown

Immunoglobulin heavy constant gamma gene evolution is modulated by both the divergent and birth-and-death evolutionary models

Supplementary material 3. Bayesian (BY) tree inferred for the 67 amino acid sequences from the multigene IgG family

Immunoglobulin heavy constant gamma gene evolution is modulated by both the divergent and birth-and-death evolutionary models

Supplementary material 4. IgG protein family phylogeny inferred by the DLTRS evolutionary model

Immunoglobulin heavy constant gamma gene evolution is modulated by both the divergent and birth-and-death evolutionary models

Supplementary material 5. Negative and positive selected sites identified by codon-based methods

Selected negative sites		Selected Positive sites	
8	CH1	16	CH1
9		20	
21		70	
29		73	
35		78	
36		84	
38		91	
40		93	
42		94	
45		95	
50		99	
51		100	
55		102	Hinge
56		103	
62		109	
63		157	
64		158	
68		159	
74		161	
77		162	
81		167	CH2
85		168	
86		169	
87		171	
89		184	
161	Hinge	190	
173	CH2	211	
174		217	
180		227	
184		229	
186		231	
188		243	
195		244	
196		262	
197		271	
198		272	
199		274	

200		300	CH3
206		311	
213		312	
214		352	
215		355	
220		376	
224		379	
226			
234			
236			
239			
240			
242			
249			
250			
263			
264			
266			
276	CH3		
302			
303			
306			
310			
330			
334			
335			
341			
344			
351			
359			
364			
368			
369			
373			
375			
377			
378			
386			
394			
399			

Codon-based methods identified sites under natural selection. Location of these sites within Ig-domains (CH1-CH3) or hinge regions are indicated.

Immunoglobulin heavy constant gamma gene evolution is modulated by both the divergent and birth-and-death evolutionary models

Supplementary material 6. Alignment of deduced IgG amino acid sequences

CH1 Domain

Mimu IgG	AKTT--PPSV	FPLAPGNQAT	SGSSVALGCL	VKNYFPEPVT	VTWNSGSVTS	GVYTFPAALD	SSGLYSLSSM	VTVPAGSWPS	KTYTCNVAHP	ASSTKVDKTI
Mita IgG	--									
Migr IgG	--									
Mimi IgG	--		...M.....							
Mira IgG	--									
Miza IgG	--	..A.....	..T.....			H.....	S..S..	Q.....		
Prco IgG	--	..A.....	..V...KG.....	..T.....	K.....	..L.....	..H...RL.....	..S.....	Q.....	V
Inin IgG	--	..A.....	...SKG.....	..T.....		..L.....	..H...L.....	..S..S..	Q.....	V
Eufl IgG	--	..A.....	...SKT.....	..T.....	..S.....	..L.....	..H...SI.Q.....	..S..T..	Q.....	...R.....V
Euma IgG	--	..A.....	...SKT.....	..T.....	..S.....	..L.....	..H...SI.Q.....	..S..T..	Q.....	...R.....V
EuFu IgG	--	..A.....	...SKT.....	..T.....	..S.....	..L.....	..H...SI.Q.....	..S..T..	Q.....	...R.....V
Prsi IgG	--	..A.....	...SK.....	..T.....	..S.....	..L.....	..H...SI.Q.....	..S..T..	Q.....	...R.....V
Poab IgG1	S-.K--G...	...SCSRS..	..GTA.....	..D.....	..S.....	TL.....	..H...V.Q.....	..V...SS.LGT	Q.....D.K	P.N...RV
Poab IgG2	S-.K--G...	...SCSRS..	..EGTA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q...I...D.K	P.N...KV
Poab IgG5	S-.K--G...	...SSRS..	..GTA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q.....D.K	P.N...KV
Poab IgG6	T-.K--G...	...SSRS..	..GTA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q.....D.K	P.N...RV
Gogo IgG1	S-.K--G...	...CSRS..	..GTE.....	..D...K...	..S.....	SAL.....	..H...V.Q.....	..V...SSRLGT	Q.....N.K	P.T...RV
Gogo IgG2	S-.K--G...	...CSRS..	..GTA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q.....N.K	P.N...RV
Patr IgG1	S-.K--G...	...CSRS..	..EGTA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...A..SS.LGT	Q.....N.K	P.N...RV
Patr IgG2	S-QIPVT.NL	LSAE.KSCD-	-----	-----	-----	---	---PPR AH.AQ.PNLV	THL.PCPRCP	EPKS.DIPP.	CPRCPEP.S-
Patr IgG3	S-.K--G...	...CSRS..	..EGTA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...L...SSTFGT	Q.....N.K	P.N...RV
Patr IgG4	S-.K--G...	...CSRS..	..EGTA.....	..D...L...	..S.....	AL.....	..H...V.Q.....	..V...L...SSTFGT	Q.....N.K	P.N...RV
Patr IgG5	S-.K--G...	...SSKS..	..GTA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q...I...D.K	P.N...KV
Patr IgG6	S-.K--G...	...V.SSRS..	..EGTA.....	..D.....	..S.....	AL.R	S.PA...V.Q.....	..V...SS.LGT	Q.....D.K	P.N...RV
Patr IgG7	S-.K--G...	...CSRS..	..EGTA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...A..SSNFGT	Q.....D.K	P.N...V
Patr IgG8	S-.K--G...	...CSRS..	..EGTA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q.....D.K	P.N...V
Hosa IgG1	S-.K--G...	...SSKS..	..GTA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q...I...N.K	P.N...KV
Hosa IgG2	S-.K--G...	...CSRS..	..E.TA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SSNFGT	Q.....D.K	P.N...V
Hosa IgG3	S-.K--G...	...CSRS..	..GTA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q.....N.K	P.N...RV
Hosa IgG4	S-.K--G...	...CSRS..	..E.TA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	D.K	P.N...RV
Mafu IgG	S-.K--G...	...SSRS..	..E.TA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q.V...V.E	P.N...V
Mamu IgG1	S-.K--G...	...SSRS..	..E.TA.....	..D.....	..S.....	L.....	..H...V.Q.....	..V...SS.LGT	Q.V...N.K	P.N...RV
Mamu IgG2	S-.K--G...	...SCSRS..	..Q.TA.....	..D.....	..S.....	AL.....	..H...Q.V.Q.....	..V...SS.LGT	Q.V...V.E	P.N...V
Mamu IgG3	S-.K--G...	...SCSRS..	..Q.TA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q.V...V.E	P.N...RV
Mamu IgG4	S-.K--G...	...SSRS..	..E.TA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q.V...V.E	P.N...RV
Mame IgG1	S-.K--G...	...SSRS..	..E.TA.....	..D.....	..S.....	L.....	..H...V.Q.....	..V...SS.LGT	Q.V...N.K	P.N...RV
Mame IgG2	S-.K--G...	...SSRS..	..Q.TA.....	..D.....	..S.....	L.....	..H...V.Q.....	..V...SS.LGT	Q.V...V.E	P.N...RV
Mame IgG4	S-.K--G...	...SSRS..	..E.TA.....	..D.....	..S.....	L.....	..H...V.Q.....	..V...SS.LGT	Q...I...V.E	P.N...RV
Mame IgG3	S-.K--G...	...SSRS..	..E.T.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q.V...V.E	P.N...RV
Masp IgG	S-.K--G...	...SSRS..	..Q.TA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q.V...N.K	P.N...KV
Ceat IgG1	S-.K--G...	...SCSRS..	..E.TA.....	..D.....	..S.....	L.....	..H...V.Q.....	..V...SS.LGT	Q.V...N.K	P.N.R...RV
Ceat IgG2	S-.K--G...	...SCSRS..	..Q.TA.....	..D.....	..S.....	L.....	..H...V.Q.....	..V...SS.LGT	Q.V...N.K	P.N.R...RV
Ceat IgG3	S-.K--G...	...SCSRS..	..Q.TA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q.....V.E	P.N.R...RV
Thge IgG	S-.K--G...	...SCSKS..	..E.TA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q.V...V.E	P.N.R...RV
Erpa IgG2	S-.K--G...	...SCSRS..	..Q.TA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q.....V.E	P.N...RV
Erpa IgG1	S-.K--G...	...SSRS..	..E.TA.....	..D.....	..S.....	L.....	..H...V.Q.....	..V...SS.LGT	Q.....N.K	P.N...RV
Pyne IgG1	S-.K--G...	...SSRS..	..E.TA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q...I...N.K	P.N...RV
Pyne IgG2	S-.K--G.L.	...SCSRSI	..PQ.TA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q.....V.E	P.N...RV
Pyne IgG3	S-.K--G.L.	...SCSRS..	..PQ.TA.M.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q.....V.E	P.N...RV
Nala IgG	S-.K--G...	...SSRS..	..E.TA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q...I...N.K	P.N...RV
Rhro IgG	S-.K--G...	...SSRS..	..E.TA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..V...SS.LGT	Q...I...N.K	P.N...RV
Trfr IgG	S-.K--G...	...SSRS..	..E.TA.....	..D.....	..S.....	AL.....	..H...V.Q.....	..F...SS.LGT	Q...I...N.K	P.N...RV
Caja IgG	S-.K--N.D.	...T.CG--	..TTT.....	..D.....	..S.....	AL.....	..H...V.R.....	..T--S.AG.	Q.....V.T	PTGF.K.VPV
Saim IgG	S-.K--K.D.	...K.CG--	..TTT.....	..D.....	..S.....	AL.....	..H...V.Q.....	..T--S.AG.	Q.....T	P.GF.KNVPV
Sabo IgG	S-NK--K.D.	...T.CC--	..TTM.....	..D...Q...	..S.....	AL.....	..H...V.Q.....	..T--S.AG.	Q.....V.T	PT.F.K.LIV
Ceca IgG	S-NK--K.D.	...T.CC--N	..TTT.....	..D...Q...	..S.....	L.....	..H...V.Q.....	..T--S.AG.	Q.....V.T	PTNF.K.V.V
Ceal IgG	S-NK--K.D.	...T.CC--N	..TTT.....	..D...Q...	..S.....	L.....	..H...V.Q.....	..T--S.AG.	Q.....V.T	PTNF.K.V.V
Saap IgG	S-NK--K.D.	...T.CC--N	..TTT.....	..D...Q...	..S.....	L.....	..H...V.Q.....	..T--S.AG.	Q.....V.T	PTNF.K.V.V
Aona IgG	S-.K--A.D.	...T.CG--	..TTT.....	..D...Q...	..S.....	AL.....	..H...V.Q.....	..T--S.AG.	Q.....V.T	PTGF.K.L.V
Atge IgG	S-.K--K.D.	...T.CG--	..TTT.....	..D...Q...	..S.....	AL.....	..H...V.Q.....	..T--S.AG.	Q.....V.T	PTGF.K.QIV
Alpa IgG1	S-.K--K.D.	...T.CG--	..TTT.....	..D...Q...	..S.....	AL.....	..H...V.Q.....	..T--S.AGT	Q.....V.T	PTGF.K.QIV

CH2 Domain

Saap IgG APDF.A.... .L.....I .R.TGS.E..E...N.... G.E..T.K.S .K.E....YVL.Y. ...T.QK..H...

Saap IgG	.-----	-----	---TP...Q.	..L..S...M	SRS.V.....	IQ....S..A	E	N....QP..	F....	D.HR.Q	Q.TVFS....			.SL..S...
Aona IgG	.-----	-----	---TP...Q.	.IL..S...M	SRS.V.....	.HS...S..A	S...E	SS....P..	.A...F....	D..R.Q	Q.KVFS....			.SL..S...
Atge IgG	.-----	-----	---P...Q.	..L.....M	SRS.VT....	IQ....S..A	E	N....P..	F....	D.LR.Q	Q.NVFS....			.SL..S...
Alpa IgG1	.-----	-----	---SP...Q.	.IL.....M	SRSKVT.---	-----	-----E	N....E...	.A...F....	D.LR.Q	Q.NVFS....			.SL..S...
Alpa IgG2	.-----	-----	---SP...Q.	.IL.....M	SRSKVT.---	-----	-----E	N....E...	.A...F....	D.LR.Q	Q.NVFS....			.SL..S...
Pipi IgG	.-----	-----	---TP...Q.	..L..S...M	SRS.V.....	IQ....S..A	E	S....QP..	F....	D.HR.Q	Q.TIFS....	M.		.SL..S...
Pldo IgG	.-----	-----	---TP...Q.	..L..S...L	SRS.V.....	IQ....S..A		NT....P..	F....	D.HR.Q	Q.SVFS....			.SL..S...

The human IgG critical amino acid residues binding to complement, and effector cell receptors are shown. They are located at the CH2 and CH3 domains. FcγRn (IgG-Fc receptors) binding sites are highlighted in light blue, while yellow sites are those involved in Cq1 (complement component 1q) binding. Yellow asterisks correspond to sites involved in binding to both FcγRn and Cq1. FcRn (neonatal Fc receptor) binding sites are highlighted in dark blue. CH (Ig) domains and hinge region were determined using the EU numbering tool. Positive selected sites within CH2 and CH3 domains are indicated.

Immunoglobulin heavy constant gamma gene evolution is modulated by both the divergent and birth-and-death

evolutionary models

Supplementary material 7. Relaxed or intensified natural selection between reference and tested branches

Reference branch	Tested Branch	RELAX	P-value	
<i>Strepsirrhini</i>	<i>Platyrrhini</i>	Intensification (K = 1.90)	0.001	Proportion of positive selected sites increased
	<i>Cercopithecoid ea</i>	Intensification (K = 1.25)	0.273	-
	<i>Hominoidea</i>	Intensification (K = 1.51)	0.002	Proportion of both positive and negative selected sites increased. Omega negative values increased in tested branch
	<i>Gogo/Hosa/Pat r</i>	Intensification (K = 1.52)	0.015	Proportion of both positive and negative selected sites increased
	<i>Poab</i>	Relaxation (K = 0.88)	0.007	Proportion negative selected sites decreased. Small increase in positively selected sites

Selection intensity (relaxation or intensification) found by the RELAX method is shown for pairwise-lineage comparisons.