Indian Journal of Biochemistry & Biophysics
Vol. 61, December 2024, pp. 765-775
DOI: 10.56042/ijbb.v61i12.11875

Adaptive evolutionary research of the guanylate binding proteins family

Gao Yixuan[1], Gao Jian[3], Gao Yuqi[1], Guo Xingru[1], An Meng[1], Gu Limeng[1] & Xie Jiuxia[2]*

[1]School of Life Science, Qufu Normal University, Qufu-273 165, China

[2]Dongying Vocational College, Dongying-257 091, China

[3]School of Earth Science and Technology, China University of Petroleum (East China), Qingdao-266 500, China

Received 29 August 2024; *revised 10 September 2024*

Guanylate binding proteins (GBPs), encoded by the GBP gene family, are a multi-gene family belonging to the immune system. However, their evolutionary characteristics in ungulates are largely unknown. In this study, we identified 259 sequences of seven GBP genes from the unannotated genomes of 37 ungulate species and performed phylogenetic and selection pressure analyses on the GBP genes using comparative genomics approaches. Phylogenetic analysis indicated that GBP2 and GBP3 genes may have widespread homology. Evolutionary analysis revealed significant purifying selection acting on GBP genes during the evolution of ungulates, indicating a high degree of conservation. The GBP3 gene may play a more important role in antiviral defense. Comparative developmental analysis between even-toed ungulates and odd-toed ungulates showed that several members of the GBP gene family (GBP3, GBP4, GBP5, GBP6, and GBP7) exhibited stronger positive selection pressure in even-toed ungulates. The importance of this study lies in filling knowledge gaps, analyzing evolutionary mechanisms, exploring functional differences, and providing a foundation and guidance for future research in related fields. This study clarifies the phylogenetic and evolutionary characteristics of the GBP gene in ungulates, providing valuable data for further research on the evolutionary characteristics of ungulates.

Keywords: Guanylate binding proteins, Even-toed ungulates, Odd-toed ungulates, Phylogenetic and evolutionary characteristics

Guanylate binding proteins (GBPs) are an ancient protein family found in various organisms ranging from plants to humans, particularly prevalent in eukaryotes where migratory immune cells and interferon-induced immune systems are absent[1]. In eukaryotic cells, the presence of type I and type II interferons confers innate resistance against invading pathogens. When the levels of interferons increase in response to viral or bacterial infections, they induce the expression of certain interferon-stimulated genes[2], and GBP is one of the highly expressed genes. It is through its strong induction by type II interferon and its interaction with guanosine nucleotides that GBP was discovered and utilized as a marker of interferon responsiveness. GBPs belong to the dynamin superfamily, which includes myxoma resistance proteins, immunity-related GTPases, and the very large IFN-inducible GTPases[3]. Like other members of this superfamily, GBPs exhibit oligomerization, enzymatic activity, and membrane-interacting capabilities. They can self-activate by oligomerization and hydrolyze GTP in the absence of GTPase-activating proteins or guanine nucleotide exchange factors.

The domain architecture of GBPs comprises three parts. The N-terminal large GTPase domain contains three GTP-binding motifs: the phosphate-binding P-loop, the DxxG motif, and the RD motif. When GTP is bound, GBP dimers are stimulated in their enzymatic activity. Following the LG domain is a helical region, which has been subdivided into the middle domain and the C-terminal α12/13 domain. The middle domain consists of two triple-helical bundles that share a helix and extend from the LG domain. The α12/13 domain runs along the middle domain and returns to the LG domain, where these two domains engage in some mechanically important interactions. During the dimerization process of the LG domain, a conserved arginine residue in the phosphate-binding P-loop is positioned in the active site and stimulates the cleavage of the terminal phosphate group by stabilizing the transition state of the reaction[4,5].

GBPs mediate a wide range of innate immune functions, defending against intracellular pathogens such as bacteria, parasites, viruses, and protozoa, as well as participating in the activation of

*Correspondence:
E-mail: 1204617561@qq.com

inflammasomes. Multiple binding partners of GBPs have been discovered, along with various proposed mechanisms of action, depending on the organism, cell type, and pathogen.

Bacteria and Parasites: GBPs respond to various intracellular bacteria and parasites, such as Shigella flexneri, Salmonella enterica, Salmonella typhimurium, Legionella pneumophila, Neisseria meningitidis, Chlamydia trachomatis, Listeria monocytogenes, Mycobacterium bovis, Leishmania donovani, and Toxoplasma gondii[6]. GBPs can localize to vacuoles containing bacteria like Chlamydia trachomatis and Toxoplasma gondii and inhibit the growth of Chlamydia trachomatis by overexpressing GBP1 or GBP2. Additionally, GBP5 expression is increased in macrophages infected with Group B Streptococcus, promoting pyroptosis (a pro-inflammatory cell death mechanism), while silencing GBP genes has the opposite effec. Recently, GBP1 has been identified as a bona fide pattern recognition receptor for bacterial lipopolysaccharide (LPS), disrupting the integrity of bacterial outer membranes by clustering LPS. The link between GBP1 hydrolysis-induced production of GMP and inflammasome activation has also been discovered, updating the model of GBP-mediated defense[7].

Antiviral: Forced expression of hGBP-1 in Huh7 cells containing the replicon of hepatitis C virus inhibits replication by 40. Furthermore, hGBP-1 exhibits moderate antiviral activity against negative-strand RNA viruses, vesicular stomatitis virus, positive-strand RNA picornaviruses, and encephalomyocarditis virus in cultured cells. In recent years, GBPs have also been shown to have antiviral effects against classical swine fever virus, murine norovirus, Newcastle disease virus, dengue virus, herpes simplex virus-1, Kaposi's sarcoma-associated herpesvirus, hepatitis E virus, influenza A virus, human immunodeficiency virus, and respiratory syncytial virus[8].

Inflammatory Response: GBPs can regulate inflammasome activation through several different mechanisms. One mechanism for promoting canonical inflammasome activation is through the lytic destruction of microbial invaders. GBP-dependent microbial lysis leads to the release of pathogen-associated molecular patterns, including microbial DNA, which is sensed by the absence in melanoma 2 (AIM2) pattern recognition receptors.

Other Functions: GBPs can control intracellular replication. Using immune fluorescence techniques and new flow cytometry technology combined with single-cell fluorescence microscopy, Wallet, Pierre, and colleagues have shown that GBPs control the intracellular replication of Francisella novicida. The main pathways involved are: (1) direct inhibition of bacterial replication and induction of bacterial lysis in the cytoplasm, and (2) indirect elimination of the replication niche of Francisella novicida by inducing inflammatory AIM2 activation. These results have also been confirmed in a tularemia mouse mode. In addition, GBP-1 exhibits potent anti-angiogenic activity and direct anti-tumor effects on tumor cells, contributing to the inhibition of tumor growth. It also plays a protective role in the mucosa by preventing cell apoptosis, inhibiting angiogenesis, and regulating T cell receptor signaling.

Evolutionary studies have shown that most species possess large clusters of GBP genes encoding multiple paralogous homologs. However, due to the evolutionary life and death processes that GBP undergoes as part of the immune system, some genes may be deleted or retained in the genome. When maintained, genes may acquire new functions (neofunctionalization), subdivide functions (subfunctionalization), or even lose functionality and become pseudogenes. There can be significant variations among different genes. For example, the GBP3 gene seems to have emerged in Simiformes through gene duplication from GBP1 and acquired a new function involved in regulating caspase-4 activation[9]. As for GBP7, it likely arose from a duplication event involving GBP4 and appears to be specific to primates. Some GBP homologs are found in different species, while others may be unique to specific taxa. However, members of the GBP gene family are generally present in the genomes of mammals and most vertebrates, often classified as gene clusters encoding multiple GBP homologs. Previous functional and evolutionary studies have mainly focused on mouse and human GBPs. The human genome contains seven highly homologous members, named HuGBP-1 to HuGBP-7, all located within a single cluster on chromosome 1. Mice have 11 GBPs (referred to as mGBP-1 to mGBP-11), distributed in two clusters on two chromosomes.

However, more data is emerging, constantly increasing our understanding of this multifunctional protein family. Even-toed ungulates and odd-toed ungulates, as orders within the mammals, hold significant value in scientific research.

Even-toed ungulates have an equal number of toes on their limbs, arranged around the axis formed by the well-developed third and fourth digits, with the limb supported by the axis above these two digits, while the second and fifth digits are reduced and small, pointing backward. On the other hand, odd-toed ungulates have their limb axis passing through the third digit, and the first and fifth digits are generally lost. They typically function with only three toes on their fore and hind limbs, and in advanced equids, there may even be only one toe. Both even-toed and odd-toed ungulates are of great importance to humans, being essential components of livestock and agriculture for thousands of years, providing the primary sources of meat, milk, and fur products. Certain wild equids can also hybridize with domesticated counterparts, serving as scientific experimental subjects for breeding improvement.

Materials and Methods

Dataset and sequence alignment

Our dataset comprised lineages of Artiodactyla and Perissodactyla, including Bovidae (n = 11), Cervinae (n = 9), Muntiacinae (n = 3), Odocoileinae and Hydropotinae (n = 7), Perissodactyla (n = 6), and Homo sapiens. The genomic data used in this study were obtained from the National Center for Biotechnology Information (NCBI: accessed on 5 April 2024 https://www.ncbi.nlm. nih.gov/). We searched for annotated GBP genes in the NCBI database, downloaded the corresponding gene sequences, and retrieved the protein-coding sequences (CDSs) of GBPs from humans, domestic horses, and domestic cattle as seed sequences. These sequences were used as queries to investigate the unannotated genomes of Artiodactyla and Perissodactyla species.We use these sequences as query sequences and keywords to study the GBP gene sequences of unannotated genomes of odd hoofed and artiodactyl animals[10]. BLASTn[11] was used to search for GBP sequences in the unannotated genomes, with an e-value cutoff of 10-5. High-scoring fragments were selected to confirm significant matches with the genomes. The target sequences were obtained using the GeneWise website (accessed on 7 May 2024 https://www.ebi.ac.uk/Tools/psa/genewise). Finally, the MEGA X software[12] was used to remove protein sequences interrupted by premature stop codons, resulting in intact genes[13].

Evolutionary analysis

First, a phylogenetic analysis of the GBP gene family was conducted. The nucleotide sequences of Artiodactyla and Perissodactyla species were translated into amino acid sequences using MEGA X software. The resulting amino acid sequences were compared[14]. The aligned sequences were then exported for further analysis. To construct a maximum likelihood (ML) phylogenetic tree, the best model was determined using the "Find Best Model" program within the IQ-Tree software[15], which included 2000 bootstrap replicates and employed the GTR + T model. The most suitable nucleotide substitution model was determined using Model Test and the Bayesian Information Criterion (BIC). Ultimately, the GTR + T model was selected[16]. The phylogenetic tree was visually enhanced and annotated using the chipolt online tool.

The CodeML program in PAML4 was utilized to calculate the non-synonymous (dN) to synonymous (dS) substitution ratio (ω) for the GBP genes. Typically, the dN/dS ratio is used to determine the positive selection pressure on genes by measuring the ratio of non-synonymous to synonymous substitutions in homologous protein-coding gene sequences. Values of $\omega > 1$, $\omega < 1$, and $\omega = 1$ indicate positive selection, negative selection, and neutral evolution, respectively. Based on the sequence alignment and phylogenetic tree obtained from TimeTree (accessed on 7 May 2024 http://www.timetree.org/), we conducted selection pressure analysis on GBP genes in Artiodactyla and Perissodactyla species using the CodeML program in PAML 4, which employs a codon-based analysis within the ML framework. For the selection pressure analysis, the same selection pressure parameters were applied to all GBP genes, and LRT (likelihood ratio test) values and p-values were adjusted multiple times[17-19].

Next, we employed four models in CodeML to examine the adaptive changes and selection pressure of genes during evolution, as well as to identify the branches or sites that play a key role in species adaptation. We used the site model to investigate the characteristics of natural selection in GBPs genes of Perissodactyla and Artiodactyla. This model allows for the estimation of different evolutionary rates and selection pressures at different sites within the same gene sequence and uses maximum likelihood methods for statistical inference. The model calculates the ω value for each site to determine if it has undergone positive selection. We applied the M7 model and the

M8 model for our analysis. The M7 model serves as the null hypothesis, assuming that all sites have ω values between 0 and 1, indicating negative selection or neutral evolution. The M8 model serves as the alternative hypothesis, assuming that in addition to negative selection and neutral evolution, some sites have undergone positive selection with ω values greater than 1. Based on the results of the likelihood ratio test (LRT) between the models, we accept the alternative hypothesis (M8 model) as the positive selection model when the p-value of the LRT is significant. For models with significant results, we further use the Bayesian Empirical Bayes method[20-22] to identify the sites undergoing positive selection, considering sites with a posterior probability greater than 0.9 as potential positively selected sites[23].

The branch model is a model in the CODEML program used to detect adaptive evolution among different lineages through LRT[24]. It is mainly divided into the one-ratio (M0) model and the free-ratio (M1) model. The former assumes that the ω values of all lineages in the phylogenetic tree are equal, while the latter allows ω values to vary among lineages[25].

Results

Phylogenetic analysis of GBPs genes

We obtained 259 GBP-related genes in 37 mammalian species, including 250 intact genes and 9 pseudogenes. We constructed a phylogenetic tree using the obtained genes and enhanced and annotated it using the chipolt online tool (Figs 1 and 2A & B). We found that the GBP1 gene is relatively

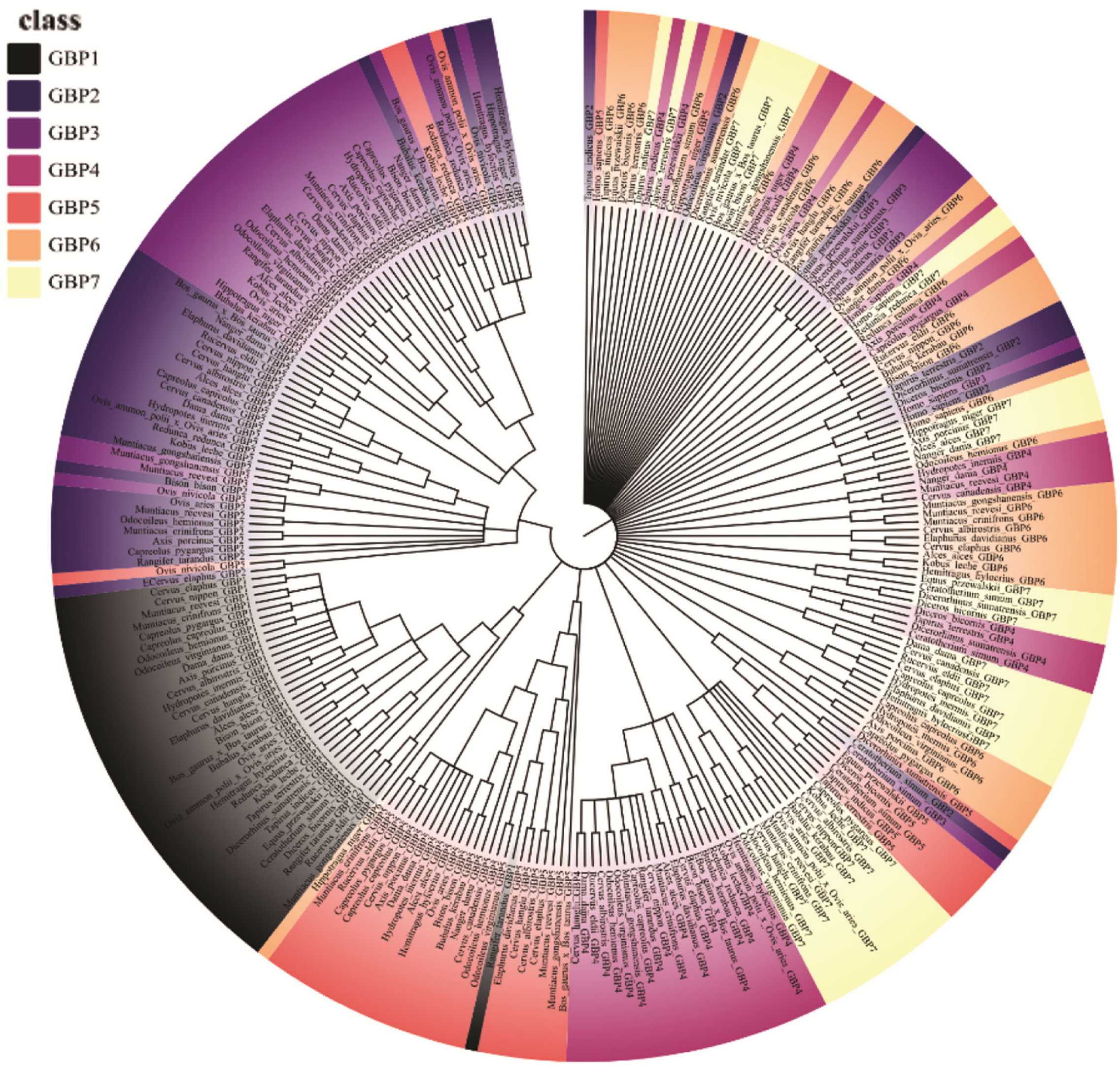

Fig. 1 — Phylogenetic tree of GBP genes in Artiodactyla

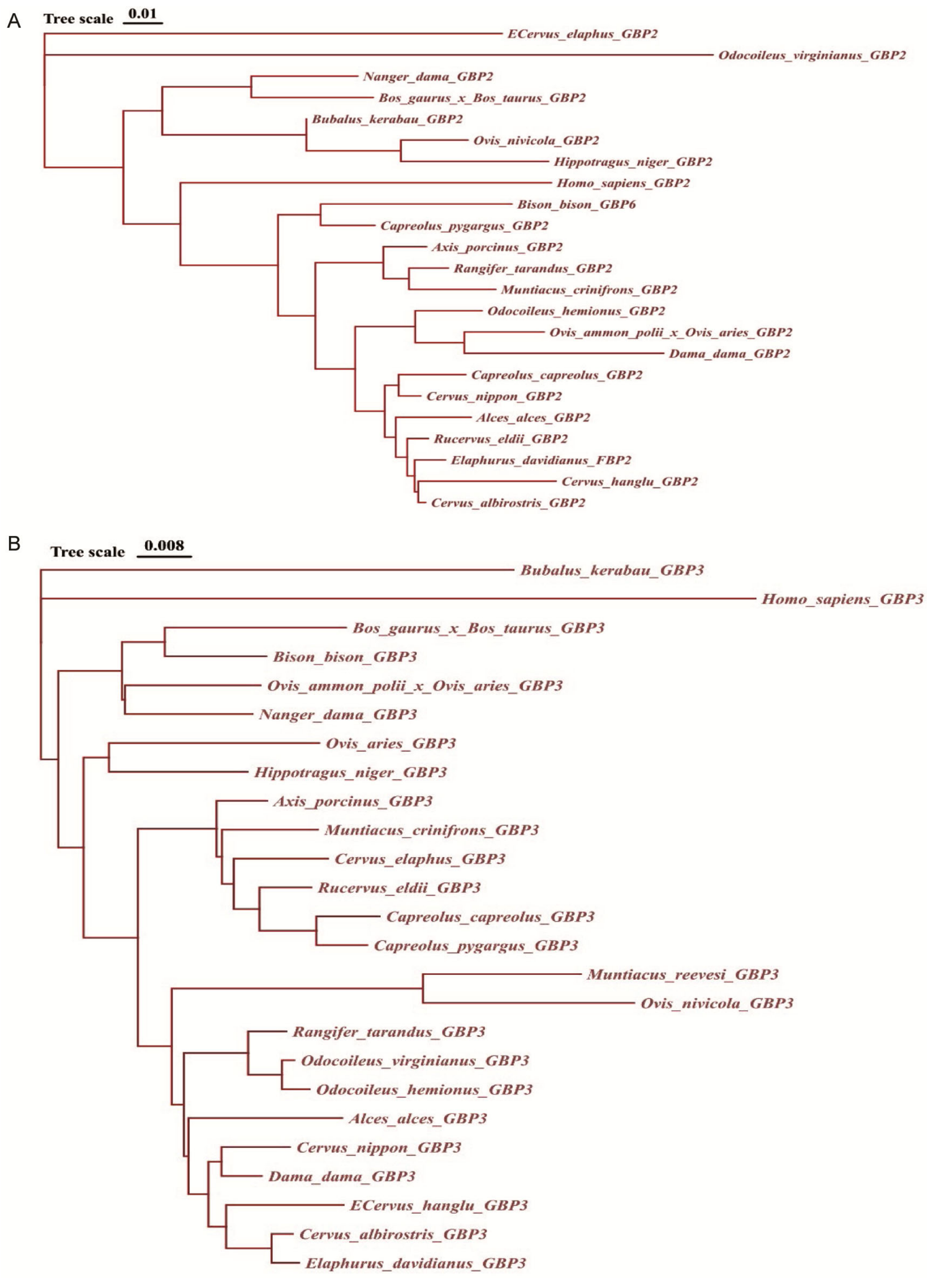

Fig. 2 — Phylogenetic tree of (A) GBP2; and (B) GBP3

independent and shows minimal duplication compared to other genes. Through the phylogenetic tree of GBPs genes in ungulates, we can infer that GBP2 and GBP3 genes may have broad homology. So we conducted a multi sequence comparison between the GBP2 gene and the GBP3 gene (Fig. 3).

Analysis of selection pressure

In the site model, a total of 281 candidate positively selected sites were identified (Table 1). In Bovidae, 7 genes had a total of 86 positively selected sites, with 52 being highly significant. In Cervinae, 7 genes had a total of 90 positively selected sites, with

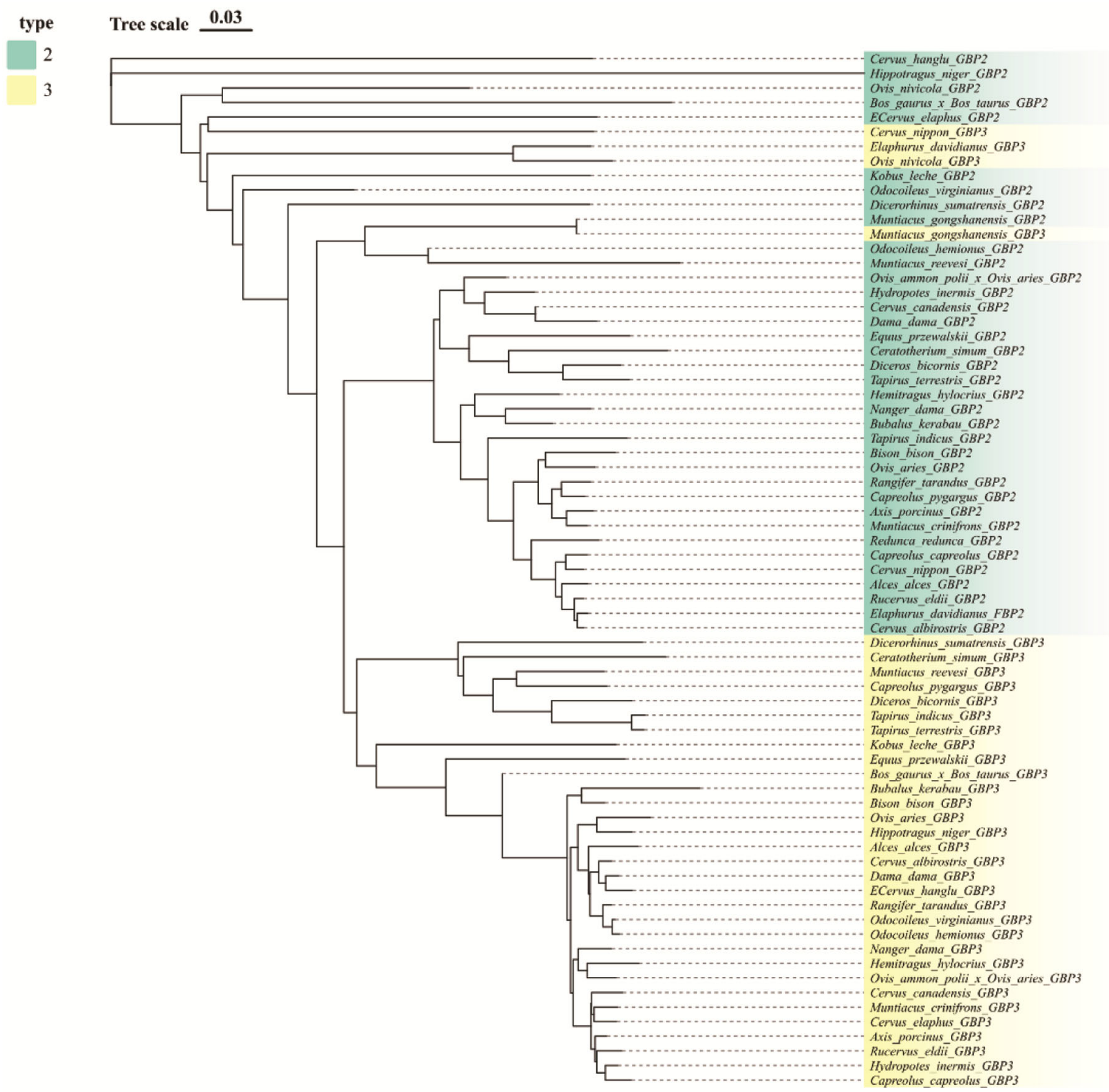

Fig. 3 — Multi sequence comparison of GBP2 and GBP3

13 being highly significant. In Muntiacinae, 7 genes had a total of 54 positively selected sites, with 16 being highly significant. In Odocoileinae and Hydropotinae, 7 genes had a total of 27 positively selected sites, with 8 being highly significant. In Perissodactyla, 7 genes had a total of 24 positively selected sites, with 10 being highly significant. Among them, GBP3 had the highest number of positively selected sites among all genes with 78 sites. Cervinae had the highest number of positively selected sites, while Perissodactyla had the fewest positively selected sites, indicating significant adaptive evolution in Cervinae.

Adaptive evolution may occur in specific lineages. To detect these branches, the branch model was used to examine the selection pressure on genes (Table 1). The ω values of the genes ranged from 0.0301 to 1.55104, indicating that these 7 genes experienced overall purifying selection pressure during the evolution of Artiodactyla.

Discussion

The guanylate-binding protein (GBP) family is a group of large GTPase molecules induced by interferon that plays an important role in the innate immune response against invading pathogens such as bacteria and viruses[26]. The human GBP family consists of 7 members[27]. In our study, GBP2, GBP3, GBP4, GBP6, and GBP7 in Redunca redunca were all found to be pseudogenes. One possible explanation is

Table 1 — Branch model testing positive selection in different branches of GBP genes

Taxa	Gene	lnL M1a	lnL M2a	Pvalues	ω0	ω	Order
Bovidae	GBP1	3304.31	3302.3693	0.69271267	0.83309	1.55104	Reduncinae
						0.80097	Caprinae
						0.87212	Antilopinae
						0.40808	Hippotraginae
						0.89416	Bovinae
	GBP2	3082.7367	3071.4542	2.0315E06	0.34504	0.34342	Reduncinae
						0.65080	Caprinae
						0.45141	Antilopinae
						0.79858	Hippotraginae
						0.33021	Bovinae
	GBP3	6291.9382	6260.3014	0.13763153	0.44991	0.44135	Reduncinae
						0.73814	Caprinae
						0.26301	Antilopinae
						0.44886	Hippotraginae
						0.34370	Bovinae
	GBP4	4981.6518	4971.1734	4.6973E06	0.67900	0.06269	Reduncinae
						0.83205	Caprinae
						0.41457	Antilopinae
						0.68834	Hippotraginae
						0.74451	Bovinae
	GBP5	2046.3827	2028.7949	3.0128E09	0.21521	0.37458	Reduncinae
						0.49604	Caprinae
						1.01362	Antilopinae
						0.06404	Hippotraginae
						0.40285	Bovinae
	GBP6	3377.2461	3369.983	0.00013821	0.637	0.37869	Reduncinae
						0.26199	Caprinae
						1.01677	Antilopinae
						0.42293	Hippotraginae
						0.89973	Bovinae
	GBP7	10501.269	10489.705	1.5154E06	0.35617	0.40654	Reduncinae
						0.76237	Caprinae
						0.26474	Antilopinae
						0.33358	Hippotraginae
						0.23104	Bovinae
Cervidae	GBP1	13055.924	13050.176	0.07418155	0.99184	0.86871	Cervus
						0.69985	Elaphurus davidianus
						1.46386	Axis Genus
						1.19709	Dama
	GBP2	6119.6536	6107.2435	6.2937E07	0.57577	0.81662	Cervus
						0.24119	Elaphurus davidianus
						0.47262	Axis Genus
						0.43671	Dama
	GBP3	7227.8102	7200.6295	0.13763153	0.56166	0.64620	Cervus
						0.89075	Elaphurus davidianus
						0.13057	Axis Genus
						0.21250	Dama
	GBP4	11120.742	11103.525	4.4084E09	0.62102	0.84857	Cervus
						0.28183	Elaphurus davidianus
						0.30127	Axis Genus
						1.24893	Dama
	GBP5	3782.9279	3775.3711	0.00010123	0.38054	0.64377	Cervus

(*Contd.*)

Table 1 — Branch model testing positive selection in different branches of GBP genes (*Contd.*)

Taxa	Gene	lnL M1a	lnL M2a	P values	ω0	ω	Order
						0.10089	Elaphurus davidianus
						0.36950	Axis Genus
						0.47556	Dama
	GBP6	2247.9428	2246.1403	0.05760565	0.77735	0.36229	Cervus
						0.19616	Elaphurus davidianus
						0.50242	Axis Genus
						0.15717	Dama
	GBP7	6941.1413	6941.1283	0.87172948	0.53695	0.30686	Cervus
						1.24334	Elaphurus davidianus
						0.45187	Axis Genus
						0.44578	Dama
Odocoileinae & Hydropotesinermis	GBP1	6942.9281	6935.1432	0.01625853	0.81311	1.19233	Rangifer tarandu
						0.51367	Odocoileus
						0.70268	Capreolus pygargus
						0.72909	Alces alce
						0.08872	Hydropot
	GBP2	1842.2118	1829.6368	5.3042E07	0.13872	0.38973	Rangifer tarandu
						0.03170	Odocoileus
						0.28171	Capreolus pygargus
						0.09895	Alces alce
						0.20518	Hydropot
	GBP3	4433.4109	4389.9483	0.13763153	0.49027	0.08037	Rangifer tarandu
						0.03010	Odocoileus
						1.07676	Capreolus pygargus
						0.16992	Alces alce
						0.35117	Hydropot
	GBP4	14113.376	14108.467	0.00172883	0.71374	0.77841	Rangifer tarandu
						0.78354	Odocoileus
						0.59111	Capreolus pygargus
						0.32420	Alces alce
						1.18061	Hydropot
	GBP5	2920.6299	2916.4313	0.00375787	0.40918	1.13662	Rangifer tarandu
						0.27800	Odocoileus
						0.20389	Capreolus pygargus
						0.62702	Alces alce
						0.83952	Hydropot
	GBP6	2252.4342	2246.8911	0.00086968	0.42445	0.80761	Rangifer tarandu
						1.39194	Odocoileus
						0.22071	Capreolus pygargus
						0.41006	Alces alce
						0.99022	Hydropot
	GBP7	7183.0359	7156.5018	3.2215E13	0.57636	0.96786	Rangifer tarandu
						0.27441	Odocoileus
						0.21747	Capreolus pygargus
						0.32003	Alces alce
						1.00841	Hydropot
Perissodactyla	GBP1	10433.166	10428.88	0.19919646	0.60180	0.71464	Equidae
						0.38913	Tapiridae
						0.26526	Rhinocerotidae
	GBP2	5254.8147	5248.6327	0.00043767	0.3782	0.58780	Equidae
						0.36018	Tapiridae
						0.26533	Rhinocerotidae

(*Contd.*)

Table 1 — Branch model testing positive selection in different branches of GBP genes (*Contd.*)

Taxa	Gene	lnL M1a	lnL M2a	P values	ω0	ω	Order
	GBP3	5908.9243	5896.5003	0.13763153	0.51322	0.94778	Equidae
						0.49075	Tapiridae
						0.45026	Rhinocerotidae
	GBP4	8595.4697	8581.8157	1.7349E07	0.43839	1.84858	Equidae
						0.33580	Tapiridae
						0.70486	Rhinocerotidae
	GBP5	3247.7282	3243.9389	0.00590628	0.3368	0.62953	Equidae
						0.20352	Tapiridae
						0.27135	Rhinocerotidae
	GBP6	2669.3351	2667.7825	0.07804265	0.17666	0.04621	Equidae
						0.29075	Tapiridae
						1.60826	Rhinocerotidae
	GBP7	5201.0122	5196.5505	0.00281536	0.40101	0.29630	Equidae
						0.38575	Tapiridae
						0.55078	Rhinocerotidae

that in this species, the functions of GBPs in immune or cellular processes have been replaced or compensated by other genes or proteins. This functional substitution or transfer may contribute to adaptation to specific immune or ecological environments. Studies have shown that GBP1 seems to be deleted in Arvicanthis niloticus. GBP1, GBP3, and GBP7 also appear to be deleted in primates. Dicerorhinus sumatrensis lacks GBP2, Ovis ammonpolii lacks GBP6, and Muntiacus gongshanensis lacks GBP4 and GBP7 genes, making them susceptible to viral infections, which may contribute to their endangered status. Especially for Dicerorhinus sumatrensis, it is listed as a critically endangered species.

To clarify the phylogenetic relationships of GBP gene family in Artiodactyla, we constructed a phylogenetic tree. In the phylogenetic tree, GBP2 and GBP3 clustered together, indicating that GBP2 and GBP3 are homologous genes that may have been generated from the same ancestral gene through gene duplication, recombination, or other mechanisms during evolution[28]. The GBP2, GBP3, and GBP5 genes of Ceratotherium simum clustered together, which may be due to gene family expansion events. During evolution, genes in a gene family may undergo gene duplication and functional divergence, leading to the formation of new genes[29]. These new genes may retain some common sequence features and cluster on adjacent branches in the evolutionary tree[30]. GBP5 and GBP3 of Ovis ammon polii x Ovis aries, Kobus leche, and Redunca redunca clustered together. GBP2 and GBP3 of Redunca redunca, Muntiacus reevesi, Muntiacus gongshanensis, Hemitragus hylocrius, and Hippotragus niger clustered together, suggesting that these genes have experienced similar evolutionary pressures in these species, leading to similar changes in their evolutionary rates and adaptability[31].

In this study, a large number of positively selected sites were detected in GBP genes, suggesting a coevolutionary phenomenon between artiodactyls and exogenous pathogenic microorganisms. This coevolution can be understood as an "arms race, " where the adaptive evolution of one species may exert selective pressure on the other species, driving their coevolution[32], supporting the "Red Queen hypothesis". Among the GBP genes, GBP3 had the highest number of positively selected sites, indicating its important functional and adaptive advantages in adaptive immune response and pathogen defense. Our findings revealed that Cervinae had the highest number of positively selected sites, while Perissodactyla had the fewest. This finding may suggest significant changes in adaptive evolution in Cervinae. On the contrary, Perissodactyla may be at a disadvantage in the arms race with viruses, making them more susceptible to viral infections. It is worth noting that Perissodactyla includes three families: Equidae, Tapiridae, and Rhinocerotidae, with the latter two being endangered species. Therefore, the lack of positively selected sites in Perissodactyla's GBP genes may contribute to their slow adaptive evolution in response to pathogens. Cervinae may face more adaptive pressures, leading to a higher number of positively selected sites and a better ability to outcompete viruses.

GBP genes in artiodactyls mainly underwent purifying selection during evolution, reflecting their functional conservation throughout animal development[33]. However, GBP4 showed more sites with ω > 1 compared to other GBP genes, suggesting its functional relevance. GBP4 may play an important role in immune adaptation and inflammatory response. Immune adaptation and inflammatory response are crucial factors for animal survival and adaptation, thus GBP4 may experience stricter selective pressure[34]. Additionally, GBP4 may play important adaptive roles in specific species, adapting to different environments and ecological pressures. In the Nangerdama, an antelope species, GBP5 and GBP6 genes showed ω > 1 values. This suggests that non-synonymous mutations are more common relative to synonymous mutations in this species, implying significant functional changes in GBP5 and GBP6 genes. Antelope species are known for their fast running and adaptation to arid environments shown that GBP6 is the most sensitive member of the GBP family in both in vivo and *In vitro* responses, suggesting these genes may play crucial roles in immune adaptation and inflammatory response, helping antelope species adapt to their unique ecological environments and providing protection against pathogens and other adversities[35].

Furthermore, we observed that multiple members of the GBP gene family (GBP3, GBP4, GBP5, GBP6, and GBP7) in artiodactyls exhibited stronger positive selection pressure, while in perissodactyls, only GBP4 and GBP6 experienced similar positive selection pressure. These observations may reflect differences in immune system function and adaptability between artiodactyls and perissodactyls. The differential ω values may indicate a more important role of the GBP gene family in immune adaptation in artiodactyls, while perissodactyls may prioritize other immune-related gene families or mechanisms.

Conclusion

In this study, we identified 259 novel GBP gene sequences in artiodactyls and analyzed the evolutionary characteristics of the artiodactyl GBP gene family, providing a phylogenetic basis for further functional research. Our findings suggest that GBP2 and GBP3 genes may have extensive homology. We also observed a significant number of positively selected sites in the GBP genes, with GBP3 having the highest number of positively selected sites, indicating its potentially important role in antiviral defense. In contrast, the GBP genes in Perissodactyla species had fewer positively selected sites, which may make them more susceptible to viral infections compared to other species. Additionally, our study revealed that the evolution of GBP genes primarily underwent purifying selection, indicating functional conservation throughout animal development. However, GBP4 gene may experience stricter selective pressure and play an important role in immune adaptation and inflammatory response.

However, this study has some limitations. Firstly, although the necessity of investigating other immune related genes and immune response pathways was mentioned, this aspect has not been explored in the research, which may lead to a less comprehensive understanding of the overall characteristics of the immune system in ungulates. In addition, the insufficient sample size is also a limiting factor, especially the lack of sufficient samples of ungulates, which may affect the accurate understanding of the evolution and immune adaptation of the GBP gene family in the entire ungulates. Therefore, despite limitations, this study lays the foundation for revealing the biological significance of the GBP gene and providing a deeper understanding of the immune system in ungulates. Future efforts will further enhance our understanding of the evolution and adaptability of the immune system in ungulates, and provide a more accurate foundation for related research.

Conflict of interest

All authors declare no conflict of interest.

References

1 Huang S, Meng Q, Maminska A & MacMicking JD, Cell-autonomous immunity by IFN-induced GBPs in animals and plants. *Curr Opin Immunol*, 60 (2019) 71.

2 Santos JC & Broz P, Sensing of invading pathogens by GBPs: At the crossroads between cell-autonomous and innate immunity. *J Leukoc Biol*, 104 (2018) 729.

3 Bae-Hoon K, Chee JD, Bradfield CJ, Park ES, Kumar P & MacMicking JD, Interferon-induced guanylate-binding proteins in inflammasome activation and host defense. *Nat Immunol*, 17 (2016) 481

4 Zhou W, Zhang Y, Huang PJJ, Ding J & Liu J, A DNAzyme requiring two different metal ions at two distinct sites. *Nucleic Acids Res*, 44 (2016) 354.

5 Wehner M & Herrmann C, Biochemical properties of the human guanylate binding protein 5 and a tumor-specific truncated splice variant. *FEBS J*, 277 (2010) 1597.

6 Tretina K, Park ES, Maminska A & MacMicking JD, Interferon-induced guanylate-binding proteins: Guardians of host defense in health and disease. *J Exp Med*, 216 (2019) 482.

7 Kutsch M & Coers J, Human guanylate binding proteins: nanomachines orchestrating host defense. *FEBS J*, 288 (2020) 5826.

8 Zhang R, Li Z, Tang YD, Su C & Zheng C, When human guanylate-binding proteins meet viral infections. *J Biomed Sci*, 28 (2021) 17.

9 Côrte-Real JV, Baldauf HM, Abrantes J & Esteves PJ, Evolution of the guanylate binding protein (GBP) genes: Emergence of GBP7 genes in primates and further acquisition of a unique GBP3 gene in simians. *Mol Immunol*, 132 (2021) 79.

10 Karimah N, Asri S & Nurhasanah a, A bioinformatic approach towards designing a human papillomavirus vaccine based on L1 capsid protein sequence of HPV45. *Indian J Biochem Biophys*, 59 (2022) 927.

11 Wenwen W, Binghui Z, Xia J, Junyi C & Shuhang W, Characteristics and Source of Dissolved Organic Matter in Lake Hulun, A Large Shallow Eutrophic Steppe Lake in Northern China. *Water*, 12 (2020) 953.

12 Rao TB, Kalagatur NK, Gali G, Boggula VK, Upadhyayula MK, Mangamuri UK & Poda S, Genomics and transgenic approaches for sheath blight resistance development in rice. *Indian J Biochem Biophys*, 60 (2023) 886.

13 Kumar S, Stecher G, Li M, Knyaz C & Tamura K, MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol*, 35 (2018) 1547.

14 Duong HV, , Repalli SK, Gupta P, Sreevathsa R, Yadava DK & Dash PK, Cloning and structural elucidation of a brassinosteroids biosynthetic gene (*Atdwarf*4) and genetic transformation of Indian mustard (*Brassica juncea* L.). *Indian J Biochem Biophys*, 59 (2022) 320.

15 Nguyen LT, Schmidt HA, von Haeseler A & Minh BQ, IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*, 32 (2015) 268.

16 Dash S, Ishani I, Lahiri D & Nag M, Evaluation of predictive machine learning models for drug repurposing against delta variant of SARS-CoV-2 spike protein. *Indian J Biochem Biophys*, 59 (2022) 879.

17 Wu X, Chen J, Wang X, Shang Y, Wei Q & Zhang H, Evolutionary Impacts of Pattern Recognition Receptor Genes on Carnivora Complex Habitat Stress Adaptation. *Animals*, 12 (2022) 3331.

18 Jyothi K, Sivaranjani V, Pavithra U, Jayavel S & Muthulakshmi L, Computational studies on new Leishmanial drug targets against Quercetin. *Indian J Biochem Biophys*, 59 (2022) 909.

19 Shamshad S and Rajagopal S, Current trends and advances of Quorum sensing inhibitors and their biotechnological applications. *Indian J Biochem Biophys*, 60 (2023) 255.

20 Dash S, Ishani I, Lahiri D & Nag M, Evaluation of predictive machine learning models for drug repurposing against delta variant of SARS-CoV-2 spike protein. *Indian J Biochem Biophys*, 59 (2022) 879.

21 Nisekhoto N, Roy D, Chettri Ai, Abdelgani-Baraka GA, Baishya B, Bidanchi RM, Bose M, Ahmed-Laskar S, Giri Ab, Buragohain P, Sailo L, Maurya K, Sonar I, Roy VK, Rao KRSS & Guruswami G, Computational toxicology and food safety assessment of Parkia timoriana phytoconstituents using quantitative structure-activity relationship (QSAR) modeling approaches. *Indian J Biochem Biophys*, 60 (2023) 896.

22 Apoorva Sharma, *In silico*-based studies on phytochemicals from native Indian plants as potential inhibitors of SARS-CoV-2. *Indian J Biochem Biophys*, 59 (2022) 653.

23 Zhang J, Nielsen R & Yang Z, Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol*, 22 (2005) 2472.

24 Man SM, David PE, Kuriakose T & Kanneganti TD, Interferon-inducible guanylate-binding proteins at the interface of cell-autonomous immunity and inflammasome activation. *J Leukoc Biol*, 101 (2017) 143.

25 Pandian CJ, Murugan NA, Rajendren SM & Jeyaraman J, Computer-Aided vaccine design for selected positive-sense single stranded RNA viruses. *Indian J Biochem Biophys*, 60 (2023) 281.

26 Daniel F, Barbara C, Marie-Charlotte D, Vesela E, Hironori B, Ambrosius PS & Eva-Maria F, Human GBP1 Differentially Targets Salmonella and Toxoplasma to License Recognition of Microbial Ligands and Caspase-Mediated Death. *Cell Rep*, 32 (2020) 108008.

27 Côrte-Real JV, Hanna-Mari B, Melo-Ferreira J, Abrantes J & Esteves PJ, Evolution of Guanylate Binding Protein (*GBP*) Genes in Muroid Rodents (Muridae and Cricetidae) Reveals an Outstanding Pattern of Gain and Loss. *Front Immunol*, 13 (2022) 752186.

28 Li L, Ji G, Ye C, Shu C, Zhang J & Liang C, PlantOrDB: a genome-wide ortholog database for land plants and green algae. *BMC Plant Biol*, 15 (2015) 161.

29 Ye Y, Liu Y, Li X, Wang G, Zhou Q, Chen Q & Tang H, An Evolutionary Analysis of B-Box Transcription Factors in Strawberry Reveals the Role of FaBBx28c1 in the Regulation of Flowering Time. *Int J Mol Sci*, 22 (2021) 11766.

30 Wang M, Manzoor MA, Wang X, Feng X, Zhao Y, He J & Cai Y, Comparative Genomic Analysis of *SAUR* Gene Family, Cloning and Functional Characterization of Two Genes (*PbrSAUR13* and *PbrSAUR52*) in *Pyrus bretschneideri*. *Int J Mol Sci*, 23 (2022) 7054.

31 Ning J, Zhou J, Wang H, Liu Y, Ahmad F, Feng X & Zhao L, Parallel Evolution of C-Type Lectin Domain Gene Family Sizes in Insect-Vectored Nematodes. *Front Plant Sci*, 13 (2022) 856826.

32 Miller SI, Ernst RK & Bader MW, LPS, TLR4 and infectious disease diversity. *Nat Rev Microbiol*, 3 (2005) 36.

33 Silva JCF, Ferreira MA, Carvalho TFM, Silva FF, Silveira SDA, Brommonschenkel SH & Fontes EPB, RLPredictiOme, a machine learning-derived method for high-throughput prediction of plant receptor-like proteins, reveals novel classes of transmembrane receptors. *Int J Mol Sci*, 23 (2022) 12176.

34 Tyrkalska SD, Candel S, Angosto D, Gómez-Abellán V, Martín-Sánchez F, García-Moreno D & Mulero V, Neutrophils mediate Salmonella Typhimurium clearance through the GBP4 inflammasome-dependent production of prostaglandins. *Nat Commun*, 7 (2016) 12077.

35 Côrte-Real JV, Baldauf HM, Melo FJ, Abrantes J & Esteves PJ, Evolution of Guanylate Binding Protein (GBP) Genes in Muroid Rodents (Muridae and Cricetidae) Reveals an Outstanding Pattern of Gain and Loss. *Front Immunol*, 13 (2022) 752186.