Aus dem Max von Pettenkofer-Institut, Lehrstuhl fürVirologie

Institut der Ludwig-Maximilians-Universität München

Vorstand: Prof. Dr. Oliver T. Keppler

Cross-species conservation of antiretroviral cell-autonomous innate immunity restriction factors

Dissertation

zum Erwerb des Doktorgrades der Naturwissenschaften

an der Medizinischen Fakultät der

Ludwig-Maximilians-Universität München

vorgelegt von

Luca Simon Schelle

aus

Heilbronn

Jahr

2024

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Betreuerin:	PD Dr. rer. nat. Hanna-Mari Baldauf
Zweitgutachterin:	Prof. Dr. rer. nat. Barbara Stecher-Letsch
Dekan:	Prof. Dr. med. Thomas Gudermann
Tag der mündlichen Prüfung:	27. März 2025

Affidavit

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Dekanat Medizinische Fakultät
Promotionsbüro

Affidavit

Schelle, Luca Simon

Surname, first name

I hereby declare, that the submitted thesis entitled:

Cross-species conservation of antiretroviral cell-autonomous innate immunity restriction factors

is my own work. I have only used the sources indicated and have not made unauthorized use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given.

I further declare that the dissertation presented here has not been submitted in the same or similar form to any other institution for the purpose of obtaining an academic degree.

München, 03.04.2025

place, date

Luca Schelle

Signature doctoral candidate

Confirmation of Congruency

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Dekanat Medizinische Fakultät
Promotionsbüro

Confirmation of congruency between printed and electronic version of the doctoral thesis

Doctoral candidate: Luca Simon Schelle

I hereby declare that the electronic version of the submitted thesis, entitled

Cross-species conservation of antiretroviral cell-autonomous innate immunity restriction factors

is congruent with the printed version both in content and format.

München, 03.04.2025
place, date

Luca Schelle
Signature doctoral candidate

Table of content

List of abbreviations

90K/LGALS3BP	Lectin Galactoside-Binding Soluble 3-Binding Protein
AIDS	acquired immunodeficiency syndrome
APOBEC	Apolipoprotein B mRNA Editing Catalytic Polypeptide-like
ASLV	Avian Sarcoma Leukosis Virus
BCA2/Rabring7	Breast cancer-associated gene/Rab7-interacting RING finger protein
BIV	Bovine Immunodeficiency Virus
BST-2/Tetherin/CD317	Bone Marrow Stromal Cell Antigen
CA	capsid protein
CH25H	Cholesterol 25-Hydroxylase
CIITA	Class II Transactivator
CNP	2',3'-cyclic-nucleotide 3'-phosphodiesterase
DAMP	danger-associated molecular pattern
EIAV	Equine Infectious Anemia Virus
ER	Endoplasmic reticulum
ERV	endogenous retroviruses
FAIDS	feline AIDS
FIV	Feline Immunodeficiency Virus
Fv	Friend virus susceptibility gene
G3BP1	GTPase-activating protein-(SH3 domain)-binding protein
GBP	Guanylate Binding Protein
HIV	human immunodeficiency viruses
HTLV	Human T-Lymphotropic Virus
HUSH	Human Silencing Hub
ICTV	International Committee on Taxonomy of Viruses
IFITM	Interferon-inducible transmembrane protein
IFN	interferon
IN	integrase
IR	immunity-related
ISG	interferon stimulated genes
JSRV	JaagSiekte Sheep Retrovirus
KO	Knock-out
LTR	long terminal repeat
MA	matrix protein
MARCH	Membrane-associated Ring Finger (C3HC4)
MLV	murine leukemia virus
MMTV	Mouse Mammary Tumor Virus
MX	Myxovirus Resistance
NC	nucleocapsid protein

PAMP	pathogen-associated molecular pattern
PBS	primer binding site
PFV	Prototype foamy virus
PIC	pre-integration complex
ppt	poly purine tract
PR	protease
PRR	pattern-recognition receptor
R5	CCR5
Rmcf	resistance to mink cell focus-forming virus
RRE	Rev response element
RT	reverse transcriptase
SAMHD	SAM And HD Domain Containing Deoxynucleoside Triphosphate Triphosphohydrolase
SERINC	Serine Incorporator
SIV	Simian Immunodeficiency Virus
SLFN	Schlafen
ss	Single-stranded
SU	surface envelope protein
TGN	Trans-Golgi network
TM	transmembrane envelope protein
TRIM	Tripartite Motif Containing
X4	CXCR4
ZAP	Zinc Finger CCCH-Type Antiviral Protein
ψ	packaging signal

List of Tables and Figures

List of Tables

List of Figures

List of publications

Thesis publications

This cumulative thesis comprises two publications:

1) Schelle L, Abrantes J, Baldauf H-M and Esteves PJ. Evolution of primate interferon-induced transmembrane proteins (IFITMs): a story of gain and loss with a differentiation into a canonical cluster and IFITM retrogenes. *Frontiers in Microbiology* (2023) 14:1213685. doi: 10.3389/fmicb.2023.1213685

2) Schelle L, Côrte-Real JV, Fayyaz S, del Pozo Ben A, Shnipova M, Petersen M, Lotke R, Menon B, Matzek D, Pfaff L, Pinheiro A, Marques JP, Melo-Ferreira J, Popper B, Esteves PJ, Sauter D, Abrantes J and Baldauf H-M. Evolutionary and functional characterization of lagomorph guanylate-binding proteins: a story of gain and loss and shedding light on expression, localization and innate immunity-related functions. *Frontiers in Immunology* (2024) 15:1303089. doi: 10.3389/fimmu.2024.1303089

Furthermore, I have presented parts of the work at several conferences (see conference list in CV).

Review articles

I have contributed to two review articles concerning R1) retroviral restriction factors and R2) cross-species conservation of guanylate binding proteins (R2) is attached in the Appendix). R2) is not part of the main components of this cumulative thesis and R1) is not included in this cumulative thesis.

R1) Kriesel, F, **Schelle L,** Baldauf, HM Same same but different - Antiviral factors interfering with the infectivity of HIV particles, *Microbes and Infection*. (2020) 22(9):416-422. doi: 10.1016/j.micinf.2020.05.009.

R2) Schelle L, Côrte-Real JV, Esteves PJ, Abrantes J, Baldauf HM. Functional cross-species conservation of guanylate-binding proteins in innate immunity. *Medical Microbiology and Immunology* (2022) 212(2):141-152 doi: 10.1007/s00430-022-00736-7

Thesis independent publications

The following publications are not part of the cumulative thesis and were conducted independently.

Stockinger P, **Schelle L**, Schober B, Buchholz PCF, Pleiss J, Nestl BM. Engineering of Thermostable β-Hydroxyacid Dehydrogenase for the Asymmetric Reduction of Imines. *ChemBioChem*. (2020) 21(24):3511-3514. doi:10.1002/cbic.202000526

Oberacker P, Stepper P, Bond DM, Höhn S, Focken J, Meyer V, **Schelle L**, Sugrue VJ, Jeunen GJ, Tim Moser T, Hore SR, von Meyenn F, Hipp K, Hore TA, Jurkowski TP. Bio-On-Magnetic-Beads (BOMB): Open platform for high-throughput nucleic acid extraction and manipulation. *PLoS Biolology*. (2019) 17(1):1–16. doi:10.1371/journal.pbio.3000107

Contributions to the publications

Contributions to publication 1)

Contributions: LS: Conceptualization; Data curation; Formal analysis; Writing – original draft; Writing – review & editing. **JA**: Funding acquisition; Writing – review & editing. **H-MB**: Supervision; Writing – review & editing; Funding acquisition. **PE**: Conceptualization; Writing – review & editing; Funding acquisition (reproduced from publication 1) see below).

Contributions to publication 2)

Contributions: LS and **JVCR** have contributed equally to this work and share first authorship. **LS**: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing (motifs, structure, expression (mRNA and protein), localization, IFN stimulation). **JVCR**: Conceptualization, Data curation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing (evolution, phylogeny, synteny) **SF**: Investigation, Formal analysis, Writing – review & editing. **AB**: Investigation, Writing – review & editing. **MS**: Investigation, Formal analysis, Writing – review & editing. **MP**: Investigation, Formal analysis, Writing – review & editing. **RL**: Formal analysis, Supervision, Writing – review & editing. **BM**: Investigation, Writing - review & editing. **DM**: Resources, Writing – review & editing. **LP**: Resources, Writing – review & editing. **AP**: Data curation, Writing – review & editing. **JM**: Data curation, Resources, Writing – review & editing. **JM-F**: Data curation, Resources, Writing – review & editing. **BP**: Resources, Supervision, Writing – review & editing. **PE**: Funding acquisition, Supervision, Writing – review & editing. **DS**: Funding acquisition, Supervision, Writing – review & editing. **JA**: Funding acquisition, Supervision, Writing – review & editing. **H-MB**: Conceptualization, Funding acquisition, Supervision, Writing – original draft, Writing – review & editing (reproduced and complemented from publication 2) see below).

Contributions to review R1) not included in this cumulative thesis

Contributions: FK: Literature research; Writing – review & editing. **LS**: Literature research; Writing – review & editing. **H-MB**: Writing and Figures – original draft Supervision; Writing – review & editing; Funding acquisition.

Contributions to review R2) included in the Appendix

Contributions: LS and **JVCR** contributed equally. **LS**: Conceptualization; Literature research; Writing and Figures – original draft (except of the parts: Evolution and conservation, GBPs in Teleosts); Writing – review & editing. **JVCR**: Conceptualization; Literature research; Writing – original draft (Evolution and conservation, GBPs in Teleosts); Writing – review & editing. **PE**: Funding acquisition. Writing – review & editing; **JA**: Funding acquisition; Writing – review & editing. **H-MB**: Supervision; Writing – review & editing; Funding acquisition.

1. Introduction

1.1 Retroviruses

According to the International Committee on Taxonomy of Viruses (ICTV), the family of *Retroviridae* can be divided into the subfamily of *Spumaretrovirinae,* which is not related to disease, and *Orthoretrovirinae*. *Orthoretrovirinae* comprises the genera of *Alpharetrovirus*, *Betaretrovirus*, *Gammaretrovirus*, *Deltaretrovirus*, *Epsilonretrovirus*, and *Lentivirus* (Lefkowitz *et al.*, 2018). Examples of *Orthoretrovirinae* and their related diseases are listed in **Table 1**.

Table 1: Examples of disease-causing *Orthoretrovirinae*. Compilation of examples of *Orthoretrovirinae* from all genera with associated disease(s).

Genus	Virus	Disease	References
Alpharetrovirus	ALV	B-cell lymphoma/leukaemia, erythroleukaemia, myeloid leukaemia	(Hulo *et al.*, 2011; Payne and Nair, 2012; Coffin *et al.*, 2021)
	RSV	Sarcomas	(Spencer and Groupé, 1962a, 1962b; Hulo *et al.*, 2011; Weiss and Vogt, 2011; Coffin *et al.*, 2021)
Betaretrovirus	MMTV	Mammary adenocarcinoma, rarely T-cell lymphoma	(Ross, 2010; Hulo *et al.*, 2011; Coffin *et al.*, 2021)
	JSRV	Ovine pulmonary adenomatosis	(Hofacre and Fan, 2010; Hulo *et al.*, 2011; Coffin *et al.*, 2021)
Gammaretrovirus	MLV	T-cell lymphoma/leukaemia, myeloid leukaemia, neurological disorders	(Fan, 1997; Münk *et al.*, 1997; Hulo *et al.*, 2011; Coffin *et al.*, 2021)
	FeLV	Aplastic anaemia, immunodeficiency syndrome, T-cell lymphoma, myeloid leukaemia	(Hulo *et al.*, 2011; Sykes and Hartmann, 2014; Coffin *et al.*, 2021)
Deltaretrovirus	HTLV	Adult T-cell leukemia, myelopathy/tropical spastic paraparesis, uveitis, infective dermatitis, chronic respiratory diseases, lymphadenitis	(Proietti *et al.*, 2005; Hulo *et al.*, 2011; Coffin *et al.*, 2021; Ramezani *et al.*, 2022)
	BLV	benign persistent B-cell lymphocytosis, rarely fatal adult lymphosarcom	(Willems *et al.*, 2000; Hulo *et al.*, 2011; Juliarena *et al.*, 2017; Coffin *et al.*, 2021)
Epsilonretrovirus	WDSV	Seasonal benign dermal sarcoma	(Rovnak and Quackenbush, 2010; Hulo *et al.*, 2011; Coffin *et al.*, 2021)

Genus	Virus	Disease	References
Lentivirus	FIV	Feline AIDS	(Burkhard and Dean, 2005; Hulo *et al.*, 2011; Liu, 2015; Coffin *et al.*, 2021)
	SIV	Simian AIDS	(Hulo *et al.*, 2011; Klatt, Silvestri and Hirsch, 2012; Coffin *et al.*, 2021; Jasinska, Apetrei and Pandrea, 2023)
	HIV	AIDS	(Hulo *et al.*, 2011; Deeks *et al.*, 2015; Coffin *et al.*, 2021)
	EIAV	equine infectious anaemia	(Leroux, Cadore and Montelaro, 2004; Hulo *et al.*, 2011; Cook, Leroux and Issel, 2013; Coffin *et al.*, 2021)

ALV: Avian Leukosis Virus; **RSV**: Rous sarcoma virus; **MMTV**: Mouse Mammary Tumor Virus; **JSRV**: JaagSiekte Sheep Retrovirus; **MLV**: Murine Leukemia Virus; **FeLV**: Feline leukemia virus; **HTLV**: Human T-Lymphotropic Virus; **BLV**: Bovine leukemia virus; **WDSV**: Walleye dermal sarcoma virus; **FIV**: Feline Immunodeficiency Virus; **SIV**: Simian Immunodeficiency Virus; **HIV**: Human Immunodeficiency Virus; **EIAV**: Equine Infectious Anemia Virus **AIDS**: acquired immunodeficiency syndrome

The disease-causing viruses human immunodeficiency virus (HIV) and murine leukemia virus (MLV) are among the most studied retroviruses in humans and mice, respectively: MLV, a simple retrovirus with a recently discovered accessory protein, of the genus of *Gammaretrovirus*, and HIV, a complex retrovirus of the genus of *Lentivirus*, are the prototypes of their genera.

Aside from exogenous retroviruses, integration into the germline can give rise to heritable proviruses, also known as endogenous retroviruses (ERVs), which constitute significant portions of vertebrate genomes (Hayward, 2017).

1.1.1 Prototypic gammaretrovirus: MLV

The genomic proviral structure of MLV is about 8.3 kb and comprises the following protein coding regions: group-specific antigens (*gag)*, polymerase (*pol)*, and envelope (*env*). Translation of the Gag-Pol polyprotein relies on readthrough of the Gag termination codon, whereas Env is translated from a spliced mRNA. Translation can start at an upstream start codon, resulting in a larger, glycosylated glyco-Gag. Further, the provirus contains long terminal repeats (LTRs) with U5, R and U3 elements and a primer binding site (PBS), packaging signal (ψ) and a polypurine tract (ppt). Env encodes the glycoproteins: surface envelope protein (SU, 70 kDa) and transmembrane envelope protein (TM, 15 kDa). Gag codes for matrix protein (MA,15 kDa), p12 (12 kDa), capsid protein (CA, 30 kDa), nucleocapsid protein (NC, 10 kDa), and glycol-Gag (80 kDa, accessory protein, further proteolytically processed). Pol encodes protease (PR, 14 kDa),

reverse transcriptase (RT, 80 kDa) and integrase (IN, 46 kDa) (**Figure 1A**) (reviewed in (Rein, 2011; Coffin *et al.*, 2021)).

Mature MLV particles are approximately 80-100 nm in size, are enveloped with trimeric Env complexes, and contain two copies of (+)-single-stranded (ss)RNA as genome (reviewed in (Rein, 2011; Coffin *et al.*, 2021)). The virion structure of MLV is depicted in **Figure 1 A**.

The MLV replication cycle includes the following steps: 1) Binding of SU to mCAT1 by the Env trimer leads to 2) fusion of TM with the host cell. 3) In the cytoplasm, reverse transcription by RT occurs along with microtubular transport towards the nucleus. 4) MLV double-stranded DNA pre-integration complex (PIC) can only enter the nucleus of dividing cells after nuclear lamina breakdown 5) followed by integration, which is facilitated by IN. 6) The provirus is transcribed, hijacking the host cell machinery, followed by 7) translation of the polyproteins (Gag, glyco-Gag, Gag-Pol, Env). Unlike the cytoplasmic proteins, Env and glyco-Gag mature in the secretory pathway (Env: proteolytic processing, disulfide bond SU-TM, glycosylation; glyco-Gag: glycosylation). 8) Virion budding and packaging of two copies of (+)-ssRNA as genome occur at the plasma membrane. 9) Release from the plasma membrane is facilitated by the ESCRT machinery and proteolytic viral maturation occurs in the released virions (reviewed in (Rein, 2011; Coffin *et al.*, 2021)). The MLV replication cycle is depicted in **Figure 2 A**.

MLV causes T cell lymphoma a few months after latency (Fan, 1997) and additionally encephalomyelopathy in mice (Münk *et al.*, 1997).

1.1.2 Prototypic Lentivirus: HIV-1

The genomic proviral structure of HIV-1 is about 9.3 kb and contains, in comparison to MLV, additionally the accessory genes *vif*, *vpr*, *vpu* and *nef*, as well as the regulatory genes *tat* and *rev*. The HIV provirus yields multiple transcripts in comparison to MLV (only genomic RNA and env transcript): the unspliced viral genomic RNA transcript (containing *gag-pol*, where *pol* genes are expressed as a result of a ribosomal frameshift) and spliced transcripts for all accessory and regulatory proteins except Vpu, which shares a spliced transcript with Env. HIV-1 also contains LTRs, PBS, ψ, ppt plus additionally a Rev response element (RRE). *Gag* encodes MA (17 kDa), CA (24 kDa), NC (7 kDa); p6 (budding protein, 6 kDa). PR (12 kDa) RT (66 kDa/51 kDa), IN (32 kDa) are translated from *pol*, whereas *env* codes for SU (gp120, 120 kDa) and TM (gp41, 41 kDa). The accessory proteins Vif, Vpr, Vpu, and Nef as well as the regulatory proteins Tat and Rev are translated from their respective transcripts (**Figure 1 B**) (reviewed in (Votteler and Schubert, 2008; Coffin *et al.*, 2021)).

Mature HIV-1 particles are approximately 120 nm in size, have an irregular icosahedral (cone-shaped) core, are enveloped with trimeric Env complexes, and contain two copies

of (+)-ssRNA as genome (reviewed in (Pornillos and Ganser-Pornillos, 2019; Coffin *et al.*, 2021). The virion structure of HIV-1 is depicted in **Figure 1 B**.

The replication cycle of HIV-1 is very similar to that of MLV but differs in some crucial details. In addition, recent technical advances have changed the textbook understanding of HIV nuclear import localization and reverse transcription timing (Burdick *et al.*, 2020; Dharan *et al.*, 2020; Francis *et al.*, 2020; Selyutina *et al.*, 2020; Müller *et al.*, 2021; Zila *et al.*, 2021; Xue *et al.*, 2023): In comparison to MLV, 1) HIV requires CD4 as a receptor and CXCR4 (X4) or CCR5 (R5) as co-receptors; 3) Reverse transcription begins in the cytoplasm in intact capsids, and (+)-strand synthesis is completed in the nucleus after import. 4) After microtubular transport, the intact cone-shaped capsid enters the nucleus through nuclear pores and accumulates in nuclear speckles, where it collapses and releases PICs for integration. 7) Additional translation of accessory and regulatory proteins occurs in the cytoplasm that influences viral replication and pathogenesis (replication cycle reviewed in (Freed, 2015; Ramdas *et al.*, 2020; Coffin *et al.*, 2021) and recent advances (Steps 3) and 4)) in (Dharan and Campbell, 2022; Muller *et al.*, 2022)). The replication cycle of HIV-1 is depicted in **Figure 2 B**.

HIV is the causative agent of the acquired immune deficiency syndrome (AIDS). The pathogenesis of HIV can be divided into the eclipse phase, the acute phase, and the chronic phase, which ultimately leads to AIDS (reviewed in (Deeks *et al.*, 2015)).

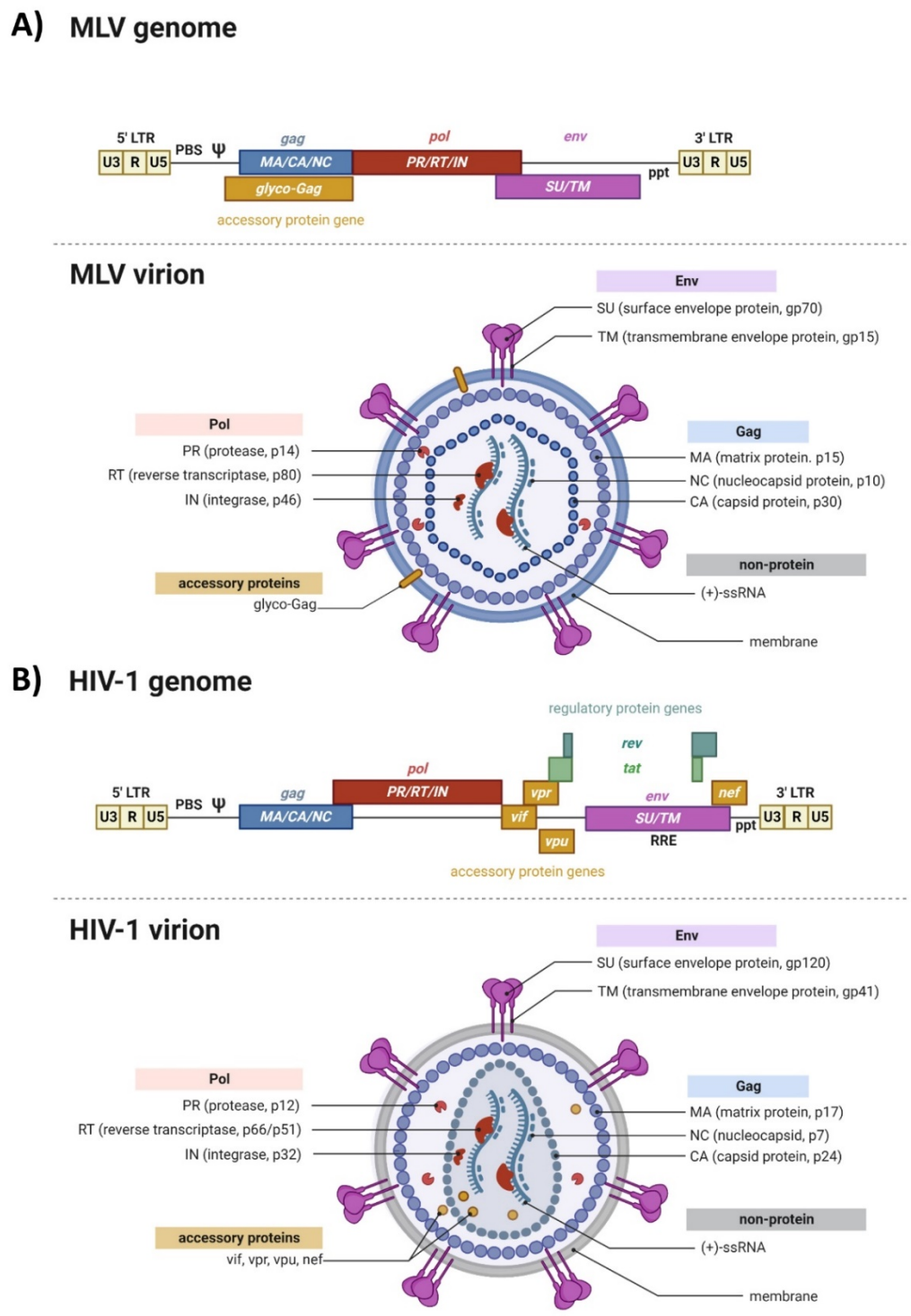

Figure 1: Retroviral provirus and virion structure. Shown are the genomic organization as well as a schematic of the virion structure for **A)** MLV and **B)** HIV. A detailed description can be found in the main text. **HIV**: Human Immunodeficiency Virus; **MLV**: Murine Leukemia Virus **Gag**: group-specific antigen, **Pol**: polymerase, **Env**: envelope, **LTR**: long terminal repeat, **PBS**: primer binding site, **ψ**: packaging, **RRE**: Rev response element, **ppt**: polypurine tract **ss**: single-stranded. *(Modified, completed and revised from "HIV genome and structure" by BioRender.com (2023))

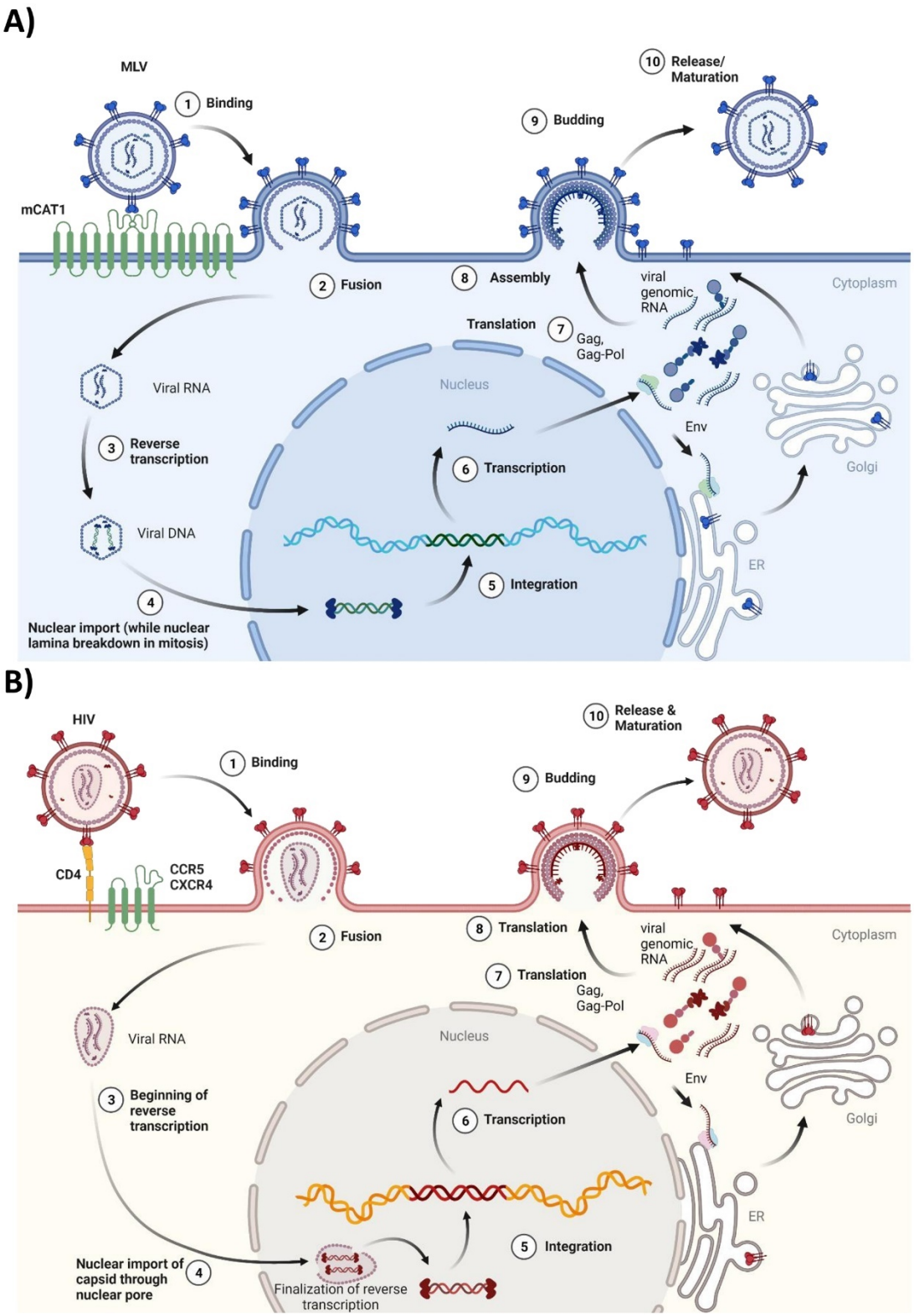

Figure 2: Retroviral replication cycle. Shown is the replication cycle for **A)** MLV and **B)** HIV**.** A detailed description can be found in the main text**. HIV:** Human Immunodeficiency Virus, **MLV**: Murine Leukemia Virus, **Gag:** group-specific antigen, **Pol**: polymerase, **Env**: envelope, **ER**: Endoplasmic reticulum *(Modified, completed and revised from ”HIV replication cycle” by BioRender.com (2023))

1.2 Innate immune system

The immune system can be divided into innate and adaptive immunity. In our studies, we investigated the retroviral restriction factors of the cell-autonomous innate immunity, which are part of the innate immune response (**Figure 3**).

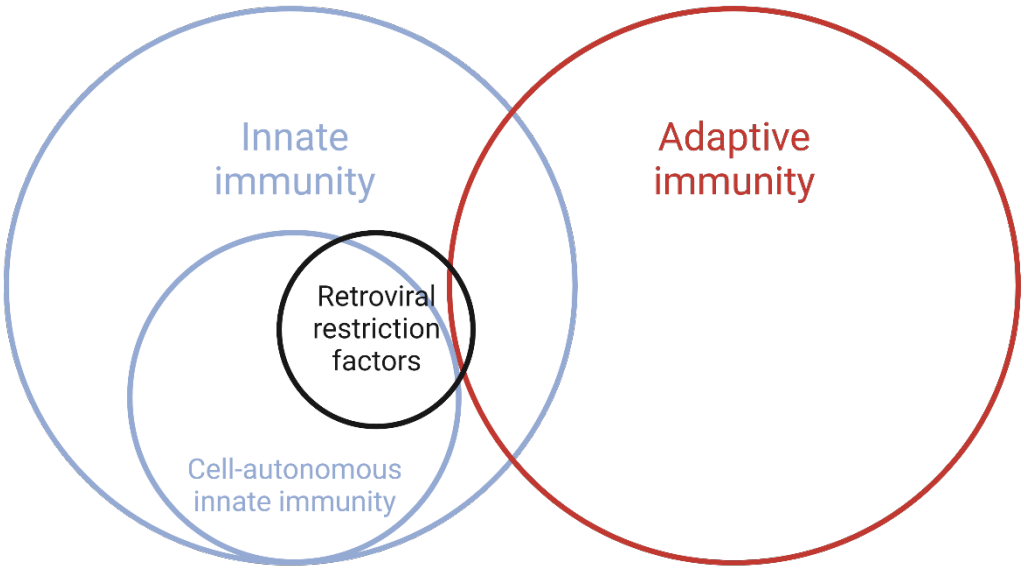

Figure 3: Schematic representation of the mammal immune system. Shown is the relationship of retroviral restriction factors within the immune system. *Created with BioRender.com.

1.2.1 Canonical innate immunity

Sensing of pathogen-associated molecular patterns (PAMPs) or indirect sensing of danger-associated molecular patterns (DAMPs) is conducted by pattern-recognition receptors (PRRs). Upon recognition of molecular patterns, signaling cascades lead to the production of pro-inflammatory cytokines and molecules and three types of interferons (IFNI, IFNII, IFNIII). These molecules, which act in an auto- and paracrine manner, lead to a cellular and humoral innate immune response, accompanied by an inflammatory response initiating cell recruitment and local inflammation and the expression of a variety of innate immunity-related genes, including interferon-stimulated genes (ISGs) (reviewed in (Akira, Uematsu and Takeuchi, 2006; Turvey and Broide, 2010; Riera Romo, Pérez-Martínez and Castillo Ferrer, 2016; Marshall *et al.*, 2018; Mantovani and Garlanda, 2023)). Many cells are involved in the innate immune response, including phagocytes (macrophages and neutrophils), dendritic cells, mast cells, basophils, eosinophils, natural killer cells, and innate lymphoid cells (reviewed in (Riera Romo, Pérez-Martínez and Castillo Ferrer, 2016; Marshall *et al.*, 2018)). This mainly recapitulates the activated innate immune response. The constitutive innate immune response includes physical

barriers or mechanisms such as autophagy and metabolite-mediated inhibition (reviewed in (Paludan *et al.*, 2021)).

ISGs control viral, bacterial, and parasitic infections by directly targeting pathways and functions required during the pathogen replication cycles or by enhancing pathogen detection. In addition, ISGs encode for proapoptotic proteins that induce cell death under certain conditions (reviewed in (Schneider, Chevillotte and Rice, 2014).

1.2.2 Cell-autonomous innate immunity

The host cytosol is not only surveilled by PRRs for the canonical innate immune response, but also contains a number of antipathogenic mechanisms, referred to as cell-autonomous immunity. In particular, cell-autonomous innate immunity intrinsically protects immune and non-immune cells from infection with pathogens (Randow, MacMicking and James, 2013; Moretti and Blander, 2017; Wein and Sorek, 2022). This cell-autonomous immunity includes retroviral restriction factors from constitutive innate immune response and/or activated innate immune response, e.g. ISGs (reviewed in (Paludan *et al.*, 2021)).

1.3 Retroviral restriction factors

1.3.1 Retroviral restriction factors in general

Retroviral restriction factors are part of the innate immunity (as described above) and form a multilayered defense that synergistically inhibits infection by retroviruses through various mechanisms covering all steps of the viral replication cycle. A compilation of described retroviral restriction factors (name, restricted step and known restricted retroviruses) is summarized in **Table 2**.

The retroviral restriction factors interferon-inducible transmembrane proteins (IFITMs) and guanylate binding proteins (GBPs) were integral part in this thesis and therefore further elucidated (elucidated in **1.6 Objectives of the thesis**).

Table 2: Compilation of retroviral restriction factors:

Replication step	Restriction factor	Affected retrovirus	Viral counter	References
Entry	SERINC3/5	HIV, SIV, EIAV, MLV	Nef, Vpu, S2, glyco-Gag	(Rosa *et al.*, 2015; Usami, Wu and Göttlinger, 2015; Timilsina *et al.*, 2020)
	IFITM1/2/3	HIV, MLV, JSRV	Vpr, glyco-Gag	(Lu *et al.*, 2011; Li *et al.*, 2013; Tartour *et al.*, 2014; Yu *et al.*, 2015; Smith *et al.*, 2019)
	CH25H	HIV		(S.-Y. Liu *et al.*, 2013)
	Visfatin	HIV (only R5 tropic)		(Bergh *et al.*, 2009; Van den Bergh *et al.*, 2012)
	Fv4	MLV		(Taylor, Gao and Sanders, 2001; Takeda and Matano, 2007)
	Rmcf, Rmcf2	Nonecotropic MLVs		(Jung *et al.*, 2002; Wu, Yan and Kozak, 2005)
Post entry events and reverse transcription	TRIM5a	HIV, SIV, MLV		(Stremlau *et al.*, 2004; Sebastian and Luban, 2005; Ganser-Pornillos and Pornillos, 2019)
	Fv1	MLV, EIAV, FFV		(Pincus, Hartley and Rowe, 1971; Yap *et al.*, 2014)
	APOBEC3 members	HIV, SIVs, MLV, MMTV, ERV	Vif, glyco-Gag	(Harris and Liddament, 2004; Esnault *et al.*, 2008; Henriet *et al.*, 2009; Stavrou *et al.*, 2014; Salas-Briceno, Zhao and Ross, 2020; Uriu *et al.*, 2021; Ajoge *et al.*, 2023)
	SAMHD1	HIV, SIV, MLV, EIAV, BIV, FIV	HIV-2/SIV Vpx, SIV Vpr	(Laguette *et al.*, 2011; Lahouassa *et al.*, 2012; Lim *et al.*, 2012; Gramberg *et al.*, 2013; White *et al.*, 2013; Baldauf *et al.*, 2017)
	P21	HIV, SIV		(Zhang, Scadden and Crumpacker, 2007; Allouch *et al.*, 2013; Shi *et al.*, 2018)
	MX2/MXB	HIV, SIV, MLV, EIAV, FIV		(Goujon *et al.*, 2013; Kane *et al.*, 2013; Z. Liu *et al.*, 2013)
Proviral transcription and translation	TRIM28	HIV, MLV, PFV, ERV		(Wolf and Goff, 2007; Rowe *et al.*, 2010; Allouch *et al.*, 2011; Yuan *et al.*, 2021)
	G3BP1	only HIV		(Cobos Jiménez *et al.*, 2015)
	HUSH complex	HIV, SIV	Vpx, Vpr	(Chougui *et al.*, 2018; Yurkovetskiy *et al.*, 2018)
	CIITA	HIV, HTLV		(Graziano *et al.*, 2018; Forlani *et al.*, 2019)
	TRIM22	HIV		(Barr, Smiley and Bushman, 2008; Singh *et al.*, 2011; Turrini *et al.*, 2015; Graziano *et al.*, 2018)
	ZAP	HIV, MLV, HTLV		(Gao, Guo and Goff, 2002; Zhu *et al.*, 2011; Miyazato *et al.*, 2019)

Replication step	Restriction factor	Affected retrovirus	Viral counter	References
	SLFN11	HIV, EIAV, MLV, FIV, PFV		(Li *et al.*, 2012; Lin *et al.*, 2016; Stabell *et al.*, 2016; Guo *et al.*, 2021)
Env translation, processing and trafficking	GBP2/5	HIV, MLV	Vpu	(Krapp *et al.*, 2016; Braun *et al.*, 2019)
	MARCH1/2/8	HIV, SIV, MLV, MMTV		(Tada *et al.*, 2015; Zhang, Lu and Liu, 2018; Zhang *et al.*, 2020; Lun *et al.*, 2021; Umthong *et al.*, 2021)
	90K/LGALS3BP	HIV, SIV		(Lodermeyer *et al.*, 2013, 2018)
	Mannose Receptor	HIV	Vpr, Nef	(Vigerust, Egan and Shepherd, 2005; Lubow *et al.*, 2020)
Assembly, budding and release	CNP	HIV, SIV		(Wilson *et al.*, 2012)
	ISG15	HIV, ASLV		(Okumura *et al.*, 2006; Pincetic *et al.*, 2010; Kuang, Seo and Leis, 2011)
	BCA2/Rabring7	HIV, SIV		(Miyakawa *et al.*, 2009; Nityanandam and Serra-Moreno, 2014; Colomer-Lluch and Serra-Moreno, 2017)
	BST-2/Tetherin/ CD317	HIV, MLV, FIV, HTLV	HIV-1 Vpu, HIV-2 Env, SIV Nef	(Neil, Zang and Bieniasz, 2008; Van Damme *et al.*, 2008; Jouvenet *et al.*, 2009; Miyagi *et al.*, 2009; Jolly, Booth and Neil, 2010; Dietrich *et al.*, 2011)

HIV: Human Immunodeficiency Virus; **MLV**: Murine Leukemia Virus; **SIV**: Simian Immunodeficiency Virus; **EIAV**: Equine Infectious Anemia Virus; **FIV**: Feline Immunodeficiency Virus; **FFV**: Feline foamy virus; **HTLV**: Human T-Lymphotropic Virus, **MMTV**: Mouse Mammary Tumor Virus; **ASLV**: Avian Sarcoma Leukosis Virus; **PFV**: Prototype foamy virus; **ERV**: Endogenous RetroVirus; **BIV**: Bovine Immunodeficiency Virus; **JSRV**: JaagSiekte Sheep Retrovirus; **SERINC**: Serine Incorporator; **IFITM**: Interferon-inducible transmembrane protein; **CH25H**: Cholesterol 25-Hydroxylase; **Rmcf**: resistance to mink cell focus-forming virus; **TRIM**: Tripartite Motif Containing; **Fv**: Friend virus susceptibility gene; **APOBEC**: Apolipoprotein B mRNA Editing Catalytic Polypeptide-like; **SAMHD**: SAM And HD Domain Containing Deoxynucleoside Triphosphate Triphosphohydrolase; **MX**: Myxovirus Resistance; **G3BP1**: GTPase-activating protein-(SH3 domain)-binding protein; **HUSH**: Human Silencing Hub; **CIITA**: Class II Transactivator; **ZAP**: Zinc Finger CCCH-Type Antiviral Protein; **SLFN**: Schlafen; **GBP**: Guanylate Binding Protein; **MARCH**: Membrane-associated Ring Finger (C3HC4); **90K/LGALS3BP**: Lectin Galactoside-Binding Soluble 3-Binding Protein; **CNP**: 2',3'-cyclic-nucleotide 3'-phosphodiesterase; **BCA2/Rabring7**: Breast cancer-associated gene/Rab7-interacting RING finger protein; **BST-2/Tetherin**: Bone Marrow Stromal Cell Antigen

1.3.2 IFITMs

IFITMs are transmembrane proteins with approximately 130 amino acids in size. IFITMs contain five topological domains: the N-terminal domain, the intramembrane domain 1 and the conserved intracellular loop (jointly the CD255 domain), the intramembrane domain 2 and the C-terminal domain (Bailey *et al.*, 2013, 2014). The topology of IFITMs in membranes is not fully understood. It might differ between the paralogs and might be depending on the characteristics of the membranes (reviewed in (Bailey *et al.*, 2014)). The first discovered *IFITMs*, namely *IFITM1*, *IFITM2* and *IFITM3*, were identified as ISGs (Friedman *et al.*, 1984). IFITMs are an ancient protein family with homologs in fish, amphibians, reptiles, birds, monotremes, marsupials, and mammals (Hickford *et al.*, 2012). IFITMs can be classified, according to their phylogeny, into three clades: immunity-related (IR-) IFITMs comprising IFITM1, IFITM2 and IFITM3 and further IFITM5 and IFITM10 (Zhang *et al.*, 2012).

The different IFITMs possess different roles and functions. IFITM5 has acquired a Ca^{2+} binding site and is involved in bone mineralization and osteoblast function (Hanagata *et al.*, 2011; Hedjazi *et al.*, 2022). The exact role and function of IFITM10 is unknown but it has been linked gastric cancer (Zhao *et al.*, 2019). The IR-IFITMs, as the name implies, are playing a role in immune responses. They act as viral restriction factors in innate immune response against RNA and DNA viruses with several observed and proposed modes of action (reviewed in (Diamond and Farzan, 2013; Bailey *et al.*, 2014; Liao *et al.*, 2019; Zhao *et al.*, 2019; Friedlová *et al.*, 2022; Gómez-Herranz, Taylor and Sloan, 2023)). Recently, additional functions in other areas of immunity like adaptive immunity have been found (reviewed in (Yánez, Ross and Crompton, 2020; Friedlová *et al.*, 2022; Gómez-Herranz, Taylor and Sloan, 2023)).

In context of innate immunity, IFITMs act as viral restriction factors, including retroviruses. They are restricting RNA and DNA viruses from several virus families, e.g. HIV-1, influenza A/B virus, West Nile virus, Dengue virus, Hepatitis C virus, Vesicular stomatitis virus, Rabies virus, Hantaan virus, Ebola virus, Marburg virus, SARS Corona virus, Reovirus, Vaccinia virus and Rana grylio virus (summarized in (Liao *et al.*, 2019; Ren *et al.*, 2020)). Their mode of action is still under discussion. It is known that IFITMs inhibit viral entry. Several mechanisms were proposed: IFITMs may change the characteristics of the endosomal/lysosomal cavity (e.g. lipid concentration, pH) thereby making these structures unfavorable for virion fusion, IFITM proteins block the formation of fusion pores by changing the membrane fluidity and accumulation of cholesterol in cell membranes. affecting the cell membrane structure or further they can restrict viral assembly, reduce infectivity of nascent virions, inhibit viral protein synthesis or stimulate effective immune responses (Diamond and Farzan, 2013; Bailey *et al.*, 2014; Liao *et al.*, 2019; Zhao *et al.*, 2019; Kriesel, Schelle and Baldauf, 2020; Ren *et al.*, 2020; Friedlová *et al.*, 2022).

1.3.3 GBPs

GBPs are an evolutionary ancient protein family belonging to the dynamin superfamily. GBPs are ISGs that are involved in cell-autonomous innate immunity against parasites, bacteria and viruses (reviewed in (Tretina *et al.*, 2019; Kutsch and Coers, 2021; Zhang *et al.*, 2021; Schelle *et al.*, 2023)).

The functional cross-species conservation of GBPs between plants, animals and humans and GBPs acting as retroviral restriction factors in these species are reviewed in detail in **Review R2** (Schelle *et al.*, 2023) in **Appendix A**.

1.4 Evolution of immune genes

1.4.1 Virus-host coevolution

1.4.1.1 Host-virus arms race

During a viral infection, the evolutionary pressures for host survival and viral replication lead to an arms race between the host and the virus. The host and the virus are constantly "fighting" and exerting a selection pressure on the opponent, which drives the arms race through adaption and counteradaption. Interacting proteins are affected by the selection pressure, including viral restriction factors and their viral antagonists (**Figure 4**). The advantage of viruses is that their mutation rate is much higher; the advantage of the host is the genome size (more antiviral mechanisms) and the diploidy (more available alleles) (reviewed in (Little *et al.*, 2010; Daugherty and Malik, 2012; Duggal and Emerman, 2012)).

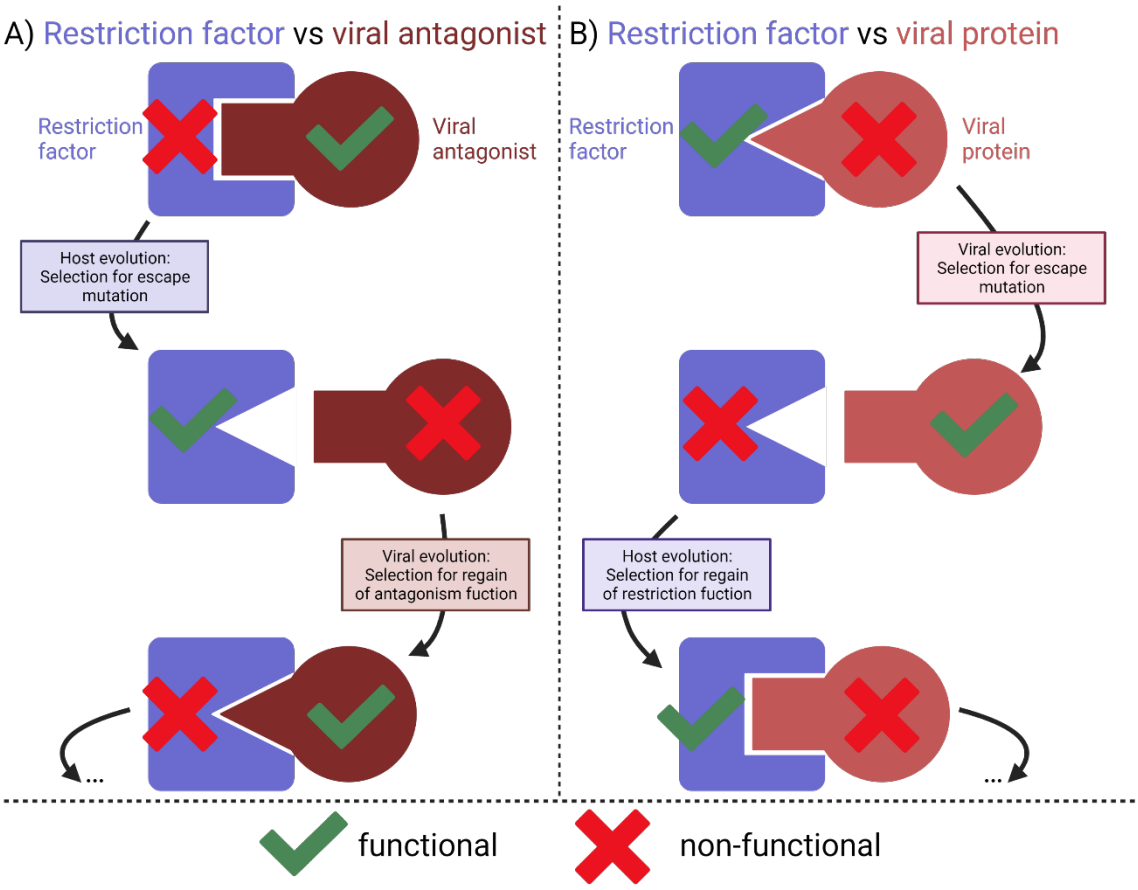

Figure 4: Figure: Virus-Host arms race. Depicted is an arms-race **A)** between an antiviral restriction factor and its viral antagonist, and **B)** between an antiviral restriction factor and a viral protein (compiled and modified from (Daugherty and Malik, 2012; Duggal and Emerman, 2012)) Created with BioRender.com.

1.4.1.2 Models of specific co-evolution

There are two current models that explain the specific host-virus co-evolution: the selective sweep co-evolution and balancing selection co-evolution ("Red Queen" hypothesis). The selective sweep co-evolution explains the rapid increase of beneficial variants by positive selection for that beneficial variant. The balancing selection co-evolution explains the balancing of allele frequencies by negative frequency-dependent selection, which is detrimental for common alleles. (reviewed in (Daugherty and Malik, 2012; Ebert and Fields, 2020)). The two models of specific co-evolution are compared in **Table 3**.

Table 3: Models of specific host-virus co-evolution: Selective sweep and balancing selection co-evolution (reproduced with permission from Springer Nature (license: 5730191346632) from (Ebert and Fields, 2020)).

Features	**Selective sweep co-evolution**	**Balancing selection co-evolution ("Red queen")**
Form of selection	Positive selection drives sweeps; selection is directional	Negative frequency-dependent selection gives common alleles a disadvantage; selection results in a balance of the frequencies of genetic variants
Functional polymorphisms	Visible only during selective sweeps	Maintained constantly and potentially for very long time periods
Underlying genetic system	Beneficial mutation in the host and parasite at any locus in the nuclear or cytoplasmic genome may sweep	Frequencies of alternative alleles at a few selected loci are balanced
Role of mutations	Mutations define the onset of new selective sweeps (hard sweeps)	Mutations are not necessary but do create rare variants, which may be selected and contribute to balancing selection or even replace a previous variant
Temporal continuity	Process can be highly stochastic and does not need to be continuous; long periods without sweeps are possible	Process must operate continuously because genetic variants may otherwise be lost. In a spatial setting, previously lost alleles may be reintroduced from other populations
Timescale of phenotypic change	Relatively slow because new mutations take a long time to reach a high enough frequency to be recognized. Sweeps starting from standing genetic variation progress more quickly	Fast because genetic variants are always at intermediate frequencies where selection results in fast changes
Population divergence	Sweeps drive population and species divergence	Population divergence is prevented in the long term, although it may occur in the short term

Features	**Selective sweep co-evolution**	**Balancing selection co-evolution ("Red queen")**
Evolutionary outcome	Creates macroevolutionary patterns (lineage divergence)	Explains high levels of genetic diversity within populations and species
Introgression among species	May introduce beneficial new alleles that can sweep	May introduce new functional variants that can contribute to balancing selection, but may create a fake picture of trans-species polymorphism

1.4.2 Evolution of multigene families

Genes involved in the immune response are often multigene families. Multigene families originate from one ancestral progenitor by gene duplication and therefore share similar sequences. In divergent evolution, multigene families evolve gradually under different selective pressures (**Figure 5 A**). Divergent evolution is insufficient to explain the observed patterns, and multigene families were thought to evolve by concerted evolution, i.e. the paralog genes evolve as a unit by genetic exchange through unequal crossing over and gene conversion (**Figure 5 B**) (Nei and Rooney, 2005). Since these two models cannot explain all evolutionary patterns, Nei *et al.* in 1997 proposed a third mechanism for the evolution of multigene families, the birth and death mechanism, i.e. genes duplicate, the duplicated genes can be maintained and diverge by neo- or subfunctionalization, or become non-functional or can be lost (**Figure 5 C**).

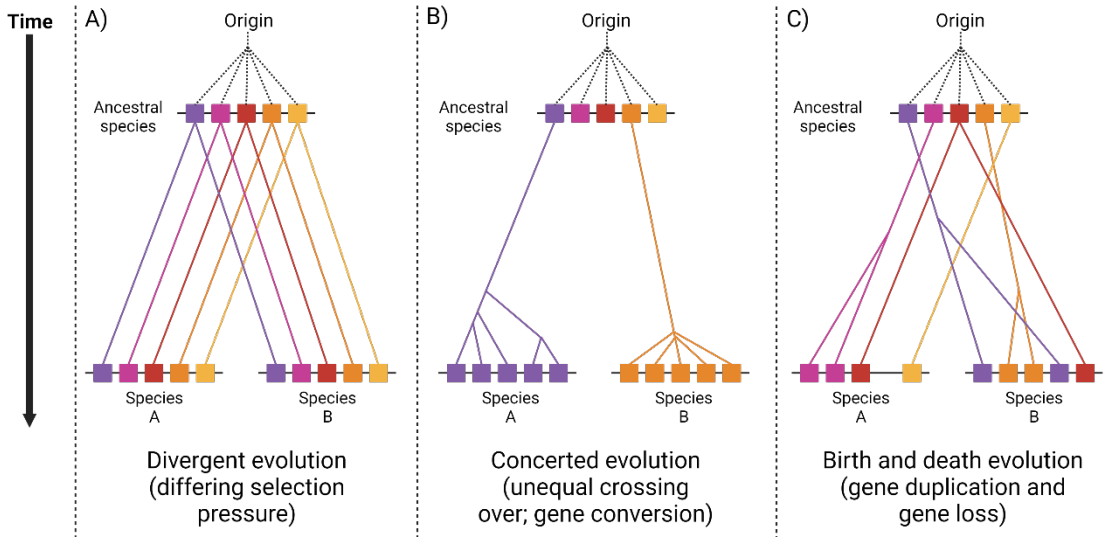

Figure 5: Models for the evolution of multigene families. Depicted are three different models for the evolution of multigene families - **A)** Divergent evolution **B)** Concerted evolution **C)** Birth and death evolution (modified from (Nei and Rooney, 2005)) Created with BioRender.com.

These are the theoretically proposed models, but in reality, it is rather not all-or-nothing. Mixed evolution processes have been observed, including two or three of the models (Nei and Rooney, 2005). In addition, other mechanisms are in place that contribute to the evolution of multigene families as e.g. retrogenes, which are described in the next chapter.

1.4.3 Role of retrogenes in evolution

Retrogenes (also referred as processed pseudogenes) are genes originating from a transcribed gene via retrotransposition of mRNA with a class 1 transposable element. These retrogenes can have several fates: acquisition of a promoter and expression (functional gene), neofunctionalization, development of a non-coding regulatory function and degeneracy. This makes them a major driver of evolution and an additional factor other than the three models of multigene family evolution described above (reviewed in (Kaessmann, Vinckenbosch and Long, 2009; Casola and Betrán, 2017; Cheetham, Faulkner and Dinger, 2020; Staszak and Makałowska, 2021; Troskie, Faulkner and Cheetham, 2021)). Recently, it has been described that retrogenes (protein expression and non-protein expressing) of antiretroviral restriction factors can contribute to the restriction of retroviruses (Yang *et al.*, 2020; Rahman and Compton, 2021; Rheinemann *et al.*, 2022). Therefore, they may also play an important role in limiting retroviral replication.

1.5 Animal models in retrovirology with focus on MLV and AIDS-related models

For MLV, an animal model is not a problem because the mouse, the most commonly used animal model, is the host species. Therefore, MLV has been used as a model virus to study retroviruses *in vivo*. Since humans are the natural host of HIV, surrogate models are used instead. Therefore, several animal models have been used to study HIV replication, therapy, cure and vaccine approaches. A model for studying lentiviruses *in vivo* is the feline immunodeficiency virus (FIV), which naturally infects cats and displays similarities in pathogenesis to HIV, eventually leading to feline AIDS (FAIDS). Downsides are differences in entry-receptor usage and in the set of accessory proteins (reviewed in (Burkhard and Dean, 2005; Hatziioannou and Evans, 2012; Policicchio, Pandrea and Apetrei, 2016). Non-human primate models, which are the most commonly used model, include several macaques with the disadvantage of using simian immunodeficiency virus (SIV) or SIV/HIV chimeras (SHIV). Several types of humanized mice are used as small-animal models for HIV, e.g. human peripheral blood lymphocytes severe combined immune deficiency, severe combined immune deficiency human, hematopoietic stem cells, and bone-liver-thymus mice (both models extensively reviewed in (Hatziioannou and Evans, 2012; Hessell and Haigwood, 2015; Kumar, Chahroudi and Silvestri, 2016; Policicchio, Pandrea and Apetrei, 2016; Wong, Jaworowski and Hearps, 2019; Weichseldorfer *et al.*, 2020; Waight *et al.*, 2022). Since all of these models have their specific advantages and disadvantages, additional suitable, especially small animal models for HIV are needed. Rabbits and tree shrews may be such putative HIV animal models because they better support HIV replication, although they also have species-specific limitations to full HIV replication. (Kulaga *et al.*, 1988; Tervo and Keppler, 2010; Luo *et al.*, 2021). Common to all non-human cells, HIV entry is species-specific and thus limited. In rabbits, TRIM5 restricts HIV at the level of reverse transcription and there is also a macrophage-specific infectivity defect after efficient particle release, the underlying mechanism of which is currently unknown (Schaller, Hué and Towers, 2007; Tervo and Keppler, 2010). In the tree shrews, viral infection is restricted by APOBEC3 (Luo *et al.*, 2021). Both species are more closely related to humans than other small animals, such as mice or rats, and can be bred, housed in animal facilities, genome-edited and are already used as animal models for various purposes, including viral infections (reviewed in (Esteves *et al.*, 2018; Soares, Pinheiro and Esteves, 2022) for rabbits and in (Kayesh *et al.*, 2021) for tree shrews). Aside from species-specific immunity, the aforementioned species-specific barriers by several mechanisms, including species-specific retroviral restriction factors, that can completely abolish retroviral replication, are general limitations of established and new animal models that need to be considered and/or overcome. The disadvantage of new animal models is that all processes have to be established and therefore involves a lot of work, but the introduction of new species

models could complement the traditionally used models to further advance HIV research in general on the way to finding an effective vaccine and/or cure. A summary of established animal models and animal models under development is visualized in **Figure 6.**

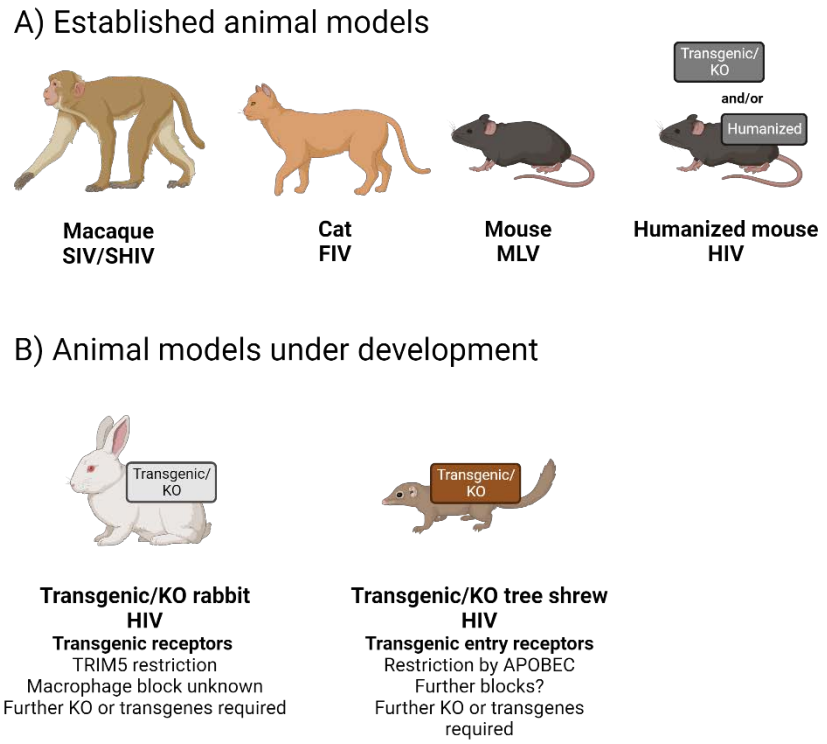

Figure 6: Retrovirus animal models and possible candidates. Shown are **A)** currently established animal models and **B)** animal models under development. **HIV**: Human Immunodeficiency Virus; **MLV**: Murine Leukemia Virus; **SIV**: Simian Immunodeficiency Virus; **FIV**: Feline Immunodeficiency Virus; Virus, **KO**: knock-out; **TRIM**: Tripartite Motif Containing; **APOBEC**: Apolipoprotein B mRNA Editing Catalytic Polypeptide-like Created with BioRender.com.

1.6 Objectives of the thesis

Immunity genes are under special selection pressure due to their co-evolution with pathogens. Restriction factors acting against retroviruses, which are part of the cell-autonomous innate immune system, are such immune genes. Data on their cross-species conservation in evolution and antiretroviral function are scarce. Therefore, they are particularly interesting to study in their cross-species conservation in evolution and function to elucidate the differences in immune response to retroviruses in closely related species and to uncover the relationships between the species and their immune mechanisms against retroviral infection. It also provides insight in currently used or future putative animal models and the prospects for translation to humans. We decided to focus on the antiretroviral factors IFITMs and GBPs, since most data have only been available for humans and mice.

For IFITMs, there has been insufficient information on the evolution of IFITMs in primates. Therefore, the objective was to investigate and clarify the evolution for primate IFITMs.

For GBPs, the evolution in primates has already been investigated. Furthermore, functional data is only available for human and mouse GBPs. Since GBPs reduce the infectivity of released particles by interfering with furin-mediated processing, and a macrophage-specific infectivity defect has been identified in rabbits, we decided to study GBPs in rabbits because of their high potential as a new animal model for HIV research. Our objective was to study the evolution and to shed light on the unexplored functionality of rabbit GBPs.

2. Summary

Retroviruses infect different species and cause disease. The immune system has several weapons to fight against pathogens. After physical barriers, innate immunity is the first line of defense against the viruses. Retroviral restriction factors are cell-autonomous innate immunity genes that are active against retroviruses, and the retroviral proteins counteract some of them. This interplay between restriction factors and the retroviruses shapes their coevolution in an arms race between virus and host cells. Multigene families of immune genes can evolve by several mechanisms: 1) canonical divergent evolution, 2) concerted evolution, 3) birth and death model of evolution. Interestingly, immune genes can vary widely in number and phylogenetic relationship between even closely related species. This complicates the need for animal models for retrovirus research, since retroviral restriction factors are different in animal models such as mouse, rabbit (which could become a model), or even in the closely related primates. Since this area of research is scarce, we investigated the cross-species conservation of retroviral restriction factors to shed light on their inter-species differences. Specifically, we focused on the cross-species conservation of 1) IFITMs and 2) GBPs as IFN-inducible cell autonomous innate immunity genes.

1) IFITMs interfere with retroviral entry by a mechanism that is not fully elucidated. We observed that primate IFITMs can be distinguished into 1) a canonical IFITM gene cluster (located on the same chromosome in a consistent arrangement) and 2) IFITM retrogenes (random and unique location within the genome due to retrotransposition). Our phylogenetic results from the canonical cluster led to the discovery of three novel groups of primate IFITMs (pIFITMs) located in the IR-IFITM clade: the prosimian pIFITMs(pro), the old world monkey pIFITMs(owm), and the new world monkey pIFITMs(nwm). Based on specific sequence features, we proposed a revised nomenclature for the primate IR-IFITM groups: IR-pIFITM1, IR-pIFITM2, IR-pIFITM3, IR-pIFITMnwm, IR-pIFITMown and IR-pIFITMpro. For pIFITM5 and pIFITM10, synteny and phylogenetic analyses suggested divergent evolution after primate evolution. For the IR-pIFITMs, the analyses reflected a combination of a birth-and-death and a concerted evolution model. In contrary to the canonical cluster, the additional IFITMs were scattered throughout the genomes. In depth sequence analyses revealed the presence of features characteristic of retrogenes which are retrotransposed by class 1 transposable elements. The IFITM retrogenes appeared to originate from more recent events. Taken together, we hypothesized that IFITM3/pro/nwm transcripts were subjected to continuous retrotransposition by class 1 transposable elements. This mechanism gave rise to the IFITM retro(pseudo)genes. Continuous pseudogenization and gene loss could explain the unique pattern of each primate species. In conclusion we suggested that the

mechanism of emergence of retro(pseudo)genes as described above represents a third mechanism of evolution for the primate IR-IFITMs with similarities to the birth-and-death model of evolution.

2) GBPs interfere with the proper maturation of the HIV and MLV glycoproteins. Evolutionarily, lagomorph GBPs, as well as human and murine GBPs, followed a pattern of gain and loss. We observed a general lack of GBP3/6/7 in the order Lagomorpha. Interestingly, we found a loss of GBP2, a substantial expansion of GBP4s and a unique duplication of GBP5 in Leporidae. Expression of leporid GBPs, determined by reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) and transcriptome data analysis, revealed that expression differed among tissues and cell types tested and that four GBPs were IFN-inducible by IFNα and/or IFNγ in primary rabbit macrophages. All rabbit GBPs could be overexpressed and localized intracellularly either continuously and/or discretely in the cytoplasm and/or nucleus, except ocGBP5L1 and rarely ocGBP5L2, which colocalized with the trans-Golgi network (TGN). Rabbit furin activity was only inhibited by GBP5L2. In conclusion, our study provided valuable insights into the evolution and the biological properties of the multifunctional family of ocGBPs. It suggested a role for GBPs in immune responses in species other than humans and mice.

In summary, the studies shed further light on the cross-species conservation of retroviral restriction factors, confirming the high variance in gene number and genetic variance between species, illustrating the very specific relationship between host immune genes and pathogens.

3. Zusammenfassung

Retroviren infizieren verschiedene Tierarten und verursachen Krankheiten. Das Immunsystem verfügt über mehrere Möglichkeiten Krankheitserreger zu bekämpfen. Neben den physischen Barrieren ist die angeborene Immunität die erste Verteidigungslinie gegen Viren. Retrovirale Restriktionsfaktoren sind zellautonome Gene der angeborenen Immunität, die gegen Retroviren aktiv sind, und die retroviralen Proteine wirken einigen von ihnen entgegen. Dieses Wechselspiel zwischen Restriktionsfaktoren und Retroviren prägt deren Koevolution in einem Wettrüsten zwischen Virus und Wirtszellen. Multigenfamilien von Immungenen können sich durch verschiedene Mechanismen entwickeln: 1) kanonische divergente Evolution, 2) konzertierte Evolution, 3) "birth & death" Modell der Evolution. Interessanterweise kann die Anzahl der Immungene und ihre phylogenetische Verwandtschaft selbst zwischen eng verwandten Arten stark variieren. Dies verkompliziert den Bedarf von Tiermodellen für die Retrovirus-Forschung, da die retroviralen Restriktionsfaktoren in Tiermodellen wie der Maus, dem Kaninchen (das zukünftig ein Tiermodell werden könnte) oder sogar in den eng verwandten Affen unterschiedlich sind. Da es in diesem Bereich nur wenige Forschungsarbeiten gibt, haben wir die artenübergreifende Konservierung der retroviralen Restriktionsfaktoren untersucht, um ihre Unterschiede zwischen den Arten aufzudecken. Insbesondere konzentrierten wir uns auf die artenübergreifende Konservierung von 1) IFITMs und 2) GBPs als IFN-induzierbare zellautonome Gene der angeborenen Immunität.

1) IFITMs stören den Eintritt von Retroviren durch einen Mechanismus, der noch nicht vollständig geklärt ist. Wir haben festgestellt, dass IFITMs bei Primaten in 1) ein kanonisches IFITM-Gencluster (auf demselben Chromosom in einheitlicher Anordnung) und 2) IFITM-Retrogene (zufällige und einzigartige Position innerhalb des Genoms aufgrund von Retrotransposition) unterschieden werden können. Unsere phylogenetischen Ergebnisse aus dem kanonischen Cluster führten zur Entdeckung von drei neuen Gruppen von Primaten-IFITMs (pIFITMs), die in der IR-IFITM-Klade angesiedelt sind: die pIFITMs(pro) der Halbaffen, die pIFITMs(owm) der Altweltaffen und die pIFITMs(nwm) der Neuweltaffen. Auf der Grundlage spezifischer Sequenzmerkmale haben wir eine überarbeitete Nomenklatur für die IR-IFITM-Gruppen der Primaten vorgeschlagen: IR-pIFITM1, IR-pIFITM2, IR-pIFITM3, IR-pIFITMnwm, IR-pIFITMown und IR-pIFITMpro. Für pIFITM5 und pIFITM10 deuten Syntenie und phylogenetische Analysen auf eine divergente Entwicklung entsprechend der Evolution der Primaten hin. Für die IR-pIFITMs ergaben die Analysen eine Kombination aus einem „birth & death" und einem konzertierten Evolutionsmodell. Im Gegensatz zu den kanonischen Clustern waren die zusätzlichen IFITMs über das gesamte Genom verstreut. Eingehende

Sequenzanalysen zeigten das Vorhandensein von Merkmalen, die für Retrogene charakteristisch sind, die durch transponierbare Elemente der Klasse 1 retrotransponiert werden. Die IFITM-Retrogene scheinen aus jüngeren Ereignissen hervorgegangen zu sein. Insgesamt stellten wir die Hypothese auf, dass IFITM3/pro/nwm-Transkripte einer kontinuierlichen Retrotransposition durch transponierbare Elemente der Klasse 1 ausgesetzt waren. Dieser Mechanismus führte zur Entstehung der IFITM-Retro(pseudo)gene. Die kontinuierliche Pseudogenisierung und der Genverlust könnten das einzigartige Genmuster der einzelnen Primatenarten erklären. Abschließend schlugen wir vor, dass der oben beschriebene Mechanismus der Entstehung von Retro(pseudo)genen einen dritten Evolutionsmechanismus für die IR-IFITMs der Primaten darstellt, der Ähnlichkeiten mit dem „birth & death“ Evolutionsmodell aufweist.

2) GBPs stören die ordnungsgemäße Reifung der Glykoproteine von HIV und MLV. Evolutionär gesehen folgten die lagomorphen, murinen und humanen GBPs einem Muster von Gewinn und Verlust. Wir beobachteten ein generelles Fehlen von GBP3/6/7 in der Ordnung Lagomorpha. Interessanterweise fanden wir einen Verlust von GBP2, eine erhebliche Zunahme von GBP4 und eine einzigartige Duplikation von GBP5 bei der Familie Leporidae. Die Expression der leproiden GBP, die mittels quantitativer Polymerasekettenreaktion (RT-qPCR) und Transkriptomdatenanalyse bestimmt wurde, ergab, dass sich die Expression in den untersuchten Geweben und Zelltypen unterschied und dass vier GBPs in primären Kaninchenmakrophagen durch IFNα und/oder IFNγ induzierbar waren. Alle Kaninchen-GBPs konnten überexprimiert werden und waren intrazellulär entweder kontinuierlich und/oder diskret im Zytoplasma und/oder im Zellkern lokalisiert, mit Ausnahme von ocGBP5L1 und selten ocGBP5L2, die mit dem trans-Golgi-Netzwerk (TGN) kolokalisierten. Die Aktivität von Kaninchen-Furin wurde nur durch GBP5L2 gehemmt. Zusammenfassend lässt sich sagen, dass unsere Studie wertvolle Einblicke in die Entwicklung und die biologischen Eigenschaften der multifunktionellen Familie der ocGBPs liefert. Die Studie deutet darauf hin, dass GBPs auch bei anderen Arten als Menschen und Mäusen eine Rolle bei Immunreaktionen spielen.

Zusammenfassend erweitern unsere Studien unser Wissen bezüglich der artenübergreifenden Konservierung retroviraler Restriktionsfaktoren und bestätigen die hohe Varianz in der Anzahl der Gene und die genetische Varianz zwischen den Arten, was die sehr spezifische Beziehung zwischen Wirtsimmungenen und Krankheitserregern verdeutlicht.

4. Publication 1)

Evolution of primate interferon-induced transmembrane proteins (IFITMs): a story of gain and loss with a differentiation into a canonical cluster and IFITM retrogenes

Luca Schelle[1], Joana Abrantes[2,3,4], Hanna-Mari Baldauf[1+], Pedro José Esteves[2,3,4,5+]

[1]Max von Pettenkofer Institute and Gene Center, Virology, National Reference Center for Retroviruses, Faculty of Medicine, LMU München, Munich, Germany

[2]CIBIO-InBIO, Research Center in Biodiversity and Genetic Resources, University of Porto, 4485-661 Vairão, Portugal

[3]BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, 4485-661, Vairão, Portugal

[4]Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, 4099-002 Porto, Portugal

[5]CITS - Center of Investigation in Health Technologies, CESPU, 4585-116 Gandra, Portugal

+ These authors share last authorship

This article was published 2023 in Frontiers in Microbiology and is available under: doi: 10.3389/fmicb.2023.1213685

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 26 July 2023
DOI 10.3389/fmicb.2023.1213685

OPEN ACCESS

EDITED BY
Terence L. Marsh,
Michigan State University, United States

REVIEWED BY
Chen Liang,
McGill University, Canada
Alex Compton,
National Cancer Institute at Frederick (NIH), United States

*CORRESPONDENCE
Hanna-Mari Baldauf
✉ baldauf@mvp.lmu.de
Pedro José Esteves
✉ pjesteves@cibio.up.pt

†These authors share last authorship

RECEIVED 03 May 2023
ACCEPTED 06 July 2023
PUBLISHED 26 July 2023

CITATION
Schelle L, Abrantes J, Baldauf H-M and Esteves PJ (2023) Evolution of primate interferon-induced transmembrane proteins (IFITMs): a story of gain and loss with a differentiation into a canonical cluster and IFITM retrogenes.
Front. Microbiol. 14:1213685.
doi: 10.3389/fmicb.2023.1213685

Evolution of primate interferon-induced transmembrane proteins (IFITMs): a story of gain and loss with a differentiation into a canonical cluster and IFITM retrogenes

Luca Schelle[1], Joana Abrantes[2,3,4], Hanna-Mari Baldauf[1*†] and Pedro José Esteves[2,3,4,5*†]

[1]Faculty of Medicine, Max von Pettenkofer Institute and Gene Center, Virology, National Reference Center for Retroviruses, LMU München, Munich, Germany, [2]CIBIO-InBIO, Research Center in Biodiversity and Genetic Resources, University of Porto, Vairão, Portugal, [3]BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Vairão, Portugal, [4]Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, Porto, Portugal, [5]CITS - Center of Investigation in Health Technologies, CESPU, Gandra, Portugal

Interferon-inducible transmembrane proteins (IFITMs) are a family of transmembrane proteins. The subgroup of immunity-related (IR-)IFITMs is involved in adaptive and innate immune responses, being especially active against viruses. Here, we suggest that IFITMs should be classified as (1) a canonical IFITM gene cluster, which is located on the same chromosome, and (2) IFITM retrogenes, with a random and unique location at different positions within the genome. Phylogenetic analyses of the canonical cluster revealed the existence of three novel groups of primate IFITMs (pIFITM) in the IR-IFITM clade: the prosimian pIFITMs(pro), the new world monkey pIFITMs(nwm) and the old world monkey pIFITMs(owm). Therefore, we propose a new nomenclature: IR-pIFITM1, IR-pIFITM2, IR-pIFITM3, IR-pIFITMnwm, IR-pIFITMowm, and IR-pIFITMpro. We observed divergent evolution for pIFITM5 and pIFITM10, and evidence for concerted evolution and a mechanism of birth-and-death evolution model for the IR-pIFITMs. In contrast, the IFITMs scattered throughout the genomes possessed features of retrogenes retrotransposed by class 1 transposable elements. The origin of the IFITM retrogenes correspond to more recent events. We hypothesize that the transcript of a canonical IFITM3 has been constantly retrotransposed using class 1 transposable elements resulting in the IFITM retro(pseudo)genes. The unique pattern of each species has most likely been caused by constant pseudogenization and loss of the retro(pseudo)genes. This suggests a third mechanism of evolution for the IR-IFITMs in primates, similar to the birth-and-death model of evolution, but via a transposable element mechanism, which resulted in retro(pseudo)genes.

KEYWORDS

interferon-induced transmembrane proteins, evolution, innate immunity, antiviral proteins, primates, transposable elements, retrogene

1. Introduction

Interferon-inducible transmembrane proteins are relatively small transmembrane proteins with around 130 amino acids (AA). These proteins are encoded by a family of interferon-stimulated genes (ISGs), *IFITM1*, *IFITM2*, and *IFITM3*, which were first discovered as interferon-inducible genes (Friedman et al., 1984), and the paralogs *IFITM5* and *IFITM10*. IFITMs are ancient proteins present in fish, amphibians, reptiles, birds, monotremes, marsupials and mammals (Hickford et al., 2012). Phylogenetically, IFITMs can be divided into three major clades: the immunity-related (IR-)IFITMs (IFITM1, IFITM2, and IFITM3), IFITM5 and IFITM10 (Zhang et al., 2012). IFITMs comprise 5 domains: the N-terminal domain, the CD255 domain, which contains intramembrane domain 1 (IM1) and conserved intracellular loop (CIL), and the C-terminus. The latter consists of intramembrane domain 2 (IM2) and the C-terminal domain (Bailey et al., 2013, 2014). Whether the IMs are intramembrane or rather transmembrane domains remains unclear as their exact topology in the membranes has not been solved and might differ between membrane types (reviewed in Bailey et al., 2014).

IFITMs are associated with several functions: the IR-IFITMs play a role in adaptive (reviewed in Yanez et al., 2020) and innate immune responses, especially against RNA and DNA viruses, with several mechanisms for viral inhibition observed and proposed (extensively reviewed in Diamond and Farzan, 2013; Bailey et al., 2014; Zhao et al., 2018; Liao et al., 2019). IFITM5 has acquired a Ca^{2+} binding site, which is important for its role in osteoblast function and bone mineralization (Hanagata et al., 2011; Hedjazi et al., 2022). The role of IFITM10 remains unclear, but it has recently been associated with gastric cancer (Liu et al., 2021).

Primates diverged into the suborders Strepsirrhini (prosimians) and Haplorrhini ~71.4–77.5 million years ago (MYA). The infraorders Simiiformes and Tarsiiformes (tarsier) originated from Haplorrhini ~61.6–71.1 MYA. At ~40.0–44.2 MYA, the Simiiformes branched to the parvorders of Platyrrhini (new world monkeys) and Catarrhini, which further divided ~26.80–30.60 MYA to Cercopithecidae (old world monkeys) and the superfamily Hominidea (apes), including Hylobatidae (gibbons) and Hominidae (great apes) (divergent times derived from Kumar et al., 2022).

Multigene families were originally believed to evolve by concerted evolution, i.e., the paralog genes would evolve as a unit by genetic exchange from unequal crossing over and gene conversion (Nei and Rooney, 2005). Nei et al. (1997) proposed the birth-and-death model of evolution for multigene families of the immune system where newly duplicated genes are either maintained in the genome and diverge functionally with neofunctionalization or subfunctionalization, or become nonfunctional or are deleted. These models are not mutually exclusive and genes can evolve in a mixed model process (Nei and Rooney, 2005).

Retrogenes or processed pseudogenes are functional retrocopies of genes originating from a parental gene by RNA-based gene duplication via retrotransposition by class 1 transposable elements. Retropseudogenes are the non-functional forms of retrogenes (reviewed in Kaessmann et al., 2009; Troskie et al., 2021). In order to be inherited, retrotransposition has to occur in the germline (Kaessmann et al., 2009). During a retrotransposition event, the mRNA of a parental gene is bound to reverse transcriptase of transposable elements; in mammals, these elements are long interspersed nuclear elements (LINEs), which recognize polyadenylated mRNA (Doucet et al., 2015). The bound mRNA is then retrotransposed to another genomic localization and integrated at a consensus cleavage site of the endonuclease by a process termed target-site primed reverse transcription (TPRT) (Luan et al., 1993; Troskie et al., 2021). Retropseudogenes are characterized by the lack of introns, and the presence of a conserved poly A signal (AATAAA), a poly A tail start and target-site duplications [5′ and 3′ untranslated region (UTR)] (Esnault et al., 2000; Kaessmann et al., 2009). The possible fate of retro(pseudo)genes has been reviewed by Troskie et al. (2021), and includes, for example, the acquisition of a promoter and expression, neofunctionalization, development of a non-coding regulatory function and degeneracy.

Some studies have addressed primate IFITM evolution (Hickford et al., 2012; Zhang et al., 2012; Compton et al., 2016; Wilkins et al., 2016; Benfield et al., 2020). In this study, we conducted a more in-depth study of IFITM evolution in primates by including more primate species (Rahman and Compton, 2021) into the analyses and considering the separation of canonical IFITMs cluster and IFITM retrogenes.

2. Results

2.1. Gene synteny of canonical *IFITM* cluster in primates

After retrieving all available primate *IFITM* sequences from the NCBI database (Accession numbers of the sequences are listed in Supplementary Table S1), we inferred the gene synteny, which is depicted in Figure 1 (right side). Genes used for synteny were located on the same chromosome or same unplaced scaffold in each species and were all flanked by the same genes (PGGHG, BAGALNT4, CTSD respectively; in gray in Figure 1), except for the *IFITMs* of *Rhinopithecus roxellana* and *Theropithecus gelada*, which were not flanked by *BAGALNT4* due to chromosomal rearrangements. This prompted us to term them the canonical IFITM cluster. Genes in red could not be aligned or were only partial mRNAs or pseudogenes, and were therefore excluded from the alignment (Figure 1).

For all the 26 species included, we observed that IFITM5 and IFITM10 consisted of single-copy genes at a conserved position in the synteny. The IR-IFITMs gene synteny was also conserved in the prosimians and apes; however, prosimians possessed two *IR-IFITMs*, with a distinct gene location and orientation rearrangement compared to *Otolemur garnettii* (Figure 1). The apes had three identically arranged *IR-IFITMs*, i.e., one more than the prosimians from which they separated around ~74 MYA (Kumar et al., 2022). For the new and old world monkeys, different numbers of *IR-IFITM* genes were observed, ranging from zero to six (Figure 1). We could not exclude that, especially in the case of single *IR-IFITMs,* additional genes might have been missed due to small size of the gene, gaps in scaffolds and/or poorer genome quality (Figure 1). In summary, we observed diversification of the gene copy number of the *IR-IFITMs* and their synteny in the apes, new and old world monkeys since the separation from the prosimians. In contrast, *IFITM5* and *IFITM10* appeared highly conserved as single copy genes present at a fixed location.

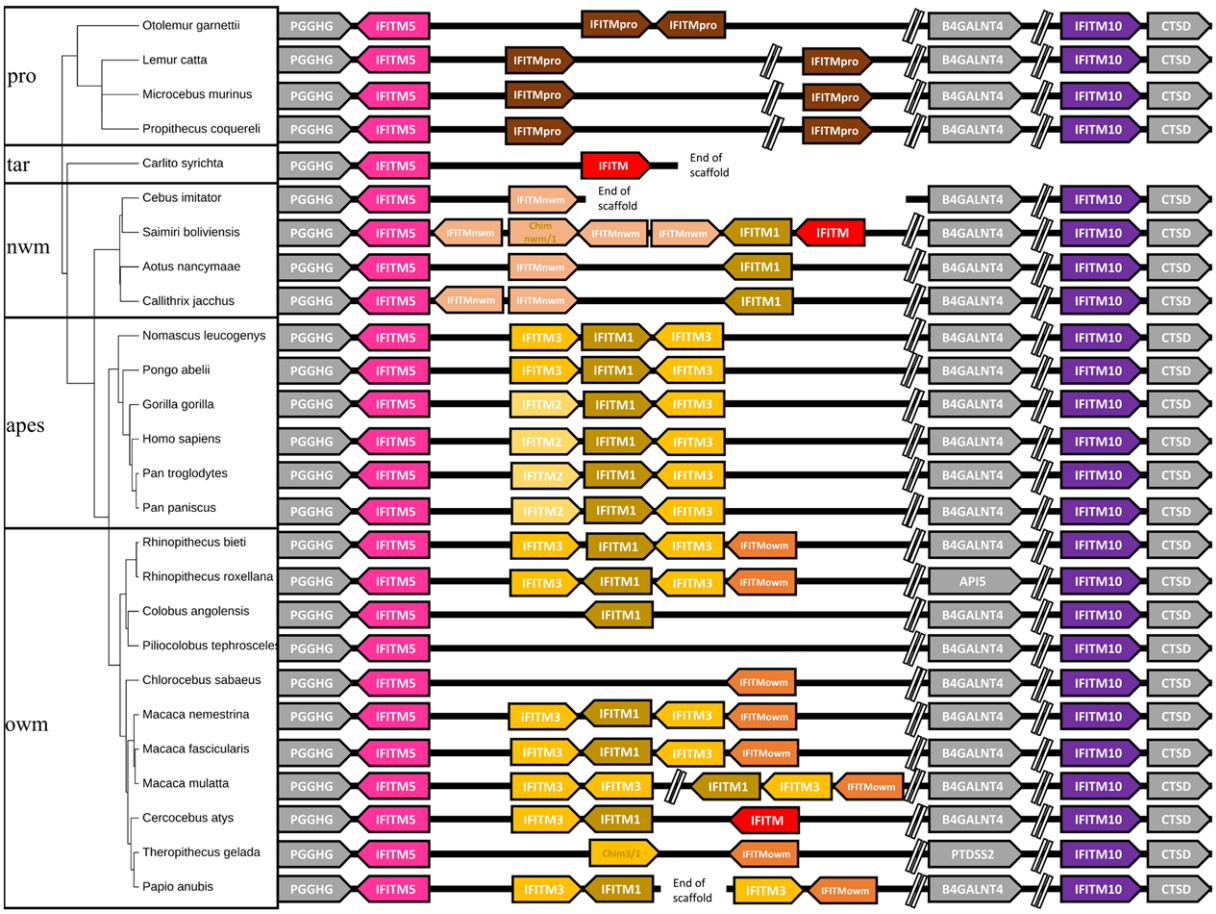

FIGURE 1
Gene synteny of primate *IFITMs* of the canonical cluster. The gene synteny of the primate *IFITMs* in the canonical cluster is displayed for the 26 analyzed primate species (right). *IFITMs* were colored following the grouping in the phylogenetic analyses (Figure 3). Arrows indicate gene orientation. Primate phylogeny (left) was constructed using timetree.org (Kumar et al., 2022). Gray: flanking genes, pink: *IFITM5*, purple: *IFITM10*, brown: *IR-pIFITMpro*, light orange: *IR-pIFITMnwm*, orange: *IR-pIFITMowm*, sand: *IR-pIFITM1*, yellow: *IR-pIFITM3*, light yellow: *IR-pIFITM2* red: not considered in the analyses, e.g., partial mRNA, Chim: Chimeric genes (see below); pro: prosimians; nwm: new world monkeys; owm: old world monkeys.

2.2. Distinction between canonical IFITMs cluster and IFITM retrogenes

For most of the primate species analyzed, in addition to the canonical cluster, we found various *IFITMs scattered* at different random positions within the genome, with most having a unique localization. In line with our observations that these genes are retrogenes (see Section 2.7), we propose that primate *IFITMs* can be classified according to their localization in the genome into canonical *IFITMs* cluster and *IFITM* retrogenes (Figure 2).

2.3. Phylogeny of canonical IFITM cluster in primates

For phylogenetic inference, only the IFITMs from the canonical cluster were used (Figure 3).

Considering the IR-IFITMs of primates (IR-p; Figure 3, accession numbers Supplementary Table S1, alignment Supplementary Figure S5), IR-pIFITM1 was only present in Simiiformes, while absent in prosimians, and formed a well-supported separate group in accordance with the primate phylogeny. The absence of IR-pIFITM1 in prosimians was unique for primates. The genes classified as *IR-pIFITM3s* did not cluster in accordance to primate phylogeny and appeared to be polyphyletic. The IR-pIFITM2 sequences clustered together (bootstrap value of 73), but they were only present in *Homo sapiens*, *Gorilla gorilla*, *Pan paniscus*, and *Pan troglodytes*. We also observed three new phylogenetic groups of primate IFITMs: one of the clusters comprised all prosimian IFITMs (pIFITM(pro)), the second included only old world monkeys IFITMs (pIFITM(owm)) and the third encompassed all NCBI annotated IFITM3 of new world monkeys (pIFITM(nwm)). Except for *Colobus angolensis* and *Piliocolobus tephrosceles*, all old world monkeys maintained a copy of the pIFITMowm, which is in addition to the pIFITM3s present in old world monkeys.

Regarding the phylogeny of the pIFITM5 (Supplementary Figure S1, accession numbers Supplementary Table S1, alignment Supplementary Figure S2) and pIFITM10 (Supplementary Figure S3, accession numbers Supplementary Table S1, alignment Supplementary Figure S4), clustering was according to the established primate phylogeny (Figure 1). The primate IFITM5s were highly

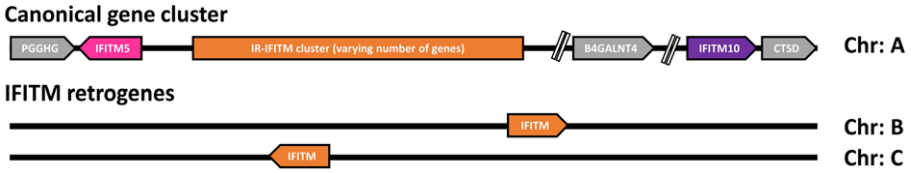

FIGURE 2
General genomic arrangement of canonical *IFITMs* cluster and *IFITMs* retrogenes. Schematic representation of the general arrangement of the proposed canonical gene cluster and *IFITM* retrogenes. The distinction between the consistently arranged canonical *IFITM* cluster on one chromosome (Chr) and the *IFITM* retrogenes, which are randomly distributed throughout the genome, is shown. Arrows indicate gene orientation. Gray: flanking genes, pink: *IFITM5*, purple: *IFITM10*, orange: *IR-IFITMs*.

conserved, with 72% (97/134) of the sites 100% conserved in all aligned species. The same applied for primate IFITM10s where 88% (115/130) of the sites were 100% identical. Indeed, the *IFITM5* and *IFITM10* genes of prosimians and tarsier, new world monkeys, old world monkeys and apes clustered into closely related separate groups, with the exception of IFITM5 of *Macaca* species (Supplementary Figure S1). This was most likely caused by a point mutation leading to an amino acid exchange (G19R), compared to the otherwise identical sequences of old world monkey IFITM5s (Supplementary Figure S2).

2.4. Sequence characteristics of primate IR-IFITM groups

To further characterize and classify the six proposed groups of primate IR-IFITMs, we investigated the AA sequences of the N-termini (Figure 4A), the CD225 middle domains (Figure 4B), and the C-termini (Figure 4C). The CD225 domain sequence was based on the alignment of all six groups, because they were highly conserved except for two AAs (Figure 4B).

We observed that the groups could be characterized by their N- and C-termini (Figure 4), as the remaining CD225 domains were highly conserved and not informative. IR-pIFITM1 and IR-pIFITMowm had shorter N-termini (20-21 AA) compared to IR-pIFITM2/3/nwm/pro, while IR-pIFITMowm also had small deletions next to the start codon (5 AA and 7 AA). The IR-pIFITM2/3/nwm/pro N-termini were of the same length, except that IRpIFITM2 had a deletion of one AA. The N-termini of IR-pIFITM2/3/nwm showed higher similarity to each other than to IR-pIFITMpro, but differed especially at positions 4–16 and 27 (Figure 4A). The IM2 domain of the C-terminus was less conserved than the CD225 domain and therefore a further determinant of the six groups, but the IMs of IR-pIFITM2/3/nwm were more similar. The C-terminal domains differed between the groups in length and sequence. IR-pIFITM1s had an elongated C-terminal domain, while the domain was lost in IR-pIFITMowm. IR-pIFITM2/3/nwm/pro had C-terminal domains of the same length but differed in sequence (Figure 4C). In summary, all primate IR-IFITM groups comprised a highly conserved CD225; yet, they can be differentiated and classified by their N- and C-termini, which were group-specific both in terms of sequence and length.

2.5. New classification of primate IFITMs

Based on our analyses, we propose a new nomenclature for the primate *IR-IFITMs* as *IR-pIFITM1*, *IR-pIFITM2*, *IR-pIFITM3* (Immunity-Related-primate), *IR-pIFITMnwm* (Immunity-Related-primate-new world monkey) *IR-pIFITMowm* (Immunity-Related-primate-old world monkey) and *IR-pIFITMpro* (Immunity-Related-primate-prosimian). The old and new nomenclature is listed in Table 1. This phylogeny-based proposed nomenclature does not specify individual genes in a species if more than one gene is present. Due to the closer relationship between paralogs of a species, caused by concerted evolution, than to orthologs, a relationship-based specification was not possible. Therefore, we suggest to specify them according to their synteny as locus (L) + number (1, 2, 3…) = L1, L2, L3… without emphasizing any phylogenetic or functional relationship.

2.6. IR-pIFITM1/3 chimeras

In *Theropithecus gelada* and *Saimiri boliviensis*, we found longer IFITMs sequences that did not align with either of the six primate groups. The alignment of these IFITMs revealed two chimeric sequences with recombination between an IR-pIFITM3/nwm at the N-termini and an IR-pIFITM1 at the C-termini (Figure 5).

2.7. Genomic localization of additional primate IFITMs

We observed that the many additional IFITMs were not localized in the canonical clusters, but rather spread throughout the genome. In prosimians, only one additional IFITM was present in *Otolemur garnettii*. For the remaining primates, variable numbers of additional IFITMs were detected, ranging from 6 to 21 genes (Table 2). We further noted an increased number of these additional IFITMs after the separation of prosimians from all other primates.

For 13 selected species, covering apes (all apes), old and new world monkeys (randomly selected representatives), tarsier (only one genome available) and prosimians (only one species with additional IFITM), we mapped the scattered IFITMs to characterize their synteny (Supplementary Table S2). The genomic localization of the scattered IFITMs appeared random and unique. Further, we observed

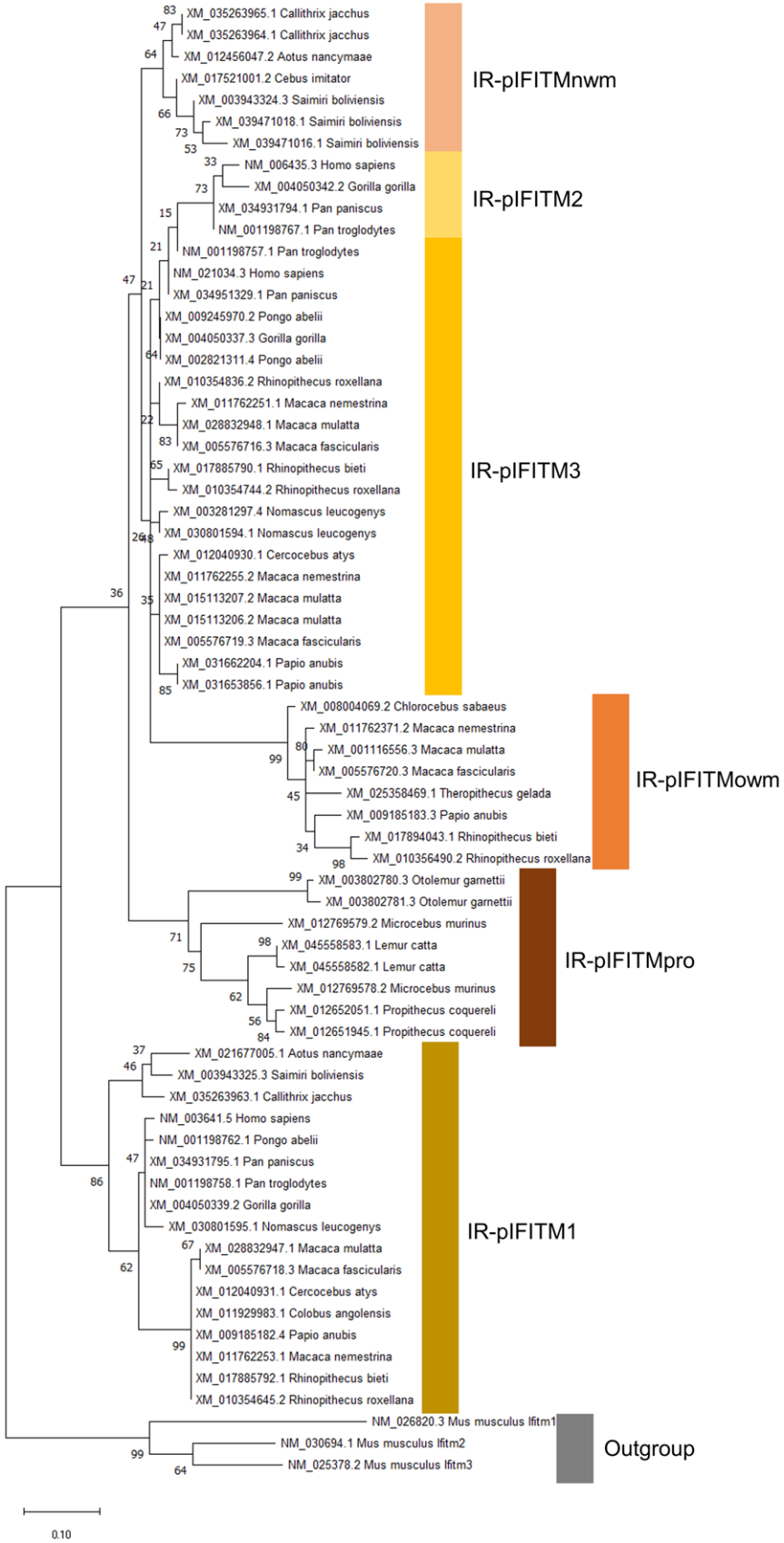

FIGURE 3
Phylogeny of IR-IFITM in 26 primate species based on AA sequences. The evolutionary history was inferred by using the Maximum Likelihood (ML) method. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The bootstrap value is shown next to the

(Continued)

FIGURE 3 (Continued)

branches. Mouse Ifitms were used as outgroup. IR-pIFITM1, IR-pIFITM2, IR-pIFITM3 (Immunity-Related-primate), IR-pIFITMnwm (Immunity-Related-primate-new world monkey) IR-pIFITMowm (Immunity-Related-primate-old world monkey), and IR-pIFITMpro (Immunity-Related-primate-prosimian).

that a considerable number was located in the intronic regions of other genes, especially in new and old world monkeys (Table 3). Only in closely related species, we observed a genomic overlap, with some IFITMs present in more than one species flanked by the same genes (mostly among apes, some among old world monkeys, one among new world monkeys and none for tarsier and prosimians; Table 4).

2.8. Additional IFITIMs are IFITM retrogenes

This random distribution and localization in introns of other genes hinted toward transposable element mechanisms and retrogenes. To test this hypothesis, we randomly picked two additional IFITMs from each analyzed species (only one if no more were available) and analyzed the genomic context. For this, we searched for features of retrogenes 200 bp upstream of the canonical start codon and 400 downstream of the canonical stop codon (Supplementary Figure S9). The results are summarized in Table 5.

We observed that all investigated sequences lacked an intron, except for one in *Carlito syrichta*. They had a consensus poly A signal, the start of the poly A tail and target-site duplications (TSDs) adjacent to the poly A tail start and upstream of the canonical start codon. These are all features of retrogenes (Esnault et al., 2000; Kaessmann et al., 2009). For the coding sequences, we also found some with premature stop codons (8/25 tested, e.g., in *Pan paniscus* and *Aotus nancymaae*), which are an indication for retropseudogenes.

Since we observed that the additional *IFITMs* were retrogenes, we compared them with genes from the canonical cluster to infer their origin. For this, we aligned two selected *IFITM* retrogene genomic sequences of each species with the mRNAs of *IR-pIFITM3, IR-pIFITMnwm* or *IR-pIFITMpro* from the canonical cluster (Supplementary Figure S10). We observed that the genomic sequences aligned with the mRNA sequences of *IR-pIFITM3, IR-pIFITMnwm* or *IR-pIFITMpro*, suggesting that these might have been the origin (parental genes) of the IFITMs retrogenes. Further, we observed that the two selected IFITM retrogenes aligned better with the canonical mRNA of *IR-pIFITM3*, *IR-pIFITMnwm* or *IR-pIFITMpro* from the same species and that even the 5′ and 3’ UTR parts aligned with only few nucleotide mismatches (Supplementary Figure S10). This suggests that the emergence from their parental gene was a recent event. In summary, the additional IFITMs are retrogenes or retropseudogenes that exhibit various retrogenic features and could have originated from parental genes in the canonical cluster in a more recent event.

3. Discussion

In this study, we examined the evolution of the IFITM protein family in primate species. Our synteny analyses suggest that primate IFITMs can be classified according to their localization within the genome into a canonical IFITM cluster, which includes *IFITM5*, *IFITM10*, *IR-IFITM*, and *IFITM* retrogenes (Figure 2). We observed that the primates *IFITM5* and *IFITM10* were present as single copy genes with conserved synteny: IFITM5 was flanked by PGGHG and IFITM10 by CTSD (Figure 1). This high conservation and the presence of a single copy are most likely related to their essential function as shown by the link between their absence or the presence of mutations and diseases (Hanagata et al., 2011; Liu et al., 2021; Hedjazi et al., 2022). In contrast, a diversification of the gene copy numbers of the IR-IFITMs (zero to six genes) and their synteny occurred in primates after their separation from prosimians around 74 MYA, which consistently possessed two copies of *IR-pIFITMpro* (Figure 1) (Kumar et al., 2022). IR-IFITMs of new and old world monkeys underwent massive rearrangements with gene expansions and losses. In contrast, apes uniformly possessed three IR-IFITM genes, arranged identically; therefore, at least one duplication event must have occurred after the separation from the prosimians. We can only speculate that the synteny is more conserved in apes and prosimians, because they have shared the same specificity for pathogens due to their close relationship. The overall high variability in the number of IR-IFITMs genes in the primate species could be related to their function in the immune response and co-evolution with species-specific pathogens as seen for other immunity-related proteins (Nei et al., 1997; Côrte-Real et al., 2020), resulting in repertoires specific for each species. In line with this, primate IFITMs might follow the birth-and-death model evolution that often occurs in immunity-related genes (Nei et al., 1997; Nei and Rooney, 2005).

In contrast to other phylogenetic studies including primate IFITMs (Siegrist et al., 2011; Hickford et al., 2012; Zhang et al., 2012; Compton et al., 2016; Wilkins et al., 2016; Benfield et al., 2020), we conducted a study including more primate species (26 species) while the others focused on smaller subsets, which improved the resolution of our phylogenetic analysis. Further, we focused our phylogenetic analyses on the IFITMs in the canonical clusters (Figure 2) with the underlying hypothesis that these IR-IFITMs suffered similar selective pressures. In contrast, we assumed that IFITM retrogenes (see below), experienced differences in the selective pressure, probably due to their redundancy, genomic localization, and pseudogenization accompanied by altered expression (Kaessmann et al., 2009; Troskie et al., 2021). The exclusion of these IFITM retrogenes allowed us to reduce bias from the altered selection pressure and improved the alignments, the basis of the phylogeny, by removing indels.

Hickford et al. (2012) focused on marsupial IFITMs and reported only the presence of canonical IFITMs with overall low similarity to other paralogs at the AA level. In line with that, Benfield and colleagues identified chiropteran IFITMs that formed a monophyletic group separated from other taxa by a relatively long branch (Benfield et al., 2020). On the other hand, Zhang et al. (2012) performed a more general evolutionary analysis of mammalian and non-mammalian IFITMs, including only six primate species. They found that all

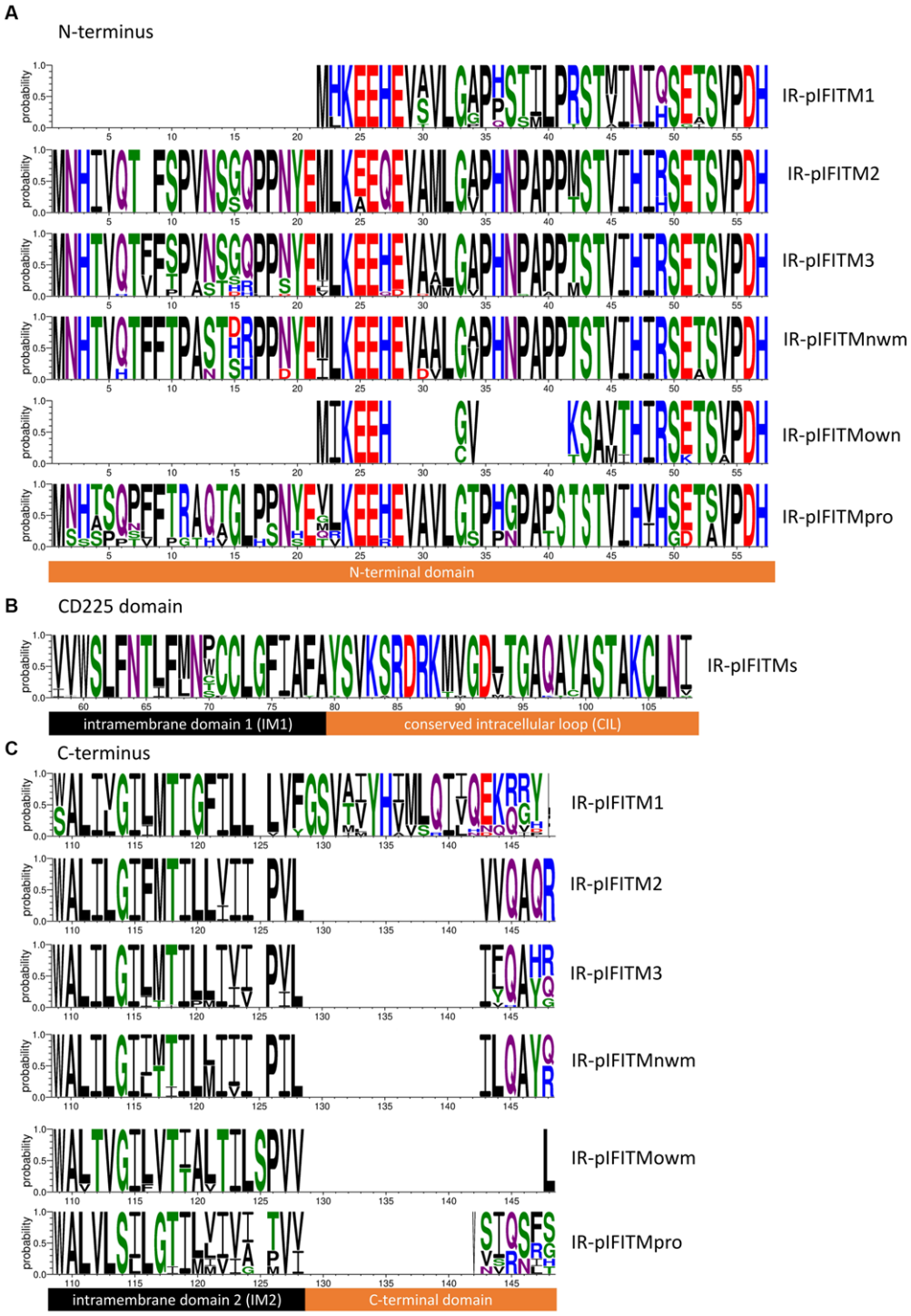

FIGURE 4
AA sequence characteristics of primate IFITM groups. Sequence logos were derived from the AA alignments of the primate IFITM groups (Supplementary Figures S6–S8) defined in Figure 3 (IR-pIFITM1, IR-pIFITM2, IR-pIFITM3, IR-pIFITMnwm, IR-pIFITMowm, and IR-pIFITMpro). **(A)** N-termini with variable lengths. **(B)** Highly conserved CD225 domains comprising IM1 domain and CIL. **(C)** C-termini including IM2 and the C-terminal domains with highly variability in length. Probability of residues is shown. Protein domains are indicated. Black: transmembrane/intramembrane domain, orange: topological domain. Logos were generated using WebLogo3 (Crooks et al., 2004).

IR-IFITM genes from the different lineages formed their own subgroups, suggesting gene duplication of IR-IFITM as an evolutionary mechanism after species separation. Focusing on the evolution of primate IFITM3s, Compton et al. (2016) identified an atypical gene locus in humans compared to bush baby species and suggested gene gain and loss events for primate evolution. A high number of pseudogenes per IFITM genes was already noted for human paralogs by Siegrist et al. (2011).

TABLE 1 New proposed classification of primate IR-pIFITMs.

Primate group	Primate	Accession number	Old classification	New classification
New world monkeys	*Callithrix jacchus*	XM_035263965.2	*IFITM3*	*IR-pIFITMnwm*
	Callithrix jacchus	XM_035263964.2	*IFITM3*	
	Aotus nancymaae	XM_012456047.2	*IFITM3*	
	Cebus imitator	XM_017521001.2	*IFITM3*	
	Saimiri boliviensis	XM_003943324.3	*IFITM3*	
	Saimiri boliviensis	XM_039471018.1	*IFITM3*	
	Saimiri boliviensis	XM_039471016.1	*IFITM3*	
Great apes	*Homo sapiens*	NM_006435.3	*IFITM2*	*IR-pIFITM2*
	Gorilla gorilla	XM_004050342.2	*IFITM2*	
	Pan paniscus	XM_034931794.1	*IFITM2*	
	Pan troglodytes	NM_001198767.1	*IFITM2*	
	Pan troglodytes	NM_001198757.1	*IFITM3*	*IR-pIFITM3*
	Homo sapiens	NM_021034.3	*IFITM3*	
	Pan paniscus	XM_034951329.1	*IFITM3*	
	Pongo abelii	XM_009245970.2	*IFITM3*	
	Gorilla gorilla	XM_004050337.3	*IFITM3*	
	Pongo abelii	XM_002821311.5	*IFITM3*	
Old world monkeys	*Rhinopithecus roxellana*	XM_010354836.2	*IFITM3*	
	Macaca nemestrina	XM_011762251.1	*IFITM3*	
	Macaca mulatta	XM_028832948.1	*IFITM3*	
	Macaca fascicularis	XM_005576716.3	*IFITM3*	
	Rhinopithecus bieti	XM_017885790.1	*IFITM3*	
	Rhinopithecus roxellana	XM_010354744.2	*IFITM3*	
Gibbon	*Nomascus leucogenys*	XM_003281297.4	*IFITM3*	
	Nomascus leucogenys	XM_030801594.1	*IFITM3*	
Old world monkeys	*Cercocebus atys*	XM_012040930.1	*IFITM3*	
	Macaca nemestrina	XM_011762255.2	*IFITM3*	
	Macaca mulatta	XM_015113207.2	*IFITM3*	
	Macaca mulatta	XM_015113206.2	*IFITM3*	
	Macaca fascicularis	XM_005576719.3	*IFITM3*	
	Papio anubis	XM_031662204.1	*IFITM3*	
	Papio anubis	XM_031653856.1	*IFITM3*	
Old world monkeys	*Chlorocebus sabaeus*	XM_008004069.2	*IFITM3*	*IR-pIFITMowm*
	Macaca nemestrina	XM_011762371.2	*IFITM2*	
	Macaca mulatta	XM_001116556.3	*IFITM3*	
	Macaca fascicularis	XM_005576720.3	*IFITM3*	
	Theropithecus gelada	XM_025358469.1	*IFITM3*	
	Papio anubis	XM_009185183.3	*IFITM3*	
	Rhinopithecus bieti	XM_017894043.1	*IFITM3*	
	Rhinopithecus roxellana	XM_010356490.2	*IFITM3*	
Prosimians	*Otolemur garnettii*	XM_003802780.3	*IFITM3*	*IR-pIFITMpro*
	Otolemur garnettii	XM_003802781.3	*IFITM3*	
	Microcebus murinus	XM_012769579.2	*IFITM3*	
	Lemur catta	XM_045558583.1	*IFITM3*	
	Lemur catta	XM_045558582.1	*IFITM3*	
	Propithecus coquereli	XM_012652051.1	*IFITM3*	
	Propithecus coquereli	XM_012651945.1	*IFITM3*	

(Continued)

TABLE 1 (Continued)

Primate group	Primate	Accession number	Old classification	New classification
New world monkeys	*Aotus nancymaae*	XM_021677005.1	*IFITM1*	*IR-pIFITM1*
	Saimiri boliviensis	XM_003943325.3	*IFITM1*	
	Callithrix jacchus	XM_035263963.1	*IFITM1*	
Apes	*Homo sapiens*	NM_003641.5	*IFITM1*	
	Pongo abelii	NM_001198762.1	*IFITM1*	
	Pan paniscus	XM_034931795.1	*IFITM1*	
	Pan troglodytes	NM_001198758.1	*IFITM1*	
	Gorilla gorilla	XM_004050339.2	*IFITM1*	
	Nomascus leucogenys	XM_030801595.1	*IFITM1*	
Old world monkeys	*Macaca mulatta*	XM_028832947.1	*IFITM1*	
	Macaca fascicularis	XM_005576718.3	*IFITM1*	
	Cercocebus atys	XM_012040931.1	*IFITM1*	
	Colobus angolensis	XM_011929983.1	*IFITM1*	
	Papio anubis	XM_009185182.4	*IFITM1*	
	Macaca nemestrina	XM_011762253.1	*IFITM1*	
	Rhinopithecus bieti	XM_017885792.1	*IFITM1*	
	Rhinopithecus roxellana	XM_010354645.2	*IFITM1*	

Shown is the old and proposed new classification of primate IFITMs. Order corresponds to phylogenetic tree (Figure 3).

Based on our phylogenetic analyses (Figure 3) and further supported by their sequence characteristics, length and AA sequences of the N- and the C-termini (Figure 4), we found six groups of primate IR-IFITMs. Therefore, we propose a new classification: IR-pIFITM1, IR-pIFITM2 and IR-pIFITM3, in line with previous studies (Hickford et al., 2012; Zhang et al., 2012; Compton et al., 2016; Benfield et al., 2020), and three new groups, the IR-pIFITMnwm, IR-pIFITMowm and IR-pIFITMpro (Figure 3). A shortcoming of our study is the lack of functional studies, especially those that have not been studied before such as pIFITMpro. However, our more in-depth evolutionary analyses might guide future functional studies.

The IR-pIFITMpro group is only present in prosimians. It is noteworthy that the two IFITMs genes of the prosimians belong to the IR-pIFITMpro group and neither IR-pIFITM1 nor IR-pIFITM3 are present. It is unclear whether the prosimian ancestor possessed IR-pIFITM1 and/or IR-pIFITM3 "progenitors," which were lost as a result of concerted evolution with the emergence of an IR-pIFITMpro group, or *vice-versa*: the birth-and-death model of evolution led to the emergence of IR-pIFITM1 and IR-pIFITM3/nwm "progenitor" in the Simiiformes. The subsequent separation of the IR-pIFITM3/nwm "progenitor" into IR-pIFITM3 and IR-pIFITMnwm could have been caused by similar mechanisms. The concerted evolution hypothesis is backed up by our finding of several highly supported subgroups (>83 bootstraps) of IR-IFITM3/nwm from the same species (Figure 3, e.g., *Callithrix jacchus* and *Papio Anubis*) and two chimeras between IR-pIFITM3/nwm and IR-pIFITM1 (Figure 5), suggesting gene conversion in new and old world monkeys and, therefore, a concerted evolution mechanism (Nei and Rooney, 2005). The IR-pIFTM2 genes are most likely a duplication of IR-pIFITM3, which gradually diverged in the apes.

Regarding the IR-pIFITMowm group, each species, except *Colobus angolensis* and *Piliocolobus tephrosceles,* had one *IR-pIFITMowm* gene. The phylogeny suggests that it probably arose by deletions from a duplication of an *IR-pIFITM3* (Figure 3), but we cannot exclude gene conversion or a chimeric origin, as it is not possible to assign an origin based on sequence motifs due to truncations at the C- and N-termini (Figure 4). One copy has been stably maintained in all but two old world monkey species, suggesting an evolutionary advantage for its presence. A possible explanation might be that IR-pIFITMowms were active against a bacterial or a viral pathogen or may have acquired a new function (neofunctionalization) and were thus maintained. Taken together, we found evidence for both concerted evolution and the birth-and-death evolution model for the canonical cluster of the IR-pIFITMs, which could indicate their evolution by a possible mixed process of both models (Nei and Rooney, 2005). The evolution of *IFITM5* and *IFITM10*, which had only one highly conserved copy at canonical positions in each species, were in line with the primate evolution (Supplementary Figures S1, S3; Figure 1).

The number of the IFITMs not in the canonical cluster was expanded in Simiiformes, probably after the separation from the prosimians (Table 2). Based on their synteny, we found that they were randomly distributed throughout the genomes and that a fraction of them were located in the intronic regions of other genes (Supplementary Table S2; Table 3). Since some IFITM genes, including human IFITM4P, have been proposed to be retrogenes (Siegrist et al., 2011; Rahman and Compton, 2021), we hypothesized that any additional primate IFITMs might also be retrogenes. Our analyses demonstrated that, along with their randomly scattered location and location within introns, all of them possessed additional features of retrogenes retrotransposed by class 1 transposable elements, such as lack of introns, the presence of

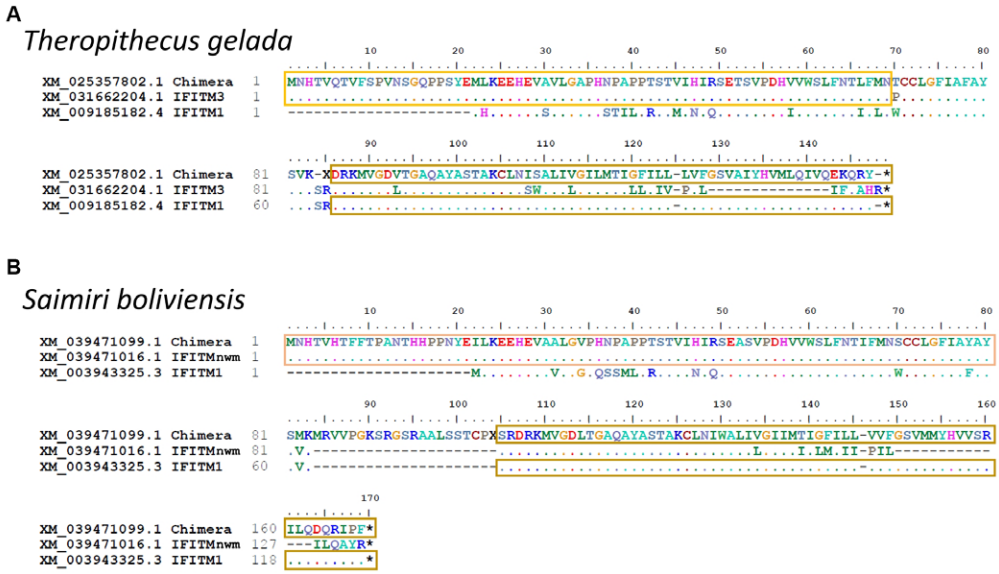

FIGURE 5
Identification of IR-pIFITM3/1 and IR-pIFITMnwm/1 chimeras. Alignment of chimeric IR-pIFITMs of *Theropithecus gelada* (owm) **(A)** and *Saimiri boliviensis* (nwm) **(B)** with IR-pIFITM3/nwm and IR-pIFITM1. For *Theropithecus gelada*, the alignment was performed with the protein sequences from its closest relative *Papio anubis* as only the chimeric gene is present in the genome. For *Saimiri boliviensis*, X represents start of frameshift omitted to emphasize identity of C-terminus. Identity to IR-pIFITM3/nwm highlighted with yellow or light orange box, respectively, and identity to IR-pIFITM1 highlighted with sand box.

TABLE 2 Number of IFITM retrogenes in primate species.

	Species	No. of additional IFITMs
Apes	*Homo sapiens*	12
	Pan paniscus	5
	Pan troglodytes	5
	Gorilla gorilla	3
	Pongo abelii	4
	Nomascus leucogenys	5
Old world monkeys	*Cercocebus atys*	10
	Chlorocebus sabaeus	14
	Macaca fascicularis	7
	Macaca mulatta	8
	Macaca nemestrina	8
	Papio anubis	11
	Theropithecus gelada	11
	Colobus angolensis	12
	Piliocolobus tephrosceles	8
	Rhinopithecus bieti	6
	Rhinopithecus roxellana	8
New world monkeys	*Aotus nancymaae*	10
	Cebus imitator	26
	Callithrix jacchus	28
	Saimiri boliviensis	20
Tarsier	*Carlito syrichta*	7
Prosimians	*Microcebus murinus*	0
	Propithecus coquereli	0
	Otolemur garnettii	1
	Lemur catta	0

TABLE 2 (Continued)

Shown is the number of IFITM retrogenes in the respective species.

conserved poly A signal (AATAAA), poly A start, and target site duplications (TSDs; 5′ and 3′ UTR) (Table 5) (Esnault et al., 2000; Kaessmann et al., 2009) and can therefore be designated as retrogenes. Sixteen of the analyzed genes had a complete coding sequence, but eight presented premature stop codons, which allowed their classification as retrogenes and retropseudogenes, respectively. In the alignment of the IFITM retrogene genomic sequences with the mRNA sequences of the canonical IR-pIFITM3, IR-pIFITMnwm or IR-pIFITMpro, we observed that the genomic sequences aligned best with the mRNA sequences of the IR-pIFITM3, IR-pIFITMnwm or IR-pIFITMpro from the same species, respectively. Furthermore, we observed that even the 5′ and 3′ UTR parts aligned with only few nucleotide mismatches with the mRNA sequences of the IR-pIFITM3, IR-pIFITMnwm or IR-pIFITMpro from the same species (Supplementary Figure S10). This suggests that the transcript of these canonical IFITMs may have been the origin (parental gene) of the retro(pseudo)gene, and that the event was recent because the TSDs and the poly A signal and tail, which degenerate over time, were mostly intact (Kaessmann et al., 2009). In conclusion, we hypothesize that the

10.3389/fmicb.2023.1213685

TABLE 3 IFITM retrogenes in primate species with location in introns of other genes.

No.	Species and location	Direction	Flanking gene	Direction	Annotated	Accession number	Direction	Flacking gene	In intron	Direction	Gene
5	Chr. 12	>	SH2B3	>	IFITM3P5 pseudo	NG_006225.2	<	BRAP	in	<	ATXN2
6	Chr. 12	<	AMIGO2	<	IFITM3P6 pseudo	NG_006230.2	<	RPAP3	in	>	PCED1B
	Pan paniscus										
2	Chr. 12	>	SH2B3	>	IFITM3P5 pseudo	XM_034934516.1	<	BRAP	in	<	ATXN2
	Macaca mulatta										
1	Chr. 16	<	CTC1	<	3 like 8 (owm P1)	XM_001112566.4	>	RANGRF	in	>	PFAS
2	Chr. 1	>	ATP5PB	>	3 like (owm P2)	XM_001106166.4	>	RAP1A	in	<	TMIGD3
4	Chr. 10	<	SLC25A17	>	3 ps (owm P4)	XR_001447798.2	>	XPNPEP3	in	<	ST13
6	Chr. 9	<	OIT1	>	3 ps	XR_001447146.2	>	MICU1	in	<	MCU
7	Chr. 11	>	SH2B3	>	P5	XR_001448216.2	<	BRAP	in	<	ATXN2
	Cercocebus atys										
2	Unplaced	>	ATP5PB	>	3 like (owm P2)	XR_001017992.1	>	RAP1A	in	<	TMIGD1
3	Unplaced		End	>	3 ps	XR_001010903.1	>	TAB3	in	>	dystropine-like
4	Unplaced	<	CTC1	<	3 like 8 (owm P1)	XM_012051598.1	>	RANGRF	in	>	PFAS
6	Unplaced	<	SLC25A17	>	3 ps (owm P4)	XR_001010110.1	>	XPNPEP3	in	<	ST13
8	Unplaced	>	SH2B3	>	P5	XR_001010401.1	<	BRAP	in	<	ATXN2
	Rhinopithecus roxellana										
1	Chr. 12	<	WTIP	>	3 ps	XR_747609.2	<	PDCD2L	in	<	UBA2
2	Chr. 19	<	C1QL1	>	3 like	XM_030922644.1	>	NMT1	in	<	DCAKD
4	Chr. 10	>	WASHC4	>	3 ps	XR_004059464.1	<	NUAK1	in	<	APPL2
5	Chr. 5	<	ATP10A	>	3 ps	XR_004057469.1	>	GABRB5	in	<	GABRB3
6	Chr. 13	<	SLC25A17	>	3 ps	XR_004052498.1	>	XPNPEP3	in	<	ST13
	Aotus nancymaae										
3	Unplaced	>	EGF	>	3 ps	XR_001104807.2	>	ENPEP	in	<	ELOVL6
	Saimiri boliviensis										
6	Unplaced	<	RPH3a	>	3 like	XM_039472141.1	>	RPL6	in	<	PTPN11
9	Unplaced	<	MTERF1	>	3 like	XM_039473093.1	<	CYP51A1	in	>	AKAP9
15	Unplaced	>	LRRFIP2	>	3 like	XM_039461847.1	>	EPM2AIP1	in	<	MLH1
19	Unplaced	<	KIAA1586	>	3 like	XM_039467413.1	>	DST	in	<	BEND6
	Carlito syrichta										
1	Unplaced	>	SVOPL	>	3 ps	XR_504221.2	End		in	<	Trim24
6	Unplaced	>	SPINK4	>	3 ps	XM_008074487.1	>	CHMP5	in	<	BAG1

in - located in introns.

10.3389/fmicb.2023.1213685

TABLE 4 IFITM retrogenes present in more than one primate species.

No.	Species and location	Direction	Flanking gene	Direction	Annotated	Accession number	Direction	Flacking gene	In intron	Direction	Gene
1	Chr. 4	<	EHPA5	>	IFITM3P1 pseudo	NG_006204.1	<	CENPC	x		
2	Chr. 12	<	AMN1	>	IFITM3P2 pseudo	NG_006205.3	>	RESF1	x		
3	Chr. 6	<	SPDEF	<	IFITM3P3 pseudo	NG_006229.1	<	ILRUN	x		
4	Chr. 7	<	NUPR2	>	IFITM3P4 pseudo	NG_006223.3	<	ZNF479	x		
5	Chr. 12	>	SH2B3	>	IFITM3P5 pseudo	NG_006225.2	<	BRAP	in	<	ATXN2
6	Chr. 12	<	AMIGO2	<	IFITM3P6 pseudo	NG_006230.2	<	RPAP3	in	>	PCED1B
7	Chr. 1	>	SYF2	>	IFITM3P7 pseudo	NG_006227.1	>	RUNX3	x		
8	Chr. 8	>	CDH7	>	IFITM3P8	NG_006224.1	>	CLVS1	x		
9	Chr. 2	>	PAPLOG	>	IFITM3P9	NG_006228.3	>	BCL11A	x		
10	Chr. 6	>	HLA-F	<	IFITM4p	NR_001590.1	>	HLA-G	x		
11	Chr.8	>	YTHDF3	>	IFITM8P pseudo	NG_005307.4	>	BHLHE22	x		
12	Chr. 11	>	MYEOV	<	IFITM9P pseudo	NG_006210.1	>	CCND1	x		
	Pan paniscus										
1	Chr. 12	<	AMN1	>	IFITM3P2 pseudo	XM_003813732.5	>	RESF1	x		
2	Chr. 12	>	SH2B3	>	IFITM3P5 pseudo	XM_034934516.1	<	BRAP	in	<	ATXN2
3	Chr. 2a	>	PAPLOG	>	IFITM3P9	XM_034953202.1	>	BCL11A	x		
4	Chr. 6	>	HLA-F related	<	IFITM4p	XM_034961680.1	>	HLA-G	x		
5	Chr.8	>	YTHDF3	>	IFITM8P pseudo	XM_034966156.1	>	BHLHE22	x		
	Pan troglodytes										
1	Chr. 4	<	EHPA5	>	IFITM3P1 pseudo	XR_001716631.2	<	CENPC	x		
2	Chr. 12	<	AMN1	>	IFITM3P2 pseudo	XM_003952225.4	>	RESF1	x		
3	Chr. 7	<	NUPR2	>	IFITM3P4 pseudo	XR_169790.4	<	ZNF479	x		
4	Chr. 2a	>	PAPLOG	>	IFITM3P9	XR_001715794.1	>	BCL11A	x		
5	Chr. 6	>	HLA-F related	<	IFITM4p	XR_002944366.1	>	HLA-G	x		
	Gorilla gorilla										
1	Chr. 12	<	AMN1	>	IFITM3P2 pseudo	XM_004052942.3	>	RESF1	x		
2	Chr. 2a	>	PAPLOG	>	IFITM3P9	XR_002004539.2	>	BCL11A	x		
	Pongo abelii										
1	Chr. 12	<	AMN1	>	IFITM3P2 pseudo	XR_656249.2	>	RESF1	x		
2	Chr. 7	<	NUPR2	>	IFITM3P4 pseudo	XR_002913425.1	<	ZNF479	x		
3	Chr. 2a	>	PAPLOG	>	IFITM3P9	XR_654203.1	>	BCL11A	x		

(Continued)

TABLE 4 (Continued)

No.	Species and location	Direction	Flanking gene	Direction	Annotated	Accession number	Direction	Flacking gene	In intron	Direction	Gene
4	Chr. 11	>	MYEOV	<	IFITM9P pseudo	XR_656019.2	>	CCND1	x		
	Nomascus leucogenys										
1	Chr. 14	>	PAPLOG	>	IFITM3P9	XR_001114400.1	>	BCL11A	x		
2	Chr. 22a	<	SPDEF	<	IFITM3P3 pseudo	XM_030803316.1	<	ILRUN	x		
5	Chr. 17	>	CHCHD2	>	IFITM3P4 pseudo	XR_004026378.1	>	VOPP1	x		
	Macaca mulatta										
1	Chr. 16	<	CTC1	<	3 like 8 (owm P1)	XM_001112566.4	>	RANGRF	in	>	PFAS
2	Chr. 1	>	ATP5PB	>	3 like (owm P2)	XM_001106166.4	>	RAP1A	in	<	TMIGD3
3	Chr. 7	>	MAPK1IP1L	<	3 like (own P3)	XM_001088204.4	>	LGALS3	x		
4	Chr. 10	<	SLC25A17	>	3 ps (owm P4)	XR_001447798.2	>	XPNPEP3	in	<	ST13
	Cercocebus atys										
2	Unplaced	>	ATP5PB	>	3 like (owm P2)	XR_001017992.1	>	RAP1A	in	<	TMIGD1
4	Unplaced	<	CTC1	<	3 like 8 (owm P1)	XM_012051598.1	>	RANGRF	in	>	PFAS
6	Unplaced	<	SLC25A17	>	3 ps (owm P4)	XR_001010110.1	>	XPNPEP3	in	<	ST13
7	Unplaced	>	MAPK1IP1L	<	3 like (own P3)	XM_012061884.1	>	LGALS3	x		
10	Unplaced	<	CHRM3	>	1 like (owm P5)	XR_001011033.1	>	ZP4	x		
	Rhinopithecus roxellana										
8	Chr. 8	<	CHRM3	>	1 like (owm P5)	XR_750288.2	>	ZP4	x		
	Aotus nancymaae										
6	Unplaced	<	AGO2	<	3 ps (nwm P)	XR_002478805.1	<	PTK2	x		
	Saimiri boliviensis										
18	Unplaced	<	AGO2	<	3 ps (nwm P)	XM_039464430.1	<	PTK2	x		

in - located in intron.
x - not located in intron.

10.3389/fmicb.2023.1213685

TABLE 5 Retrogene features of selected primate IFITM retrogenes.

Species	Accession number	Lack of intron	Poly A signal	Poly A tail	Target site duplications (TSDs)	Premature STOP
Otolemur garnettii	XR_001161573.1	Yes	Yes	Yes	Yes	No
Carlito syrichta	XR_504221.2	Yes	Yes	Yes	Yes	No
	XM_008052862.1	?*	Yes	Yes	Yes	Yes/No*
Saimiri boliviensis	XM_039468571.1	Yes	Yes	Yes	Yes	No
	XM_039478903.1	Yes	Yes	Yes	Yes	No
Aotus nancymaae	XR_002477520.1	Yes	Yes	Yes	Yes	Yes
	XR_001106643.2	Yes	Yes	Yes	Yes	Yes
Rhinopithecus roxellana	XR_748909.2	Yes	Yes	Yes	Yes	No
	XM_030922644.1	Yes	Yes	Yes	Yes	No
Cercocebus atys	XR_001017992.1	Yes	Yes	Yes	Yes	No
	XR_001011714.1	Yes	Yes	Yes	Yes	No
Macaca mulatta	XM_001112566.4	Yes	Yes	Yes	Yes	No
	XR_001438791.2	Yes	Yes	Yes	Yes	Yes
Nomascus leucogenys	XR_004026378.1	Yes	Yes	Yes	Yes	No
	XR_004027821.1	Yes	Yes	Yes	Yes	No
Pongo abelii	XR_002913425.1	Yes	Yes	Yes	Yes	No
	XR_656019.2	Yes	Yes	Yes	Yes	No
Gorilla gorilla	XR_002005707.2	Yes	Yes	Yes	Yes	No
	XM_004052942.3	Yes	Yes	Yes	Yes	No
Pan troglodytes	XR_002913425.1	Yes	Yes	Yes	Yes	No
	XR_169790.4	Yes	Yes	Yes	Yes	Yes
Pan paniscus	XM_034961680.1	Yes	Yes	Yes	Yes	Yes
	XM_034966156.1	Yes	Yes	Yes	Yes	Yes
Homo sapiens	NG_006210.1	Yes	Yes	Yes	Yes	Yes
	NG_006230.2	Yes	Yes	Yes	Yes	Yes

Shown are the features of retrogenes.*Not distinguishable if short part of intron or insert.

transcript of a canonical IR-pIFITM3/nwm/pro has been constantly retrotranspositioned by class 1 transposable elements, building the retro(pseudo)genes. The unique species-specific pattern was caused by constant pseudogenization and/or loss of the IFITM retro(pseudo)genes (Figure 6). The reason for the preferential integration of IR-pIFITM3/nwm/pro transcripts remains unclear but enrichment of retro(pseudo)gene mRNAs was observed in LINE-1 ribonucleoproteins (mediating retrotransposition) (Mandal et al., 2013). We hypothesize that the high abundance of their mRNAs in the germline might have favored their binding and retrotransposition (Zhang et al., 2003, 2004). This might be caused either by interferon induction (Friedman et al., 1984) as an innate immunity response to specific pathogens or their general expression in germline cells, which has been shown for mouse ifitms (Tanaka and Matsui, 2002). However, an unknown mechanisms could have also played a role since LINE-1 RNA is preferentially retrotranspositioned compared to other mRNAs (Esnault et al., 2000; Kulpa and Moran, 2006). It is also possible that other mRNA properties play a role similar to the poly A tail requirement for retrotransposition (Doucet et al., 2015).

The maintenance of a high number of such retro(pseudo)genes in higher primate species is also unclear. Indeed, in some cases, it could have compensated or caused the loss of the canonical IFITMs (e.g., *Piliocolobus tephrosceles*). In other cases, it might represent an additional selective advantage by their expression in response to a viral infection. This was recently shown for human IFITM4P, a retropseudogene, which is not coding for a protein (Xiao et al., 2021). However, the rate of retrotransposition and therefore the emergence of retro(pseudo)genes could be simply exceeding the rate at which pseudogenization and gene loss occur in higher primates.

In conclusion, we found evidence for concerted evolution and birth-and-death evolution model for the canonical cluster IR-pIFITMs. For the IFITM retro(pseudo)genes, we propose a new hypothesis for their origin and pattern (Figure 6) through a third mechanism of evolution, similar to the birth-and-death model of evolution, but via a transposable element mechanism leading to IFITM retro(pseudo)genes. Primate IFITMs were thus the result of a mixed evolutionary process combining three different mechanisms.

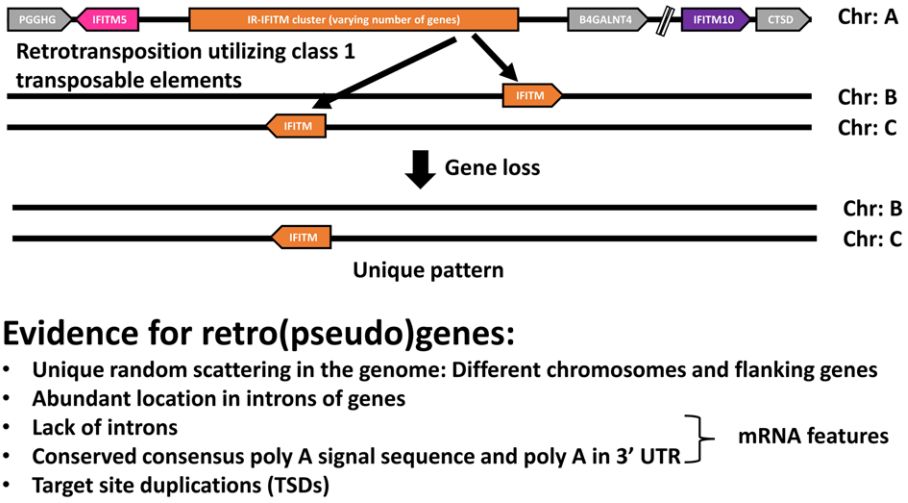

FIGURE 6
Hypothesis for the origin and pattern of the IFITM retrogenes. Schematic representation of our hypothesis: the transcript of a canonical IR-pIFITM3/nwm/pro is constantly retrotranspositioned by class 1 transposable elements originating the retro(pseudo)genes. The unique pattern of each species is caused by constant pseudogenization and loss of the retro(pseudo)genes. Evidence supporting the hypothesis are listed.

4. Materials and methods

4.1. Gene synteny analysis

Primate IFITM sequences were retrieved from https://www.ncbi.nlm.nih.gov/; BLASTn analysis ensured that all available sequences per species were included. Accession numbers of all retrieved sequences are found in Supplementary Table S1. The NCBI Genomic Data Viewer[1] was used to determine the genomic localization and orientation of the IFITMs in the 26 analyzed primate species. The primate phylogeny was obtained using Timetree.org (Kumar et al., 2022).

4.2. Sequence alignments

Sequences were initially aligned using MEGA11 (Tamura et al., 2021) and MUSCLE algorithm (Edgar, 2004). Alignments were then visually inspected and manually corrected in BioEdit (Hall, 1999).

4.3. Phylogenetic analysis

For AA sequences, the evolutionary history was inferred using the Maximum Likelihood (ML) method. The percentage of trees in which the associated taxa clustered together is shown next to the branches and was obtained by conducting 1,000 bootstrap replicates. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Joining and BioNJ algorithms to a matrix of pairwise distances estimated using the JTT model (Jones et al., 1992), and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites [5 categories (+G)]. The trees were drawn to scale, with branch lengths measured in the number of substitutions per site. All positions with less than 95% site coverage were eliminated, i.e., fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position (partial deletion option). Analyses were conducted in MEGA11 (Tamura et al., 2021).

4.4. Sequence logos

For the generation of the sequence logos, WebLogo 3[2] was used (Crooks et al., 2004). Alignments (Supplementary Figures S6–S8) were used as input.

4.5. Transposable element features analysis

We considered random unique localization, localization in introns of other genes, lack of introns, conserved poly A signal (AATAAA), poly A tail start, target-site duplications (5′ and 3′ UTR) (Kaessmann et al., 2009), and full coding sequences as features for retro(pseudo) genes. Localization (random, unique, in introns) was obtained from our synteny data. For the other features, we analyzed the genomic sequence of the IFITMs 200 bp upstream of the canonical start codon and 400 bp downstream of the canonical stop codon. Lack of introns was obtained from the annotations found at NCBI and genomic sequence. Sequences were manually inspected for canonical start codon, canonical stop codon, premature stop codon, poly A signal (AATAAA), poly A start and TSDs.

1 https://www.ncbi.nlm.nih.gov/genome/gdv/

2 https://weblogo.threeplusone.com/

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding authors.

Author contributions

LS: conceptualization, data curation, formal analysis, writing—original draft, writing—review and editing. JA: funding acquisition, writing—review and editing. H-MB: supervision, writing—review and editing, funding acquisition. PE: conceptualization, writing—review and editing, funding acquisition. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Fundação para a Ciência e Tecnologia (FCT) - Portugal, supported the Assistant Researcher grant of JA (CEECIND/00078/2017) and the Principal Researcher grant of PE (CEECIND/01495/2020). This work was co-funded by the project NORTE-01-0246-FEDER-000063, supported by Norte Portugal Regional Operational Programme (NORTE2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF). H-MB acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG) (BA-6820/1-1). H-MB and JA acknowledge the project-related personal exchange (PPP) program of the FCT/German Academic Exchange Service (DAAD) (57518622). This work was supported by the young Society for Virology (jGfV) lab rotation scholarship granted to LS.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2023.1213685/full#supplementary-material

References

Bailey, C. C., Kondur, H. R., Huang, I. C., and Farzan, M. (2013). Interferon-induced transmembrane protein 3 is a type II transmembrane protein. *J. Biol. Chem.* 288, 32184–32193. doi: 10.1074/jbc.M113.514356

Bailey, C. C., Zhong, G., Huang, I. C., and Farzan, M. (2014). IFITM-family proteins: the Cell's first line of antiviral defense. *Annu. Rev. Virol.* 1, 261–283. doi: 10.1146/annurev-virology-031413-085537

Benfield, C. T., MacKenzie, F., Ritzefeld, M., Mazzon, M., Weston, S., Tate, E. W., et al. (2020). Bat IFITM3 restriction depends on S-palmitoylation and a polymorphic site within the CD225 domain. *Life Sci. Alliance* 3:e201900542. doi: 10.26508/lsa.201900542

Compton, A. A., Roy, N., Porrot, F., Billet, A., Casartelli, N., Yount, J. S., et al. (2016). Natural mutations in IFITM3 modulate post-translational regulation and toggle antiviral specificity. *EMBO Rep.* 17, 1657–1671. doi: 10.15252/embr.201642771

Côrte-Real, J. V., Baldauf, H. M., Abrantes, J., and Esteves, P. J. 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;(October 2020):79–81. doi: 10.1016/j.molimm.2021.01.025

Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004

Diamond, M. S., and Farzan, M. (2013). The broad-spectrum antiviral functions of IFIT and IFITM proteins. *Nat. Rev. Immunol.* 13, 46–57. doi: 10.1038/nri3344

Doucet, A. J., Wilusz, J. E., Miyoshi, T., Liu, Y., and Moran, J. V. (2015). A 3' poly(a) tract is required for LINE-1 Retrotransposition. *Mol. Cell* 60, 728–741. doi: 10.1016/j.molcel.2015.10.012

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Esnault, C., Maestre, J., and Heidmann, T. (2000). Human LINE retrotransposons generate processed pseudogenes. *Nat. Genet.* 24, 363–367. doi: 10.1038/74184

Friedman, R. L., Manly, S. P., McMahon, M., Kerr, I. M., and Stark, G. R. (1984). Transcriptional and posttranscriptional regulation of interferon-induced gene expression in human cells. *Cells* 38, 745–755. doi: 10.1016/0092-8674(84)90270-8

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.

Hanagata, N., Li, X., Morita, H., Takemura, T., Li, J., and Minowa, T. (2011). Characterization of the osteoblast-specific transmembrane protein IFITM5 and analysis of IFITM5-deficient mice. *J. Bone Miner. Metab.* 29, 279–290. doi: 10.1007/s00774-010-0221-0

Hedjazi, G., Guterman-Ram, G., Blouin, S., Schemenz, V., Wagermaier, W., Fratzl, P., et al. (2022). Alterations of bone material properties in growing Ifitm5/BRIL p.S42 knock-in mice, a new model for atypical type VI osteogenesis imperfecta. *Bone* 162:116451. doi: 10.1016/j.bone.2022.116451

Hickford, D., Frankenberg, S., Shaw, G., and Renfree, M. B. (2012). Evolution of vertebrate interferon inducible transmembrane proteins. *BMC Genomics* 13:155. doi: 10.1186/1471-2164-13-155

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* 8, 275–282. doi: 10.1093/bioinformatics/8.3.275

Kaessmann, H., Vinckenbosch, N., and Long, M. (2009). RNA-based gene duplication: mechanistic and evolutionary insights. *Nat. Rev. Genet.* 10, 19–31. doi: 10.1038/nrg2487

Kulpa, D. A., and Moran, J. V. (2006). Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. *Nat. Struct. Mol. Biol.* 13, 655–660. doi: 10.1038/nsmb1107

Kumar, S., Suleski, M., Craig, J. M., Kasprowicz, A. E., Sanderford, M., Li, M., et al. (2022). TimeTree 5: An expanded resource for species divergence times. *Mol. Biol. Evol.* 39:msac174. doi: 10.1093/molbev/msac174

Liao, Y., Goraya, M. U., Yuan, X., Zhang, B., Chiu, S. H., and Chen, J. L. (2019). Functional involvement of interferon-inducible transmembrane proteins in antiviral immunity. *Front. Microbiol.* 10:1097. doi: 10.3389/fmicb.2019.01097

Liu, P., Zhang, Y., Zhang, S., Peng, C., Yang, W., Li, X., et al. (2021). Integrative overview of IFITMs family based on bioinformatics analysis. *Intractable Rare Dis. Res.* 10, 165–172. doi: 10.5582/irdr.2021.01041

Luan, D. D., Korman, M. H., Jakubczak, J. L., and Eickbush, T. H. (1993). Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cells* 72, 595–605. doi: 10.1016/0092-8674(93)90078-5

Mandal, P. K., Ewing, A. D., Hancks, D. C., and Kazazian, H. H. Jr. (2013). Enrichment of processed pseudogene transcripts in L1-ribonucleoprotein particles. *Hum. Mol. Genet.* 22, 3730–3748. doi: 10.1093/hmg/ddt225

Nei, M., Gu, X., and 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. doi: 10.1073/pnas.94.15.7799

Nei, M., and Rooney, A. P. (2005). Concerted and birth-and-death evolution of multigene families. *Annu. Rev. Genet.* 39, 121–152. doi: 10.1146/annurev.genet.39.073003.112240

Rahman, K., and Compton, A. A. (2021). The indirect antiviral potential of Long noncoding RNAs encoded by IFITM pseudogenes. *J. Virol.* 95, e00680–e00621. doi: 10.1128/JVI.00680-21

Siegrist, F., Ebeling, M., and Certa, U. (2011). The small interferon-induced transmembrane genes and proteins. *J. Interf. Cytokine Res.* 31, 183–197. doi: 10.1089/jir.2010.0112

Tamura, K., Stecher, G., and Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* 38, 3022–3027. doi: 10.1093/molbev/msab120

Tanaka, S. S., and Matsui, Y. (2002). Developmentally regulated expression of mil-1 and mil-2, mouse interferon-induced transmembrane protein like genes, during formation and differentiation of primordial germ cells. *Mech. Dev.* 119, S261–S267. doi: 10.1016/S0925-4773(03)00126-6

Troskie, R. L., Faulkner, G. J., and Cheetham, S. W. (2021). Processed pseudogenes: a substrate for evolutionary innovation: Retrotransposition contributes to genome evolution by propagating pseudogene sequences with rich regulatory potential throughout the genome. *BioEssays* 43:e2100186. doi: 10.1002/bies.202100186

Wilkins, J., Zheng, Y. M., Yu, J., Liang, C., and Liu, S. L. (2016). Nonhuman primate IFITM proteins are potent inhibitors of HIV and SIV. *PLoS One* 11:e0156739. doi: 10.1371/journal.pone.0156739

Xiao, M., Chen, Y., Wang, S., Liu, S., Rai, K. R., Chen, B., et al. (2021). Long noncoding RNA IFITM4P regulates host antiviral responses by acting as a competing endogenous RNA. *J. Virol.* 95:e0027721. doi: 10.1128/JVI.00277-21

Yanez, D. C., Ross, S., and Crompton, T. (2020). The IFITM protein family in adaptive immunity. *Immunology* 159, 365–372. doi: 10.1111/imm.13163

Zhang, Z., Carriero, N., and Gerstein, M. (2004). Comparative analysis of processed pseudogenes in the mouse and human genomes. *Trends Genet.* 20, 62–67. doi: 10.1016/j.tig.2003.12.005

Zhang, Z., Harrison, P. M., Liu, Y., and Gerstein, M. (2003). Millions of years of evolution preserved: a comprehensive catalog of the processed pseudogenes in the human genome. *Genome Res.* 13, 2541–2558. doi: 10.1101/gr.1429003

Zhang, Z., Liu, J., Li, M., Yang, H., and Zhang, C. (2012). Evolutionary dynamics of the interferon-induced transmembrane gene family in vertebrates. *PLoS One* 7:e49265. doi: 10.1371/journal.pone.0049265

Zhao, X., Li, J., Winkler, C. A., An, P., and Guo, J. T. (2018). IFITM genes, variants, and their roles in the control and pathogenesis of viral infections. *Front. Microbiol.* 9:3228. doi: 10.3389/fmicb.2018.03228

5. Publication 2)

Evolutionary and functional characterization of lagomorph guanylate-binding proteins: a story of gain and loss and shedding light on expression, localization and innate immunity-related functions

Luca Schelle1[†], João Vasco Côrte-Real[1,2,3,4†], Sharmeen Fayyaz[5,6], Augusto del Pozo Ben[1], Margarita Shnipova[1], Moritz Petersen[5], Rishikesh Lotke[5], Bhavna Menon[1], Dana Matzek[7], Lena Pfaff[7], Ana Pinheiro[2,4], João Pedro Marques[2,4,8], José Melo-Ferreira[2,3,4], Bastian Popper[7], Pedro José Esteves[2,3,4,9], Daniel Sauter[5], Joana Abrantes[2,3,4*], Hanna-Mari Baldauf[1*]

[1]Max von Pettenkofer Institute and Gene Center, Virology, National Reference Center for Retroviruses, Faculty of Medicine, LMU München, Munich, Germany

[2]CIBIO-InBIO, Research Center in Biodiversity and Genetic Resources, University of Porto, Vairão, Portugal

[3]Department of Biology, Faculty of Sciences, University of Porto, Porto, Portugal

[4]BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Vairão, Portugal

[5]Institute for Medical Virology and Epidemiology of Viral Diseases, University Hospital Tübingen, Tübingen, Germany

[6]National Institute of Virology, International Center of Chemical and Biological Sciences, University of Karachi, Karachi, Pakistan

[7]Biomedical Center (BMC), Core facility Animal Models (CAM), Faculty of Medicine, LMU München, Munich, Germany

[8]ISEM, University of Montpellier, CNRS, EPHE, IRD, Montpellier, France

[9]CITS - Center of Investigation in Health Technologies, CESPU, Gandra, Portugal

[†]These authors have contributed equally to this work and share first authorship

This article was published 2024 in Frontiers in Immunology and is available under: doi.org/10.3389/fimmu.2024.1303089

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 29 January 2024
DOI 10.3389/fimmu.2024.1303089

OPEN ACCESS

EDITED BY
Shibnath Mazumder,
University of Delhi, India

REVIEWED BY
Gregory D. Wiens,
Agricultural Research Service (USDA), United States
Junfa Yuan,
Huazhong Agricultural University, China

*CORRESPONDENCE
Joana Abrantes
jabrantes@cibio.up.pt
Hanna-Mari Baldauf
baldauf@mvp.lmu.de

†These authors have contributed equally to this work and share first authorship

RECEIVED 27 September 2023
ACCEPTED 04 January 2024
PUBLISHED 29 January 2024

CITATION
Schelle L, Côrte-Real JV, Fayyaz S, del Pozo Ben A, Shnipova M, Petersen M, Lotke R, Menon B, Matzek D, Pfaff L, Pinheiro A, Marques JP, Melo-Ferreira J, Popper B, Esteves PJ, Sauter D, Abrantes J and Baldauf H-M (2024) Evolutionary and functional characterization of lagomorph guanylate-binding proteins: a story of gain and loss and shedding light on expression, localization and innate immunity-related functions. *Front. Immunol.* 15:1303089. doi: 10.3389/fimmu.2024.1303089

Evolutionary and functional characterization of lagomorph guanylate-binding proteins: a story of gain and loss and shedding light on expression, localization and innate immunity-related functions

Luca Schelle[1†], João Vasco Côrte-Real[1,2,3,4†], Sharmeen Fayyaz[5,6], Augusto del Pozo Ben[1], Margarita Shnipova[1], Moritz Petersen[5], Rishikesh Lotke[5], Bhavna Menon[1], Dana Matzek[7], Lena Pfaff[7], Ana Pinheiro[2,4], João Pedro Marques[2,4,8], José Melo-Ferreira[2,3,4], Bastian Popper[7], Pedro José Esteves[2,3,4,9], Daniel Sauter[5], Joana Abrantes[2,3,4*] and Hanna-Mari Baldauf[1*]

[1]Max von Pettenkofer Institute and Gene Center, Virology, National Reference Center for Retroviruses, Faculty of Medicine, LMU München, Munich, Germany, [2]CIBIO-InBIO, Research Center in Biodiversity and Genetic Resources, University of Porto, Vairão, Portugal, [3]Department of Biology, Faculty of Sciences, University of Porto, Porto, Portugal, [4]BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Vairão, Portugal, [5]Institute for Medical Virology and Epidemiology of Viral Diseases, University Hospital Tübingen, Tübingen, Germany, [6]National Institute of Virology, International Center of Chemical and Biological Sciences, University of Karachi, Karachi, Pakistan, [7]Biomedical Center (BMC), Core facility Animal Models (CAM), Faculty of Medicine, LMU München, Munich, Germany, [8]ISEM, University of Montpellier, CNRS, EPHE, IRD, Montpellier, France, [9]CITS - Center of Investigation in Health Technologies, CESPU, Gandra, Portugal

Guanylate binding proteins (GBPs) are an evolutionarily ancient family of proteins that are widely distributed among eukaryotes. They belong to the dynamin superfamily of GTPases, and their expression can be partially induced by interferons (IFNs). GBPs are involved in the cell-autonomous innate immune response against bacterial, parasitic and viral infections. Evolutionary studies have shown that GBPs exhibit a pattern of gene gain and loss events, indicative for the birth-and-death model of evolution. Most species harbor large *GBP* gene clusters that encode multiple paralogs. Previous functional and in-depth evolutionary studies have mainly focused on murine and human GBPs. Since rabbits are another important model system for studying human diseases, we focus here on lagomorphs to broaden our understanding of the multifunctional GBP protein family by conducting evolutionary analyses and performing a molecular and functional characterization of rabbit GBPs. We observed that lagomorphs lack *GBP3, 6* and *7*. Furthermore, *Leporidae* experienced a loss of *GBP2*, a unique duplication of *GBP5* and a massive expansion of *GBP4*. Gene expression analysis by reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) and transcriptome data revealed that leporid *GBP* expression varied across tissues. Overexpressed rabbit GBPs localized either

uniformly and/or discretely to the cytoplasm and/or to the nucleus. *Oryctolagus cuniculus* (oc)GBP5L1 and rarely ocGBP5L2 were an exception, colocalizing with the trans-Golgi network (TGN). In addition, four ocGBPs were IFN-inducible and only ocGBP5L2 inhibited furin activity. In conclusion, from an evolutionary perspective, lagomorph GBPs experienced multiple gain and loss events, and the molecular and functional characteristics of ocGBP suggest a role in innate immunity.

KEYWORDS

GBP, evolution, innate immunity, antiviral proteins, cross-species conservation, lagomorphs, *Oryctolagus cuniculus*

1 Introduction

The survival of uni- and multicellular organisms depends on their ability to detect and eliminate invading pathogens (1), relying thereby on basic forms of immunity, such as Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) in bacteria, to complex immune systems in mammals (1). Upon infection, type I and type II IFN are produced, resulting in the expression of numerous IFN-stimulated genes (2). Several of these genes enhance the efficacy of cell-autonomous immunity (3, 4), including guanylate-binding proteins (GBPs), which are specialized for host defense against intracellular pathogens ranging from bacteria to viruses (3, 5).

The GBP family belongs to the large dynamin GTPase superfamily, which includes myxoma resistance (Mx) proteins, immunity-related GTPases, and the very large IFN-inducible GTPases. These proteins share structural and biochemical similarities such as the GTPase domain (6, 7). Mammalian GBP proteins vary in size from ~65 to 73 kDa and are mainly localized to the cytosol (5, 8). They possess a large GTPase domain at the N-terminus representing motifs for guanine nucleotide binding, specifically GxxxxGK and x(V/L)RD (9–13), followed by a middle domain and the GTPase effector domain at the C-terminus (14). Human GBP1, 2 and 5 also harbor a CaaX motif at the C-terminus, which is important for isoprenylation and enables membrane anchoring (14).

The human genome encodes seven *GBPs* (*GBP1-7*) in a single cell cluster (15). It has been described that each *GBP* originated from the same common ancestor. Following the first duplication round, one gene evolved a CaaX motif, giving origin to modern day human *GBP1/2/3/5*. The second gene gave rise to human *GBP4/6/7*, which are characterized by the L182V replacement in the GTP-binding motif (TLRD) (15). *GBP1*, *2* and 3 are closely related members, with human *GBP1* and *3* sharing 87% amino acid similarities, while human *GBP2* shares 77% and 76% identity with human *GBP1* and *3*, respectively (15). On the other *GBP* branch, the most closely related genes are *GBP4* and *GBP7*, sharing 81% identity (15). We have recently studied the evolution of GBPs in primates (16) and found that *GBP3* evolved from a duplication of *GBP1* only in *Simiiformes*, while the duplication of *GBP4* gave rise to *GBP7*, which is only present in primates (16). In contrast, *GBP4* and *GBP5* are no longer present in the genomes of Old World monkeys (16). We have further extended evolutionary analyses to muroid *GBPs*, which are separated into two gene clusters and proposed a new nomenclature, as primate *GBP1*, *GBP3* and *GBP7* are absent from muroid genomes (17). In contrast, murine *Gbp2*, *Gbp5* and *Gbp6* might be true orthologs of their primate counterparts. Orthologs are genes in different species that evolved from a common ancestral gene through speciation and may retain the same function throughout evolution. Identification of orthologs is a critical process for reliable prediction of gene function in newly sequenced genomes. More importantly, four *Gbps* are exclusive to muroids, but absent from *Mus musculus* (17). Thus, in line with the proposed birth-and-death model of evolution, our analyses revealed that GBPs underwent duplications, deletions, and neofunctionalizations, raising even more awareness to conduct in-depth evolutionary analyses for GBPs of different species. Beyond primates and muroids, information on the evolution and function of GBPs is scarce. In addition to humans, the role of GBPs in innate immunity has been described in plants, invertebrates, teleosts, mice, pigs, and *Tupaia* (14).

Within Lagomorpha, there are two families, *Leporidae* (hares and rabbits) and *Ochotonidae* (pikas), which diverged approximately ~37 million years ago (MYA) (18). The *Ochotonidae* family is restricted to the genus *Ochotona*, which is further divided into four subgenera (*Pika, Logotona, Conothoa* and *Ochotona*) and the divergence time between these subgenera is ~7 to 14 MYA (13, 19–21). The *Leporidae* family is divided into two groups, hares and rabbits, which diverged around 12 MYA (22). The hare group only contains one genus, *Lepus*, while the rabbit group comprises ten distinct genera (23, 24). The genus *Oryctolagus* is one of the most studied due to its importance in the Mediterranean ecosystem as prey for endangered species and also for its importance in biomedical research, particularly in immunology and infectious diseases (24, 25). Furthermore, the genetic diversity of innate immunity genes between rabbits and humans is lower than between mice and humans, suggesting that the European rabbit might be a better model to study such genes (26).

In this study, we aimed to characterize the evolutionary history and intrinsic functions of lagomorph GBPs, going beyond their description in murines and primates, to broaden the understanding of the GPB family. For this, we combined evolutionary analyses with *in vitro* assays, shedding light on species-specific mRNA and protein expression profiles and evolutionary patterns. In addition, we wanted to establish links to cell-autonomous innate immunity functions of GBPs.

2 Results

2.1 Absence of *GBP3/6/7* in lagomorphs; loss of *GBP2*, unique duplication of *GBP5* and expansion of *GBP4s* in leporids

We analyzed 204 *GBP* sequences belonging to muroids, primates, lagomorphs, *Tupaia*, elephant and chicken. Before conducting the evolutionary analysis, the *GBP*s alignment (see Supplementary Table 1 for accession numbers and see Supplementary Data for *GBP* alignment) was screened for recombination and gene conversion using GARD (Genetic Algorithm for Recombination Detection; 27). No gene conversion or recombination events were detected (data not shown). Thirty-one sequences were excluded because they did not encode a functional protein or the sequence was truncated (see Supplementary Table 1 for accession numbers).

The ML phylogenetic tree showed that lagomorphs do not have *GBP3*, *6* and *7*, as none of the sequences grouped with the corresponding human counterpart (Figure 1). *Ochotonidae* appear to harbor one copy of *GBP2* in their genome (Figure 1). Interestingly, *GBP2* is not present in *Oryctolagus cuniculus* nor could be found in the genome of *Lepus* (data not shown), indicating that *GBP2* was lost in the ancestor of *Leporidae* at least 12 MYA (22). Moreover, lagomorphs diverged from the common ancestor with rodents about 62-100 MYA (29, 30), which may explain why muroid *Gbp2* and *Ochotonidae GBP2* cluster together and not with primate *GBP2* despite the low bootstrap value (<0.6) (Figure 1). This group was named as *Ochotonidae GBP2* because in a previous study, muroid *Gbp2* clustered with primate *GBP2* (17). A summary of the gain and loss of *GBPs* in lagomorph is presented in Figure 2A. *GBP1* is present in *Leporidae* and *Ochotonidae*, with one copy in each species, similar to primates (Figure 1). The *GBP5* cluster was extremely robust with a bootstrap value of 1.00 (Figure 1). Lagomorph *GBP5* was present in all species with *Ochotonidae* having only one copy, whereas *Oryctolagus cuniculus* had two copies. Moreover, this duplication was also present in *Lepus* (data not shown), suggesting a duplication of *GBP5* after the split of *Ochotonidae* and *Leporidae* and before the split of *Lepus* and *Oryctolagus* (~12 MYA; 22; ~37 Mya; 18). A major cluster, designated as *GBP4*, underwent an expansion in *Oryctolagus cuniculus* with seven copies of the gene (*GBP4 XM_017345575* was not included in the analysis) (Figure 1), while *Ochotona curzoniae* and *princeps* presented two copies. From Maximum Likelihood (ML) tree, *Oryctolagus cuniculus GBP7* did not cluster with lagomorph *GBP4* but was at a basal position of the cluster of primate *GBP4* and *7*. However, the low bootstrap value (<0.6) indicated that the phylogenetic relationship could not be fully resolved. Despite this, the nomenclature of this gene might be incorrect since *GBP7* is only present in primates (16, 17) and it did not cluster with primate *GBP7* in the ML tree (Figure 1). As such, we designated it *ocGBP4*; however, throughout the manuscript we named it *ocGBP4L6* (locus 6). No *GBP6* could be found in lagomorphs, as no lagomorph *GBP* clustered with primate and muroid *GBP6* (Figure 1). The most likely explanation is that *GBP6* was deleted from the lagomorph genome after the split from rodents since it is present in rodents. One might speculate that the expansion of *GBP4* in lagomorphs could be a compensation mechanism for the loss of *GBP6*. Interestingly, a group with *GBP* sequences from both *Ochotona* species was found at a basal position from the *GBP4*, *6* and *7* group (Figure 1). The origin of this group was puzzling, and it could be explained by a duplication event of the ancestral gene of *GBP4/6/7* originating from this group in *Ochotonidae* which then underwent an accelerated mutation rate. We designated this group as *GBP4/6/7* (Figure 1). Based on the evolutionary analysis, we suggest a new nomenclature for genes that appeared to be misclassified (see Table 1).

Considering the synteny of the lagomorph *GBP* genes, the gene cluster was located in a single chromosome, similar to primates (15, 16) (Figure 2B). Both *Ochotona* species presented the same synteny (Figure 2B). In all three lagomorph species, the *GBP* gene cluster was flanked by *KYAT3* and *LRRC8B*, as described elsewhere (16, 17). In conclusion, lagomorph *GBP* genes showed patterns of gain and loss and shared similarities with primate and muroid *GBPs*. However, they evolved independently after the separation from other mammals.

2.2 Conserved *GBP*-specific motifs in the lagomorphs

In order to shed light on the protein structure of lagomorph GBPs, we analyzed GBP-specific motifs. Except for *Ochotona curzoniae* (XM_ 040998558), all *GBPs* share a GxxxxGK guanine nucleotide binding motif. The TLRD/TVRD motif, important for guanine base contact, is present in all *GBPs*, except for *Oryctolagus cuniculus (oc)GBP4 L3* (XM_017345575), which encodes a truncated GBP with only 129 amino acids (see Supplementary Table 2; see *GBP* alignment, Supplementary Data). Most of the *GBPs* in the main *GBP1/2/5* cluster possess a TLRD motif instead of a TVRD motif (15). We observed that lagomorph *GBPs* from the *GBP1/2/5* group contain the TLRD motif, while *GBP2* from *Ochotona princeps* (XM_004582068) harbors an AVRD motif instead (see Supplementary Table 2). Lagomorph *GBPs* from the major *GBP4/6/7* cluster possess a TVRD or an AVRD motif. An exchange of a threonine for an alanine has also been observed in rodents (9). Interestingly, *ocGBP4 L6* (XM_008264918) carries a cysteine instead of a threonine (CVRD) (see Supplementary Table 2). In summary, lagomorph GBPs have in general similar guanine nucleotide binding motifs and motifs for guanine base contact as described for other mammalian GBPs.

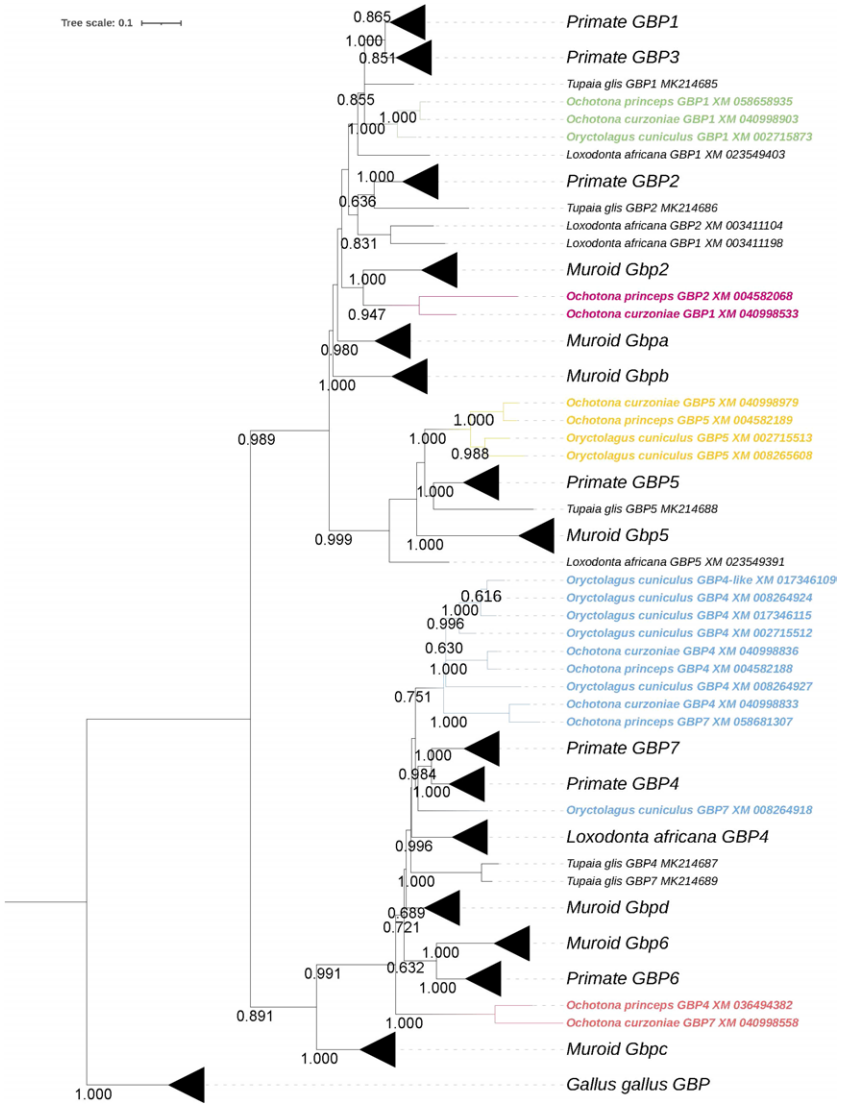

FIGURE 1
Phylogeny of lagomorph *GBPs*. The evolutionary history was inferred by using the Maximum Likelihood (ML) method. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The bootstrap values are shown next to the branches and only values >0.6 are shown (iTOL was used for tree visualization; 28).

2.3 Presence of different motifs with high probability of occurrence with a phylogeny-specific prenylation motif

As the European rabbit (*Oryctolagus cuniculus*) has been widely used as an animal model in biomedical research, we focused on the analysis of ocGBPs. We analyzed the ocGBP sequences for protein sequence motifs using the ProSite Scan tool (31–33). The results of the analysis (Supplementary Table 2) are summarized in Table 1 (Protein sequence motifs) and Table 2 (Protein sequence motifs with a high probability of occurrence). We observed that the G domain was the only conserved motif that was predicted with high confidence. With low confidence, the C-terminal glutamic acid-rich and nuclear localization signals were also found in the majority of analyzed ocGBPs (Table 1). In addition, protein sequence motifs were predicted with high occurrence, including sites of N-glycosylation, phosphorylation, ATP/GTP-binding motifs (P-loops), amidation, and N-myristoylation, which were found in varying numbers in the analyzed GBPs (for location, number and sequence motif see Supplementary Table 3). In all analyzed rabbit GBPs, the conserved P-loops were in accordance with the conserved G domain. Furthermore, prenyl group binding sites (CaaX motifs) were found only at the C-termini of ocGBP1 and ocGBP5 L2 (Table 2). However, we cannot rule out that alternative splicing might occur in rabbit GBPs and that it could impact some important motifs and dysregulate function. In summary, ocGBP paralogs have acquired individual protein sequence motifs but shared a highly conserved G domain and similar putative post-translational modification sites (PTMs).

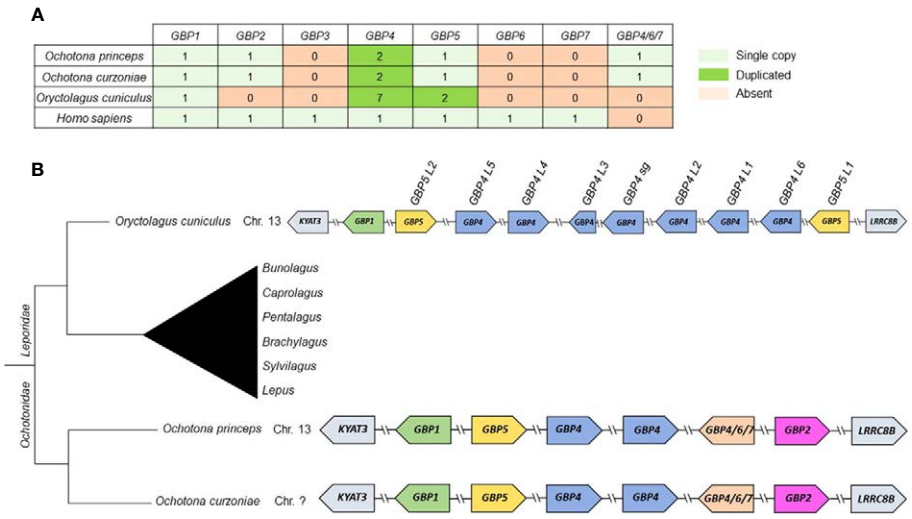

FIGURE 2
Summary of gain/loss and synteny of lagomorph *GBPs*. **(A)** Number of *GBP* copies for each lagomorph species. **(B)** Gene synteny of lagomorph GBPs. The gene synteny of the lagomorph GBPs is displayed for the analyzed lagomorph species (right). GBPs are colored following the grouping in the phylogenetic analyses (Figure 1). Arrows indicate gene orientation. Lagomorph phylogeny is shown (left) and the diagram is not to scale.

2.4 Conserved predicted tertiary structure of ocGBPs among the phylogenetic subgroups

Since structural data are available only for human GBPs, the tertiary structure of ocGBPs was predicted using AlphaFold (Figure 3). ocGBP4 L3 was excluded due to its length. We found that all ocGBPs shared a similar structure with hGBP1/2 and hGBP5, which have been crystallized without GTPase effector domain (GED) (PDB accession numbers: 6K1Z, 7E58, 7E59). ocGBP1 appeared to have the same architecture as hGBP1 (Figure 3). For ocGBP4L1/L2/L4/L5/sg/L6, we observed two additional short α-helices at the C-terminus (blue arrow in Figure 3), with ocGBP4L4 having an extended α13 helix (blue arrows in Figure 3). For ocGBP5, the large globular domain (LGD) and the middle domain (MD) appeared to be similar to those of hGBP5. The GED was predicted as an elongated α-helix in an "open" state conformation (yellow arrow Figure 3), as proposed for the active conformation of hGBP1 (34–38). In conclusion, the structure of the GBPs seem to be highly conserved in the LGDs and MDs, while the GED is variable between phylogenetic groups but specific within them.

2.5 Varying endogenous expression levels of ocGBPs

To gain more insight into ocGPBs, we examined their gene expression profiles. We established and validated RT-qPCRs for ocGBPs (data not shown) and ocFurin as control, and analyzed mRNA levels in various rabbit tissues, primary cells and cell lines, including overexpression of ocGBPs in the rabbit kidney cell line (RK13 cells; Figure 4A). We also analyzed the transcriptome of *Oryctolagus cuniculus* for the presence of the ocGBPs (Figure 4B) ocFurin was ubiquitously expressed in all samples analyzed. We detected a distinct pattern of GBP expression levels. mRNA levels for *ocGBP4L1/4L2/4L4/5L1* were lower in most tissues, primary

TABLE 1 Protein sequence motifs of ocGBPs.

ProSite Identifier	PS51715| G_GB1_RHD3	PS50313| GLU_RICH	PS50079| NLS_BP
Motif	GB1/RHD3-type guanine nucleotide-binding (G) domain	Glutamic acid enriched region	Bipartite nuclear localization signal
ocGBP1	Yes (high conf.)	No	Yes (2x, low conf.)
ocGBP4L1	Yes (high conf.)	Yes (low conf.)	Yes (2x, low conf.)
ocGBP4L2	Yes (high conf.)	Yes (low conf.)	Yes (1x, low conf.)
ocGBP4L3	Yes (high conf.)	No	No
ocGBP4L4	Yes (high conf.)	Yes (high conf.)	Yes (1x, low conf.)
ocGBP4L5	Yes (high conf.)	Yes (low conf.)	Yes (1x, low conf.)
ocGBP4sg	Yes (high conf.)	Yes (low conf.)	Yes (1x, low conf.)
ocGBP5L1	Yes (high conf.)	No	No
ocGBP5L2	Yes (high conf.)	Yes (low conf.)	No
ocGBP4L6	Yes (high conf.)	No	Yes (1x, low conf.)

conf., confidence.

TABLE 2 Protein sequence motifs with a high probability of occurrence.

ProSite Identifier	PS00004| CAMP_PHOSPHO_SITE PS00005| PKC_PHOSPHO_SITE PS00006| CK2_PHOSPHO_SITE PS60007| TYR_PHOSPHO_SITE_2	PS00017|ATP_GTP_A	PS00009| AMIDATION	PS00008| MYRISTYL	PS00294| PRENYLATION
Motif	Phosphorylation sites	ATP/GTP-binding site motif A (P-loop)	Amidation site	N-myristoylation site	Prenyl group binding site (CAAX box)
ocGBP1	Yes (22x)	Yes	Yes	Yes (3x)	Yes (CVIS)
ocGBP4L1	Yes (18x)	Yes	Yes	Yes (6x)	No
ocGBP4L2	Yes (14x)	Yes	Yes	Yes (9x)	No
ocGBP4L3	Yes (3x)	Yes	No	Yes (2x)	No
ocGBP4L4	Yes (20x)	Yes	No	Yes (8x)	No
ocGBP4L5	Yes (17x)	Yes	No	Yes (8x)	No
ocGBP4sg	Yes (15x)	Yes	No	Yes (10x)	No
ocGBP5L1	Yes (14x)	Yes	No	Yes (3x)	No
ocGBP5L2	Yes (18x)	Yes	No	Yes (3x)	Yes (CILL)
ocGBP4L6	Yes (19x)	Yes	Yes	Yes (4x)	No

cells and cell lines examined than those of *ocGBP1/4L3/4L5/4sg/5L2/4L6*. In comparison, *ocGBP5L1* only showed higher expression in lung and kidney tissues and in the rabbit skin fibroblast cell line Rab9. On average, *ocGBP1/4L3/4sg/4L5/5L2/4L6* were 76-fold more expressed compared to the low expressors (Figure 4A). These results were largely consistent with the transcriptome data, where *ocGBP1/4L5/4sg/5L2/4L6* transcripts were also present in most of the tissues examined, and a higher number of tissues lacked detectable expression of *ocGBP4L1/4L2/4L4/5L1* (Figure 4B). Notably, *ocGBP4L3* mRNA was only found in the testis in the transcriptome data, whereas the RT-qPCR data showed expression comparable to other GBPs tested in almost all tissues and cell lines analyzed. However, this result of *ocGBP4L3* should be taken with caution due to its short length. In summary, ocGBPs differed in their endogenous mRNA expression levels.

2.6 Cloned ocGBP proteins are expressed in RK13 cells

To functionally characterize ocGBPs, we cloned individual ocGBPs into an expression plasmid with an HA-tag at the N-terminus. Due to the lack of ocGBP-specific antibodies, we analyzed the overexpression of ocGBPs using HA-specific antibodies. Therefore, rabbit RK13 cells were transfected with the individual ocGBPs and protein levels were determined by flow cytometry (Figure 5A) and Western blot (Figure 5B). We observed that all ocGBPs were expressed albeit at different expression levels; ocGBP4L1/4L5/4sg/4L6 were expressed to a higher level than ocGBP4L2/4L3/4L4/5L1/5L2; ocGBP1 showed an intermediate phenotype (Figure 5). In addition, Western blot analysis revealed the expected molecular weight for each ocGBP, ranging from 15-65 kDa (Figure 5B). In summary, all ocGBPs could be overexpressed at the protein level with differential expression between paralogs.

2.7 Varying intracellular localization patterns of ocGBPs

Since GBPs paralogs have been described to perform multiple functions (reviewed in 14) and to differ in their subcellular localization (34, 40, 41), we examined the intracellular localization of overexpressed ocGBPs in RK13 cells using confocal immunofluorescence microscopy. We observed that the rabbit paralogs localized to different intracellular compartments, with distinct patterns (Figure 6). ocGBP1 was distributed throughout the cytoplasm with a continuous and distinct globular localization. ocGBP4L1/4L5/4sg were evenly distributed in the cytoplasm and additionally found in the nucleus. ocGBP4L2 was localized in globular structures in the cytoplasm. ocGBP4L3/4L4 were found in distinct spots in the cytoplasm and nucleus. ocGBP4L6 was distributed in different spots in the cytoplasm and additionally found in the nucleus. We observed that ocGBP5L1 and ocGBP5L2 each co-localized with the TGN, whereas ocGBP5L2 rarely did so - it preferentially localized uniformly or polarized in the cytoplasm. In short, ocGBPs differed in their intracellular localization – some localized either uniformly and/or discretely within vesicle- or aggregate-like structures in the cytoplasm and/or nucleus and/or co-localized with the TGN.

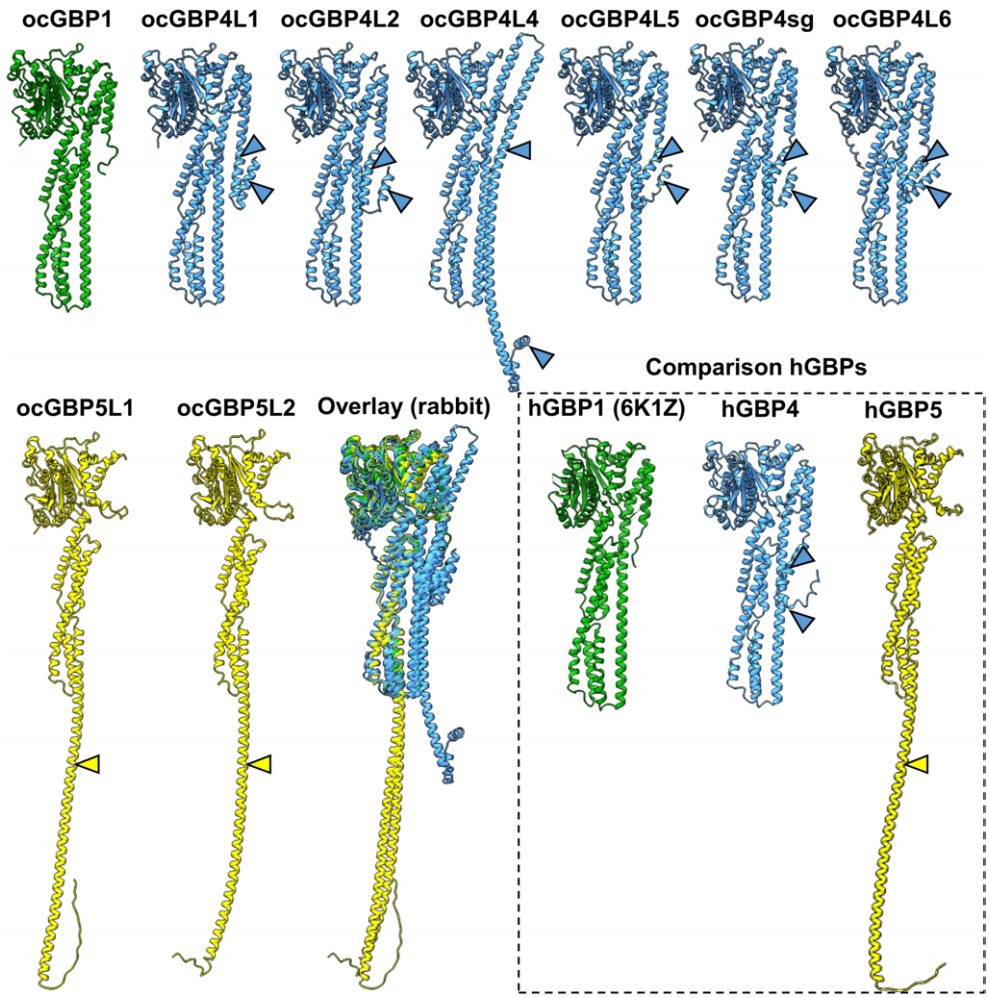

FIGURE 3
Prediction of ocGBP tertiary structures. Best predicted model per GBP is shown. Rabbit GBPs structures were predicted using AlphaFold. For comparison, as comparison structural data of hGBP1 (PDB: 6K1Z) and the predictions for hGBP4 and hGBP5 are also shown. GBPs are colored following the grouping in the phylogenetic analyses (Figure 1).

2.8 Selected ocGBPs are inducible by IFNα and IFNγ

As a next step, we tested whether the expression of *ocGBPs* could be induced by IFN treatment. In the absence of rabbit specific reagents, we used hIFNα2 as a surrogate for ocIFNα since they share 64% aa identity. Using hIFNα2 and ocIFNγ as stimuli, we first screened two different cell lines (RK13, SIRC) and primary cells (data not shown), but IFN-inducibility was observed only in primary rabbit macrophages (Figure 7). We also observed that *ocGBP4L3* and *ocGBP4L5* were not IFN-inducible. For *ocGBP4L1/4L2/4L4/5L1*, IFNs did not induce them above the limit of detection (LoD). These *ocGBPs* also showed low mRNA levels in tissues, primary cells and cell lines compared to the other *ocGBPs* (Figure 4A). In contrast, *ocGBP1/4sg/5L2/4L6* expression was significantly induced upon IFN treatment. Specifically, the mRNA expression of *ocGBP1* was induced 194-fold and 143-fold by hIFNα2 and ocIFNγ, respectively, whereas ocIFNγ-mediated induction of *ocGBP4sg* was only 43-fold. The mRNA expression levels of *ocGBP5L2* were induced only about 6-fold by hIFNα2, and *ocGBPL6* was induced 3-fold by ocIFNγ. In summary, four out of ten *ocGBPs* were IFN-inducible in our experimental setup, suggesting that they might be involved in innate immunity as described for human and muroid GBPs.

2.9 Only ocGBP5L2 inhibits the activity of rabbit furin

Human GBP2 and GBP5 have been shown to interfere with human furin activity (41). The cellular proprotein convertase furin has previously been described to be hijacked by several viruses for the proteolytic processing and activation of their glycoproteins (42). Consequently, hGBP2-/5-mediated furin inhibition prevents the production of fully infectious progeny virions. To determine whether ocGBPs also have the ability to affect the functionality of rabbit furin, we adapted the protocol recently developed by Braun et al. (41) to overexpress synthesized AU-1 tagged ocFurin with

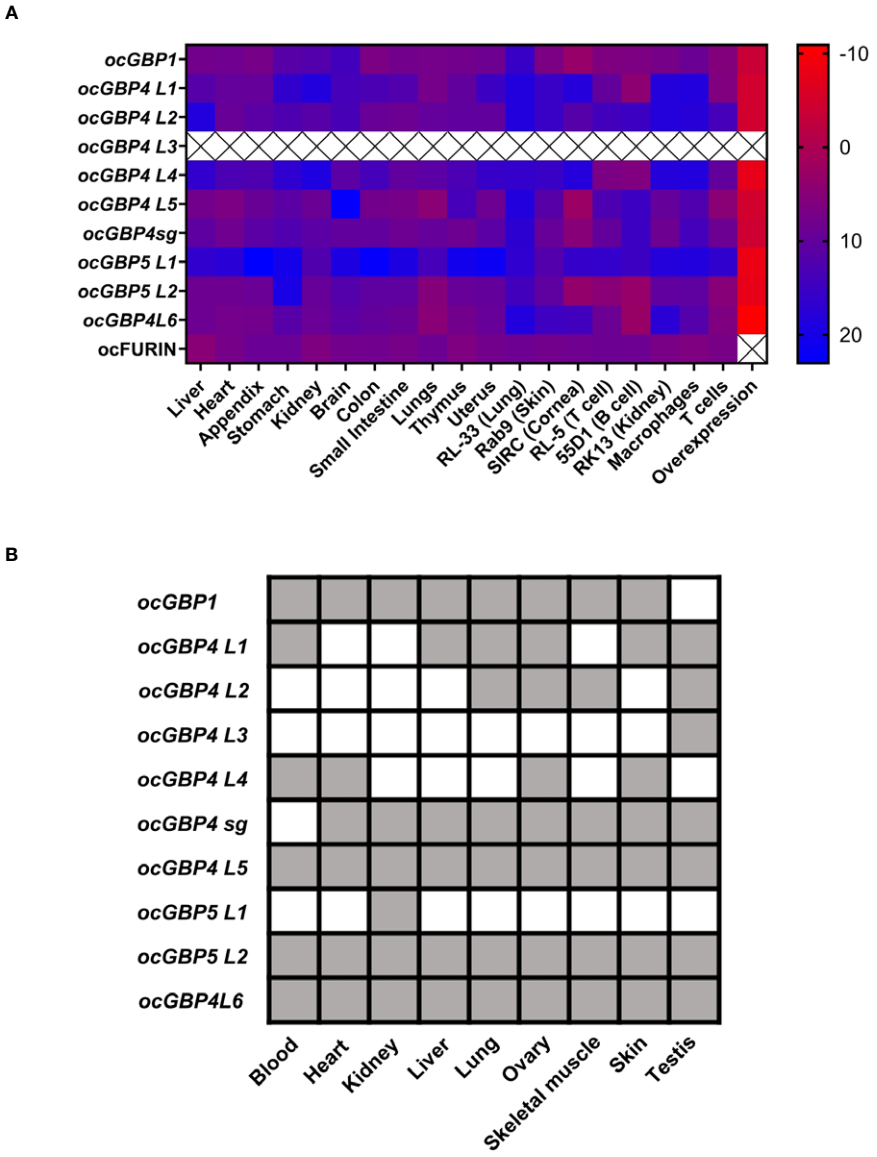

FIGURE 4
Differential mRNA expression levels for rabbit GBPs. **(A)** Heat map of RT-qPCR mRNA expression analysis of ocGBPs and ocFurin in several tissues, primary cells, cell lines and overexpression in RK13 cells: ΔCt values to the reference gene *ActinB* are displayed (CtGBP – CtActB). Tissues of four female New Zealand white rabbits and three primary cells, cell lines and overexpression were analyzed. Scale: from red (low ΔCt value, i.e., higher expression of target gene) to blue (higher ΔCt, i.e., lower expression of target gene). **(B)** Rabbit transcriptome was retrieved from (39) and blasted for *GBP* mRNA expression using the BLAST tool from NCBI. Gray color means present, white means absent.

ocGBPs in HEK293T cells, using human furin together with hGBP5 as a positive control. Interestingly, only ocGBP5L2 inhibited ocFurin activity to a similar extent as hGBP5 for hFurin (Figure 8). Thus, ocGBP5L2 might be able to interfere with glycoprotein processing of various furin-dependent viruses. However, further studies need to investigate whether ocGBP5L2 inhibits viruses via supressing furin activity.

3 Discussion

GBPs are important players in the innate immune response against bacterial, parasitic, and viral infections. However, the breadth of their evolution and mode of action have been mainly addressed in humans and mice (reviewed in 7, 14, 43–45). Here, we expanded the current knowledge of GBP paralogs by analyzing the evolution of lagomorph GBPs and performing functional characterization of European rabbit GBPs.

GBP3 and *7* have been exclusively found in anthropoids and primates (16). Consistent with this, we observed that these genes are absent from lagomorph genomes. Nonetheless, we found that lagomorph *GBPs* underwent a pattern of gain and loss events, similar to those described for other immunity-related genes, including GBPs (46, 47). Despite this similarity, the evolution of the lagomorph *GBP* genes, in particular in leporids, differed from that of other mammals (15–17) with a massive expansion of *GBP4*, especially

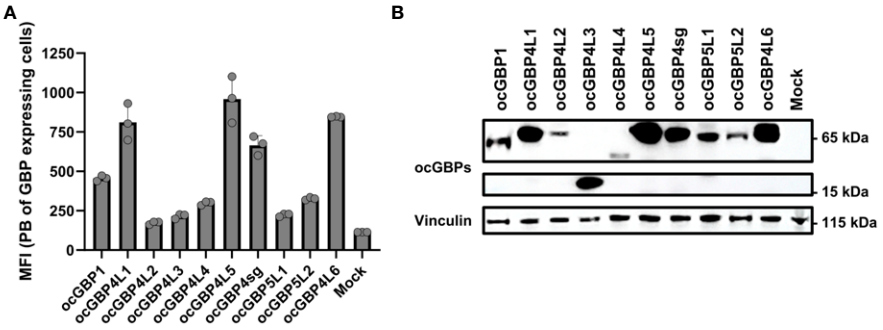

FIGURE 5
Protein expression of overexpressed rabbit GBPs in a rabbit cell line. **(A)** RK13 cells were transfected with ocGBP expression plasmids. Two days post-transfection, cells were permeabilized and protein expression was determined via flow cytometry. Shown are the mean fluorescence intensities (MFI ± SD) of HA-positive cells stained with PB-coupled antibodies (n = 3). **(B)** RK13 cells were transfected with rabbit GBP expression plasmids. Two days post-transfection, protein expression was determined using Western blot. Membranes were probed for HA tag (GBP) and Vinculin (housekeeping protein). Shown is a representative Western blot. PB, pacific blue.

in *Leporidae*. Leporids also present a unique duplication of *GBP5* compared to other mammals and lost *GBP2* (16, 17). However, *GBP2* was still present in *Ochotonidae*, suggesting conservation of *GBP* genes from the common ancestor of rodents and lagomorphs, but we also observed species-specific deletions or expansions of *GBP* genes after speciation. The resulting patterns appeared to be specific to different phylogenetic subgroups and might have been caused by host-pathogen co-evolution and/or host-specific fitness advantages against highly lethal pathogens. In addition, the unusual duplication of *GBP5* in leporids and the expansion of *GBP4* might have compensated for the loss of *ocGBP2* and *6*, respectively (Figures 1, 2). Alternatively, they may have been neofunctionalized or acquired tissue-specific functions. Additionally, it has been described in humans that the recruitment of caspase-4 to the surface of *Salmonella* depends on GBP1 with the auxiliary role of GBP2 and 4 (44, 48), indicating that *Leporidae GBP4* expansion could be a compensation not only for the loss of *GBP6*, but also for the loss of *GBP2*. Comparing the evolutionary history to those of humans and mice (15–17), we could possibly identify ortholog groups, such as *GBP1, GBP4/7*-like and *GBP5*. In addition, by establishing their synteny, we clearly found similar genes flanking the GBP gene cluster as in primates and muroids (15–17). Thus, our data highlight the need for species-specific evolutionary analyses to be able to compare and translate findings from one species to another.

Similar sequences (the Supplementary Data GBP alignment), motifs (Tables 1, 2) and tertiary structures (Figure 3), further backed up by their phylogenetic grouping, might imply similar functions as described for human and murine GBPs (reviewed in 7, 14, 43–45). The highly conserved G domain suggests that GTP binding and hydrolysis is an important feature of GBP proteins in general, which has already been described for other mammals (GTP hydrolysis of human GBPs reviewed in 44; GTPase domains and involvement in function reviewed in 14). Furthermore, the presence of an NLS motif (most of the ocGBPs with predicted NLS also partially localized to the nucleus, see below) and, in the same proteins, the presence of a Glu-rich domain (Table 1) could imply their involvement in gene regulation, although these motifs were predicted with low confidence. Several high-probability sequence motifs (Table 2) and putative post-translational modifications may imply tightly regulated protein expression, function, and localization, which has been described for other GBP paralogs (4, 34, 40, 49–53).

For some of the expanded genes, specifically *ocGBP4L1/4L2/4L4/5L2*, we observed that they were consistently expressed to a lower level in most tissues, primary cells and cell lines (Figure 4). High expression of *ocGBP4L5*/4*sg* might induce a "dosage effect" of *ocGBP4L1/4L2/4L4/5L2*, but the diversity of many ocGBP4s could still have an evolutionary advantage. They may also be tissue-specific factors that are expressed and required only in certain tissues at certain timepoints (Figure 4).

We observed that all overexpressed ocGBPs differed in their expression levels and yielded the expected molecular weight (Figure 5). Similar to the mRNA expression, there was a distinct pattern of ocGBPs with higher and lower expression levels. We saw a correlation between lower expression and localization (see below), but not with IFN inducibility.

Varying localization of GBPs has been described in the context of human GBPs (34, 40, 41). Rabbit GBPs localized either uniformly and/or discretely within vesicular or aggregate-like structures in the cytoplasm and/or nucleus or co-localized with the TGN (Figure 6). We further observed that the phylogenetically coherent ocGBP4L1/4L2/4L3/4L4/4L4/4sg/4L6 and ocGBP5L1/5L2 clusters localized according to their protein expression levels (Figure 5), with the ocGBPs with lower protein expression forming aggregates (ocGBP4L2/4L3/4L4). We speculate that such aggregation might be harmful for homeostasis and, therefore, locally restricted. ocGBP5L1 and rarely ocGBP5L2 co-localized with the TGN (Figure 6) as described for hGBP5, for which the localization was suggested to be required for its antiviral activity (41). This is not expected since ocGBP5L1, unlike ocGBP5L2, does not have a CaaX motif (Table 2). Of note, for ocGBP5L2, we rarely observed this co-localization, as we more often observed a uniform localization to the cytoplasm or a polarized localization in the cytoplasm (Figure 6). In contrast to human GBP5 (41), ocGBP5L1 co-localized with the Golgi, which suggests that the prenylation is not the only determinant and other described modifications, such as N-myristoylation present in both ocGBP5s, could also play a role (Table 2). Since ocGBP5L2 only rarely localized to the TGN, we

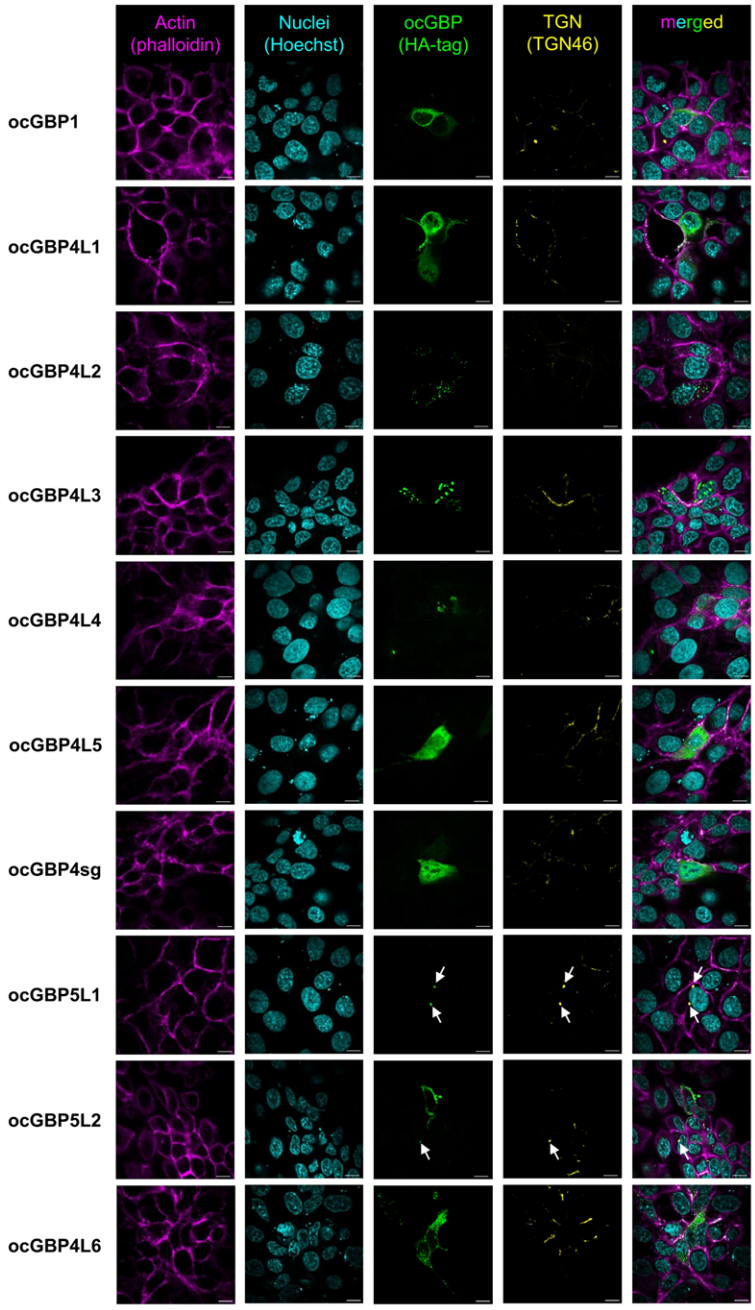

FIGURE 6
Intracellular localization of ocGBPs in RK13. RK13 cells were transfected with GBP expression plasmids. Two days post-transfection, localization was determined via immunofluorescence microscopy. The following colors were used: pink (phalloidin, actin filaments), yellow (TGN46, trans-Golgi network), indigo (Hoechst, Nucleus), green (HA-tag, GBPs). Shown are representative images out of 5 -10 imaged positions. 100x magnification, scale bars indicate 10 µm.

speculate that ocGBP5L1 and ocGBP5L2 may form heterodimers as described for other GBPs (34, 40) and thus increase the affinity to bind to the TGN for antiviral activity.

We found that the mRNA levels of ocGBP1, ocGBP4sg, ocGBP5L2 and ocGBP4L6 were significantly induced by IFN treatment in primary rabbit macrophages (Figure 7). In addition, ocGBP5L2 inhibited the activity of ocFurin (Figure 8). This would suggest that despite their genetic diversity compared to muroid and human GBPs, they play a similar role in immune responses as those described for mouse and human GBPs. The cause of the differentially induced expression of GBPs by IFNs could be their involvement in different functions in the innate immune response, as observed for human GBPs with specific paralogs involved in the response to different (classes of) pathogens or in inflammatory and cancer pathways (reviewed in 7, 14, 43–45). For hGBPs, one explanatory approach is the difference in 5' regulatory elements for IFN-dependent transactivators between the different paralogs obtained from CHIP-seq ENCODE data (43).

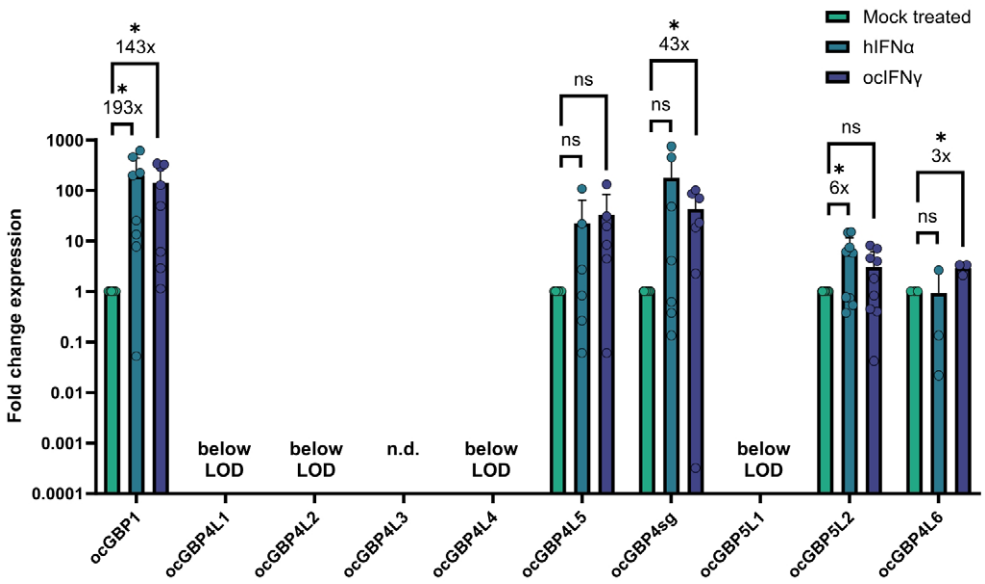

FIGURE 7
ocGBP mRNA levels in primary rabbit macrophages (Mφ) after IFN stimulation. Mφ were stimulated with human IFNα2, rabbit IFNγ, or mock-treated. mRNA levels were measured via RT-qPCR. Shown are relative expression levels that were normalized to the mock-treated samples ($2^{-\Delta\Delta Ct}$ method). Fold change in mRNA levels compared to mock-treated are displayed (mean ± SD, 3 donors with n = 3 each). Asterisks indicate significance * $p \leq 0.05$. n.s, not significant; n.d. not determined.

We observed that the different rabbit paralog groups (ocGBP1 induced by both IFN, ocGBP4 by IFNγ and ocGBP5 by IFNα) have distinct group-specific structural features (Tables 1, 2 and Figure 3). This could imply a similar but distinct function for the different paralog groups. This is contradicted by the different IFN induction within the groups (Figure 7), but could be explained by the loss/gain of an IFN-dependent 5' regulatory element in the gene duplication process, so that these genes may have acquired new or additional

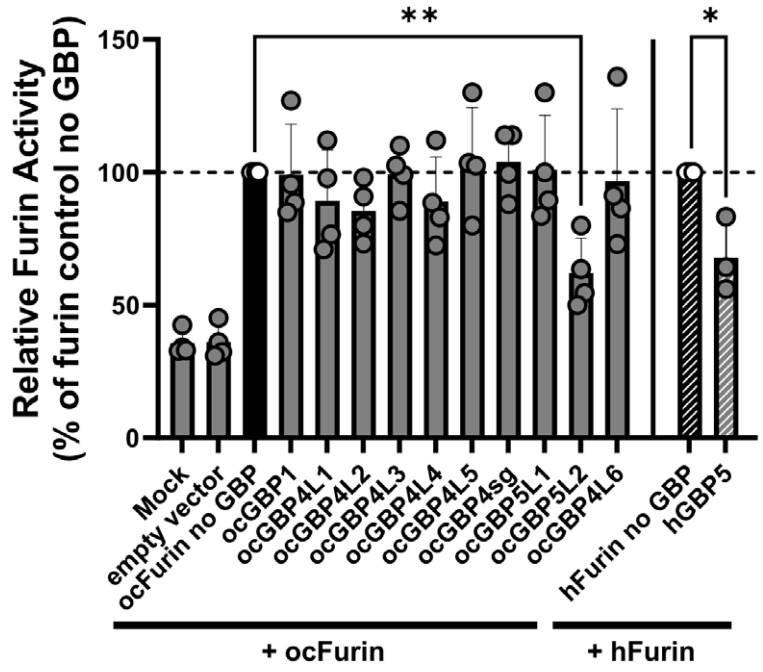

FIGURE 8
ocGBP5L2 interferes with ocFurin activity. HEK293T cells were co-transfected with furin-expression plasmids and GBP-expression plasmids. Two days post-transfection, the activity of furin secreted into the supernatant was measured after adding the AMC substrate. Shown are mean values of three independent experiments (human GBP5) or four independent experiments (rabbit GBPs). Error bars indicate SD. Asterisks indicate significance * $p \leq 0.05$.** $p < 0.01$.

functions, or may still be functional at constitutionally lower levels of expression. Despite the structural similarity to ocGBP5L1, only ocGBP5L2 inhibited furin activity (Figure 8). Therefore, the CaaX motif may be essential for furin inhibition.

In conclusion, our work adds valuable information to the evolution of ocGBPs and their characteristics, and implicates implicates a role of ocGBPs in innate immunity, which needs to be evaluated in future studies.

4 Materials and methods

4.1 Synteny

Syntenic positions and transcription orientations of lagomorph *GBP* were inferred by visual inspection of their genomes in publicly available databases NCBI (https://www.ncbi.nlm.nih.gov/genome/gdv/) and Ensembl (https://www.ensembl.org/index.html).

4.2 Phylogeny

Gene sequences annotated as *GBP* were retrieved in the timeframe March to August 2021 from publicly available databases. A total of 202 sequences were retrieved (see Supplementary Table 1): 19 sequences were retrieved from three different lagomorph species (*Ochotona princeps*, *Ochotona curzoniae* and *Oryctolagus cuniculus*), 41 sequences from 6 species of primate origin, including *Homo sapiens*; 123 sequences from 12 rodent species; *Tupaia glis* (5 sequences), *Loxodonta africana* (7 sequences) and *Gallus gallus* (7 sequences), the latter of which was used as outgroup. To ensure that all *GBP* sequences from all the species were included, a subsequent BLAST analysis was performed. Sequences that did not encode a functional protein or presented partial mRNA sequences were excluded from the analysis (accession numbers available in Supplementary Table 1). Alignment of the sequences was performed in BioEdit software (54) using the Clustal ω method (55) followed by visual inspection and correction. Alignment of GBP protein sequences can be found in the Supplementary Data GBP alignment. In addition, the alignment was screened for gene conversion/recombination using GARD (27). Phylogenetic relationships were inferred in MEGAX (56) using the Maximum Likelihood (ML) method and the Jones-Taylor-Thornton matrix-based substitution model + G + I as determined by MEGAX (57). To assess the robustness of the tree branches, 1000 bootstrap replicates were used. The trees were drawn to scale, with branch lengths measured in the number of substitutions per site. All positions with less than 95% site coverage were eliminated, i.e., fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position (partial deletion option).

4.3 ProSite Scan

The ProSite Scan tool was used to identify (functional) protein sequence motifs (31–33). Protein sequences of the ocGBPs were included to scan them against the PROSITE collection of motifs. The scan was performed at high sensitivity.

4.4 Protein structure modeling with AlphaFold

For structure prediction, ChimeraX (https://www.rbvi.ucsf.edu/chimerax) (58) was used with the structure prediction AlphaFold tool with the corresponding ocGBP protein sequences. Computations were performed on Google Colab using ColabFold, an open source, optimized version of AlphaFold 2 (59). The resulting prediction models were visualized using ChimeraX (58).

4.5 Rabbit organ and serum preparation

Four 36-to 40-days old female New Zealand white rabbits (*Oryctolagus cuniculus*) were ordered from Charles River (France) and housed for an additional acclimation week prior to organ removal in a specific pathogen-free (SPF)-barrier. The laboratory conditions and husbandry of the animals were identical to a recently published study (60). They were euthanized by slow intravenous injection of a lethal dose of sodium-pentobarbital (100 mg/kg Narcoren, Boehringer Ingelheim, Ingelheim am Rhein, Germany). The following organs were collected: spleen, liver, heart, appendix, stomach, kidney, brain, colon, small intestine, lungs, thymus and uterus. The organs were frozen in liquid nitrogen and homogenized to frozen powder, which was stored at -80°C prior further processing. For the preparation of rabbit serum, rabbit blood was collected from the heart and incubated at 37°C for 30 min, followed by 30 min on ice. The blood was then centrifuged at 12 000 g for 10 min. The experiments have been approved by the institutional ethical review committee (LMU Munich, Biomedical Center, Core facility animal models) and are in accordance with the local government authorities Az.5.1-5682 (LMU/BMC/CAM) as well as European (RL2010/63EU) and German animal welfare legislation.

4.6 Preparation of splenocytes and macrophage and T cell differentiation

Rabbit splenocytes were prepared by mashing the spleens through a 40 µm cell strainer (LABSOLUTE) in 1x PBS until only rigid scaffolds (capsules) were left. The cells were subsequently pelleted for 5 min at 500 g and the remaining red blood cells were lysed with 4 ml ACK lysis buffer (8.29 g/l NH_4CL (Carl Roth), 1 g/l $KHCO_3$ (Carl Roth), 0.0367 g EDTA (CHEMSOLUTE)) for 5 min at room temperature (RT) and washed in 1x PBS for 5 min at 500 g. The procedure was repeated until lysis was complete. Splenocytes were cultivated in RPMI 1640 GlutaMAX™ (Gibco) supplemented with 10% (v/v) FCS (Sigma-Aldrich) and 1% (v/v) Penicillin-Streptomycin (10,000 units Penicillin and 10 mg Streptomycin per ml, Sigma-Aldrich) at standard conditions (37°C; 5% CO_2; 90% humidity). For T cell differentiation, splenocytes were maintained at 2 x 10^6 cells/ml with 100 U/ml human

recombinant IL-2 (Biomol #50442) and 5.0 µg/ml Concanavalin A (Sigma-Aldrich #C2010) for four days and then only cultivated in IL-2 containing medium. For the differentiation of rabbit macrophages, 2 x 10^6 cells/ml rabbit splenocytes were seeded into 12-well plates with 2% (*v/v*) rabbit serum for one week until heterogeneous differentiation could be observed.

4.7 Cell culture cell lines

SIRC (Cornea, ATCC CCL-60), RAB-9 (Skin, ATCC CRL-1414), RK13 (Kidney, ATCC CCL-37) and RL-33 (Lung, tebu-bio JCRB0131) cell lines were cultured in monolayers in MEM GlutaMAX™ (Gibco) supplemented with 10% heat-inactivated fetal calf serum (FCS, Sigma Aldrich) and 1% Penicillin/Streptomycin (P/S, Sigma Aldrich). 55D1 (B-cell line, (61); kind gift of Dr. Katherine L. Knight) and RL-5 (T-cell line; 62) were cultured in RPMI 1640 GlutaMAX™ Medium (Gibco) supplemented with 10% FCS (Sigma Aldrich) and 1% P/S (Sigma Aldrich). HEK293T cells (Kidney, human, DSMZ ACC 635) were cultured in monolayers in DMEM GlutaMAX™ (Gibco) supplemented with 10% FCS (Sigma Aldrich) and 1% P/S (Sigma Aldrich). All cells were cultured at 37°C and 5% CO_2 and 90% humidity.

4.8 RT-qPCR

RNA was extracted from the samples using NucleoZol (Macherey-Nagel); the remaining genomic DNA was digested using TURBO DNA-free™ Kit (Invitrogen) and cDNA was subsequently generated using the High-Capacity RNA-to-cDNA™ Kit (Applied Biosystems). All three steps were conducted according to the manufacturer's instructions. One ng/µl cDNA was prepared for cell lines and primary cells and 10 ng/µl for tissues samples. Analysis of gene expression was performed using PowerUp™ SYBR™ Green Master Mix (Applied Biosystems) with a Quantstudio 3 Real-Time PCR system (Applied Biosystems). The reactions were set up using 5 µl PowerUp™ SYBR™ Green Master Mix, 2 µl nuclease-free water, 0.5 µl of 10 µM forward and reverse primer each (Table 3) and 2 µl of respective cDNA to a total reaction volume of 10 µl. The following thermal cycling conditions were used: hold stages at 50°C for 2 min and at 95°C for 2 min, 40 cycles with denaturation at 95°C for 1 s and annealing/elongation at 60°C for 30 s. Finally, the melting curve was performed with 95°C for 1 s and 60°C for 20 s with a rate of 0.1°C/s from 60°C to 95°C. Ct values were used to determine gene expression in relation to the reference gene. Optimal qPCR primers were designed using primer3 (https://primer3.ut.ee/) (80-120 bp amplicon length, 20 bp optimal length and 60°C optimal T_m, Table 1) (63, 64). One primer of each primer pair was spanning an exon-exon junction. Results were analyzed as ΔCt = Ct (ocGBP) – Ct (Actin β (ActB)).

4.9 BLAST analysis for rabbit transcriptome

Rabbit transcriptome was generated as part of the rabbit genome paper (39) and the deposited data were analysed using the BLAST tool from NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

4.10 Cloning of ocGBPs in expression plasmid

Template cDNAs for amplification of the rabbit *GBPs* were prepared as described above. pCG vector was used as backbone template. Rabbit *GBPs* were amplified using the primer pairs in Table 4 by PCR (Tables 5, 6) with the cDNAs as template. An HA-tag for detection purposes was added at the N-terminus. Since ocGBP5L1 could not be amplified from rabbit cDNAs, it was ordered from Twist Bioscience (accession number: XM_002715873.3). The final pCG-HA-GBP plasmids were obtained via Gibson assembly using NEBuilder® HiFi DNA Assembly Master Mix (NEB) according to the manufacturer's protocol and sequence-verified using Sanger sequencing.

4.11 Transfection for protein overexpression

For heterologous expression of the ocGBPs, 1.2 x 10^5 RK13 cells, seeded one day prior to transfection in a 12-well plate, were transfected with 1.5 µg of each pCG-HA-ocGBP plasmid, respectively, using TurboFect transfection reagent (Thermo Fisher) in 100 µl of unsupplemented medium, according to manufacturer's instructions. The mixture was incubated at RT for 45 min and then added dropwise to the cells. Cells were incubated at 37°C, 5% CO_2 and 90% humidity and expression was analyzed two days post-transfection.

4.12 Flow cytometric analyses

For flow cytometry detection of ocGBP expression, transfected cells were intracellularly stained for the HA-tag. Briefly, detached RK13 cells were fixed with 100 µL pre-warmed PFA (4% in 1x PBS, AppliChem) at RT for 10 min. The cells were washed once with 1 x PBS (Sigma Aldrich), and the supernatant was aspirated. For permeabilization, 100 µl of pre-cooled BD Phosflow Perm Buffer III was added to each well and incubated for 2 min on ice. Cells were washed twice with 1 x PBS, then resuspended and stained with 50 µl of Pacific Blue™ anti-HA.11 epitope tag antibody (1:100; #901526, Biolegend) in staining buffer (1x PBS (pH 7.2), 1% (*v/v*) FCS (Sigma-Aldrich), 0.09% NaN_3(Carl Roth)) at RT for 45 min in the dark. Cells were washed once with staining buffer and then resuspended in 200 µl of staining buffer for subsequent analysis using the BD FACSLyric™ Flow Cytometer. Data were analyzed using the FlowJo software.

TABLE 3 RT-qPCR primers for lagomorph gene expression analyses.

	Accession Number	Primer	Sequence 5′-3′
ocGBP1	XM_002715873.3	q_ocGBP1_f	AGCAAGGGGTCTTTTCTAAACC
		q_ocGBP1_r	TCTTCAGCCTGTATCCCTTTCC
ocGBP4 L1	XM_008264927.2	q_ocGBP4_L1_f	CGAAAGAAACTTACCGACACCAT
		q_ocGBP4_L1_r	CGAAAGCCGCCTAAGTTCAG
ocGBP4 L2	XM_017346115.1	q_2_ocGBP4_L2_f	CCTGTAGTAGTAGTGGCCATTGT
		q_2_ocGBP4_L2_r	CAGAGGGAAGCCATGTTTCTG
ocGBP4 L3	XM_017345575.1	q_3_ocGBP4_L3_f	TCTTAACCAGATATCTCAGCCTGT
		q_3_ocGBP4_L3_r	GGGAAGCCATGTTTCTGTCCT
ocGBP4 L4	XM_017346109.1	q_2_ocGBP4_L4_f	GCACAAGCTGAAGGCTCAAA
		q_2_ocGBP4_L4_r	TCTCTTCTGTTAGCCGCTTGA
ocGBP4 L5	XM_002715512.3	q_2_ocGBP4_L5_f	AGAAGATGGAGCGGGAAAGG
		q_2_ocGBP4_L5_r	AGCATTTCTTCTTGGACCTTCAG
ocGBP4sg	XM_008264924.2	q_2_ocGBP4_sg_f	AGCACAAGCTGAAGGTTCAAA
		q_2_ocGBP4_sg_r	GCTGCCATATCTTCTGTTATCCG
ocGBP5 L1	XM_008265608.2	q_2_ocGBP5_L1_f	AGAGGTGTGGCAAATGGAGA
		q_2_ocGBP5_L1_r	ATTGCAGCCTCCTCCTGG
ocGBP5 L2	XM_002715513.3	q_3_ocGBP5_L2_f	AGAGGTGCGACAAATGGAGA
		q_3_ocGBP5_L2_r	CTCTGAGCCTCTTCCTGGAG
ocGBP4L6	XM_008264918.2	q_2_ocGBP4L6_f	CCAGGAGAACATCACCCAGT
		q_2_ocGBP4L6_r	AGCAGGTCTTCTTGGATCTTCA
Actin beta (ActB)	NM_001101683.1	ActB_L_f	TCCTGGGCATGGAGTCGT
reference gene		ActB_L_r	GTGTTGGCGTACAGGTCCT

TABLE 4 PCR primers to clone rabbit GBPs.

Primer	Sequence 5'-3'
pCG_amp_f	ACGCGTCGGATCCTGAGAAC
pCG_amp_HA_r	AGCGTAATCTGGAACATCGTATGGGTACATTCTAGAAGGCCTACGCGCTTC
gib_3.1_pCG_rbGBP1_f	GTACCCATACGATGTTCCAGATTACGCTATGACCTCAGAGATCCACATG
gib_3.1_pCG_rbGBP1_r	CTGAAGTTCTCAGGATCCGACGCGTTTAGCTTATAACACATCTTCTCCTTGG
gib_3.4L1_pCG_rbGBP4_L1_f	GTACCCATACGATGTTCCAGATTACGCTATGGCAACCGAATTTATGAATG
gib_3.4L1_pCG_rbGBP4_L1_r	CTGAAGTTCTCAGGATCCGACGCGTCTATTTAATTTGTGAACTGATAAATCGC
gib_3.4L2_pCG_rbGBP4_L2_f	GTACCCATACGATGTTCCAGATTACGCTATGGCAACTGAATTCACCATG
gib_3.4L2_pCG_rbGBP4_L2_r	CTGAAGTTCTCAGGATCCGACGCGTCTATGCAGTTGTTAAAGTCTGGT
gib_3.4L3_pCG_rbGBP4_L3_f	GTACCCATACGATGTTCCAGATTACGCTATGGCAACTAATATCACCATGAAG
gib_3.4L3_pCG_rbGBP4_L3_r	CTGAAGTTCTCAGGATCCGACGCGTTTAAACTGTAAGAGCACAGTTGAG
gib_3.4L4_pCG_rbGBP4_L4_f	GTACCCATACGATGTTCCAGATTACGCTATGGCGACTGATATCACC
gib_3.4L4_pCG_rbGBP4_L4_r	CTGAAGTTCTCAGGATCCGACGCGTCTATAACTTTCTTAACAGCCTTGA

(Continued)

TABLE 4 Continued

Primer	Sequence 5'-3'
gib_3.4L5_pCG_rbGBP4_L5_f	GTACCCATACGATGTTCCAGATTACGCTATGGCAACTGATATCACCATG
gib_3.4L5_pCG_rbGBP4_L5_r	CTGAAGTTCTCAGGATCCGACGCGTTCAGTCTTTAGATTTTGAACCAAG
gib_3.4sg_pCG_rbGBP4sg_f	GTACCCATACGATGTTCCAGATTACGCTATGGCAACTGATACTACCATG
gib_3.4sg_pCG_rbGBP4sg_r	CTGAAGTTCTCAGGATCCGACGCGTCTATAAAATTCTTCGACTCAGTCTTAAC
gib_3.5L1_pCG_rbGBP5_L1_f	GTACCCATACGATGTTCCAGATTACGCTATGGCCTCGGAGATCCTC
gib_3.5L1_pCG_rbGBP5_L1_r	CTGAAGTTCTCAGGATCCGACGCGTTTATCTCTTTGGTGAAAAGAAAGTTCCA
gib_3.5L2_pCG_rbGBP_ L2_f	GTACCCATACGATGTTCCAGATTACGCTATGGCCTTGGAGATCCTC
gib_3.5L2_pCG_rbGBP_ L2_r	CTGAAGTTCTCAGGATCCGACGCGTTTAGAGTAAGATGCAATCATCATTTGG
gib_3.7_pCG_rbGBP4L6_f	GTACCCATACGATGTTCCAGATTACGCTATGGACACCACAAATCCTG
gib_3.7_pCG_rbGBP4L6_r	CTGAAGTTCTCAGGATCCGACGCGTCTATTTTATTTGTGTGCTCAACATTTTC

4.13 Western blot

Cells were lysed in 50 µl Hunt lysis buffer (20 mM Tris-HCl pH 8.0 (Carl Roth), 100 mM sodium chloride (Carl Roth), 1 mM EDTA (CHEMSOLUTE), 0.5% NP-40 (AppliChem) containing 1 x cOmplete protease inhibitor cocktail (Roche)). Lysates were cleared by centrifugation at 20,000 *g* and 4°C for 20 min. Protein concentration was quantified with the Quick Start Bradford 1 x and Quick Start Bovine Serum Albumin Standard Set (both Bio-Rad) according to the manufacturer's instructions using CLARIOstar (BMG Labtech). 4 x Laemmli buffer (Bio-Rad) supplemented with 50 mM DTT (Carl Roth) was added to a final concentration of 1 x to the samples, which were then denatured at 95°C for 5 min. NuPage 4-12% Bis-Tris gels were used (Invitrogen). 15 µg total protein was loaded per sample. Gel electrophoresis was performed in 1 x MOPS-SDS running buffer using a Mini Gel Tank (both Invitrogen) at 100 V for 90 min. Afterwards, the proteins were transferred onto nitrocellulose membranes (0.45 µm, Bio-Rad) in 1 x Tris-Glycine Transfer Buffer (25 mM Trizma base, 192 mM Glycine, 20% (*v/v*) methanol (CHEMSOLUTE)) using the Mini Blot Module (Invitrogen) at constant voltage (14 V, 75 min). Membranes were blocked in TBS-T (20 mM Tris, 150 mM NaCl, 0.1% (*v/v*) Tween 20 (Carl Roth) containing 5% (*w/v*) powdered milk (Carl Roth). Proteins were stained with the following primary antibodies overnight: HA tag polyclonal antibody (SG77, #71-5500, 1:250, Invitrogen) and Vinculin recombinant rabbit monoclonal antibody (42H89L44, #700062, 1:1000, Invitrogen). The following day, membranes were washed three times with 1 x TBS-T, incubated for 1 h at RT in horseradish peroxidase (HRP)-coupled secondary antibody peroxidase AffiniPure goat anti-rabbit IgG (H+L) (#AB_2313567, Jackson Immunoresearch), diluted 1:10000 in 1 x TBS-T containing 5% (*w/v*) powdered milk. After three washing steps in 1 x TBS-T, blots were visualized with the SuperSignal™ West Femto Maximum Sensitivity Substrate (Thermo Scientific) according to the manufacturer´s instruction using the FUSION FX (Vilber).

TABLE 5 PCR reaction components for GBP cloning.

COMPONENT	VOLUME (µl)	FINAL CONCENTRATION
5X Phusion™ HF Buffer (Thermo Scientific)	10 µl	1X
10 mM dNTPs (Thermo Scientific)	1 µl	200 µM
10 µM Forward Primer (Eurofins)	2.5 µl	0.5 µM
10 µM Reverse Primer (Eurofins)	2.5 µl	0.5 µM
Template DNA	1 pg–10 ng (plasmid or viral); 50 ng– 250 ng (genomic)	< 1,000 ng
Phusion™ High–Fidelity DNA Polymerase (Thermo Scientific)	0.5 µl	0.02 U/µl
Nuclease-Free Water	add to 50 µl	

4.14 Immunofluorescence

RK13 cell were seeded on 13 mm glass cover slips (VWR) and transfected as above. Transfected cells were fixed with 100 µl pre-

TABLE 6 PCR thermocycler program to clone rabbit GBPs.

Temperature [°C]	Time [s]	cycles
98	60	
98	10	
Ta*	30	35 x cycles
72	75	
72	360	
10	hold	

*optimal annealing temperature of the respective primer pairs.

warmed PFA (4% in 1x PBS, AppliChem) at RT for 15 min, permeabilized with 100 µl 0.1% Triton X-100 (Carl Roth) in 1x PBS for 5 min at RT and blocked with 1 x PBS with 2% (*w/v*) BSA (Carl Roth) for 30 min at RT. First, the actin filaments were stained using phalloidin Atto-647N (10 µM in MeOH, #AD647N-81, 1:60, ATTO-TEC) in 1x PBS for 30 min at RT. For TGN staining, primary antibody sheep anti human TGN46 (#AHP500G, 1:1000, Bio-Rad) was first incubated in 1 x PBS with 2% (*w/v*) BSA for 60 min at RT. Subsequently, cells were stained with secondary antibody donkey anti-sheep IgG (H+L) cross-adsorbed secondary antibody Alexa Fluor 568 (#A-21099, 1:2000, Invitrogen) and directly-coupled anti-HA.11 epitope tag antibody Alexa Fluor 488 (#901514, 1:1000, Biolegend) in 1 x PBS with 2% (*w/v*) BSA for 1 h at RT. Lastly, nuclei were counterstained with 1 µg/ml Hoechst 33342 solution (#62249, Thermo Scientific) for 15 min at RT. Cells were then washed using Millipore water to get rid of salts. Cover slips were mounted with ProLong Gold Antifade Mountant (Invitrogen) on microscope slides (Carl Roth) and dried for 24 h at RT before microscopy analyses were performed using the Yokogawa Spinning Disk Field Scanning Confocal System CSU-W1 (Nikon) with 100 x magnification and following filters: Filter block 1: EX 387/11 EM 416 LP, Filter block 2 EX 469/35 EM BA 525/39, Filter block 3: EX 559/34 EM 639/69, Filter block 4: EX 628/40 EM 692/40.

4.15 Interferon stimulation

RK13, SIRC and M0 macrophages (prepared as described above) were stimulated with either with 20 ng/ml human IFNα2 (#592702, Biolegend), rabbit IFNγ (#RP0136U-005, Kingfisher Biotech) or mock-treated, and harvested 24 h post stimulation. Cells were prepared as described above for RT-qPCR, which was conducted using the same primers and protocol. Results were analyzed using $2^{-\Delta\Delta Ct}$ method.

4.16 Furin activity measurement

ocFurin was synthesized by basegene (Leiden, Netherlands) and subcloned into the pCG C-AU-1IRES BFP vector (41). To determine furin activity in HEK293T cells, the assay was essentially performed as previously described (41) by co-transfecting cells in a 96-well cell culture plate with 50 ng furin-expressing plasmids and 75 ng GBP-expressing plasmids. Two days post-transfection, 20 µL of cell culture supernatant was incubated with the Pyr-Arg-Thr-Lys-Arg-7-Amido-4-methylcoumarin (AMC) substrate (1 nmol), and furin activity was determined for 5 h with an interval of 2 min using a Cytation3 imaging reader (355 nm excitation and 460 nm emission).

4.17 Statistics

Statistical analyses were performed using GraphPad Prism 9 using Students' t-test.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding authors.

Ethics statement

The animal study was approved by the local government authorities Az.5.1-5682 (LMU/BMC/CAM) as well as European (RL2010/63EU) and German animal welfare legislation. The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

LS: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. JVCR: Conceptualization, Data curation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. SF: Investigation, Formal analysis, Writing – review & editing. AB: Investigation, Writing – review & editing. MS: Investigation, Formal analysis, Writing – review & editing. MP: Investigation, Formal analysis, Writing – review & editing. RL: Formal analysis, Supervision, Writing – review & editing. BM: Investigation, Writing - review & editing. DM: Resources, Writing – review & editing. LP: Resources, Writing – review & editing. AP: Data curation, Writing – review & editing. JM: Data curation, Resources, Writing – review & editing. JM-F: Data curation, Resources, Writing – review & editing. BP: Resources, Supervision, Writing – review & editing. PE: Funding acquisition, Supervision, Writing – review & editing. DS: Funding acquisition, Supervision, Writing – review & editing. JA: Funding acquisition, Supervision, Writing – review & editing. H-MB: Conceptualization, Funding acquisition, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by Fundação para a Ciência e Tecnologia (FCT) through the PhD fellowship of JVCR (DFA/BD/4965/2020), the Assistant Researcher grant of JA (CEECIND/00078/2017), the principal Researcher grant of JM-F (2021.00150.CEECIND) and the Principal Researcher grant of PE (CEECIND/01495/2020). MP was funded by the Interdisciplinary Doctoral Program in Medicine of the University Hospital Tübingen and the German Center for Infection Research (DZIF, TI 07.003_Petersen). DS was supported by the SPP1923 and the Heisenberg Program of the DFG (SA 2676/3-1; SA 2676/1-2). H-MB acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG) (BA-6820/1-1).

Acknowledgments

We thank Dr. Oliver Keppler for continuous support. We would also like to thank Corinna Bay and Isabell Haußmann for excellent technical support. In addition, we want to thank Drs. Christine Josenhans and Barbara Adler for fruitful discussions.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1303089/full#supplementary-material

References

1. Meunier E, Broz P. Interferon-inducible GTPases in cell autonomous and innate immunity. *Cell Microbiol* (2016) 18(2):168–80. doi: 10.1111/cmi.12546

2. Santos JC, Broz P. Sensing of invading pathogens by GBPs: At the crossroads between cell-autonomous and innate immunity. *J Leukoc Biol* (2018) 104(4):729–35. doi: 10.1002/jlb.4mr0118-038r

3. Man SM, Place DE, Kuriakose T, Kanneganti TD. Interferon-inducible guanylate-binding proteins at the interface of cell-autonomous immunity and inflammasome activation. *J Leukoc Biol* (2017) 101(1):143–50. doi: 10.1189/jlb.4MR0516-223R

4. Kohler KM, Kutsch M, Piro AS, Wallace GD, Coers J, Barber MF. A rapidly evolving polybasic motif modulates bacterial detection by guanylate binding proteins. *mBio* (2020) 11(3):e00340–00320. doi: 10.1128/mBio.00340-20

5. Ngo CC, Man SM. Mechanisms and functions of guanylate-binding proteins and related interferon-inducible GTPases: Roles in intracellular lysis of pathogens. *Cell Microbiol* (2017) 19(12):e12791. doi: 10.1111/cmi.12791

6. Kim B-H, Chee JD, Bradfield CJ, Park E-S, Kumar P, MacMicking JD. Interferon-induced guanylate-binding proteins in inflammasome activation and host defense. *Nat Immunol* (2016) 17(5):481–9. doi: 10.1038/ni.3440

7. Praefcke GJ. Regulation of innate immune functions by guanylate-binding proteins. *Int J Med Microbiol* (2018) 308(1):237–45.

8. Huang S, Meng Q, Maminska A, MacMicking JD. Cell-autonomous immunity by IFN-induced GBPs in animals and plants. *Curr Opin Immunol* (2019) 60:71–80. doi: 10.1016/j.coi.2019.04.017

9. Kresse A, Konermann C, Degrandi D, Beuter-Gunia C, Wuerthner J, Pfeffer K, et al. Analyses of murine GBP homology clusters based on in silico, in *vitro* and in *vivo* studies. *BMC Genomics* (2008) 9(1):158. doi: 10.1186/1471-2164-9-158

10. Vestal DJ, Jeyaratnam JA. The guanylate-binding proteins: emerging insights into the biochemical properties and functions of this family of large interferon-induced guanosine triphosphatase. *J Interferon Cytokine Res* (2011) 31(1):89–97. doi: 10.1089/jir.2010.0102

11. Kravets E, Degrandi D, Weidtkamp-Peters S, Ries B, Konermann C, Felekyan S, et al. The GTPase activity of murine guanylate-binding protein 2 (mGBP2) controls the intracellular localization and recruitment to the parasitophorous vacuole of Toxoplasma gondii. *J Biol Chem* (2012) 287(33):27452–66. doi: 10.1074/jbc.M112.379636

12. Legewie L, Loschwitz J, Steffens N, Prescher M, Wang X, Smits SHJ, et al. Biochemical and structural characterization of murine GBP7, a guanylate binding protein with an elongated C-terminal tail. *Biochem J* (2019) 476(21):3161–82. doi: 10.1042/bcj20190364

13. Loschwitz J, Steffens N, Wang X, Schäffler M, Pfeffer K, Degrandi D, et al. Domain motions, dimerization, and membrane interactions of the murine guanylate binding protein 2. *Sci Rep* (2023) 13(1):679. doi: 10.1038/s41598-023-27520-8

14. Schelle L, Corte-Real JV, Esteves PJ, Abrantes J, Baldauf HM. Functional cross-species conservation of guanylate-binding proteins in innate immunity. *Med Microbiol Immunol* (2022) 212(2):141–52. doi: 10.1007/s00430-022-00736-7

15. Olszewski MA, Gray J, Vestal DJ. In silico genomic analysis of the human and murine guanylate-binding protein (GBP) gene clusters. *J Interferon Cytokine Res* (2006) 26(5):328–52. doi: 10.1089/jir.2006.26.328

16. Côrte-Real JV, Baldauf H-M, 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* (2021) 132:79–81. doi: 10.1016/j.molimm.2021.01.025

17. Côrte-Real JV, Baldauf H-M, 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* (2022) 13:752186. doi: 10.3389/fimmu.2022.752186

18. Smith AT. The world of pikas. In: Alves PC, Ferrand N, Hackländer K, editors. *Lagomorph biology: evolution, ecology, and conservation*. Berlin, Heidelberg: Springer Berlin Heidelberg (2008). p. 89–102.

19. Lanier HC, Olson LE. Inferring divergence times within pikas (Ochotona spp.) using mtDNA and relaxed molecular dating techniques. *Mol Phylogenet Evol* (2009) 53(1):1–12. doi: 10.1016/j.ympev.2009.05.035

20. Lissovsky AA. Taxonomic revision of pikas Ochotona (Lagomorpha, Mammalia) at the species level. *Mammalia* (2014) 78(2):199–216. doi: 10.1515/mammalia-2012-0134

21. Melo-Ferreira J, Lemos de Matos A, Areal H, Lissovsky AA, Carneiro M, Esteves PJ. The phylogeny of pikas (Ochotona) inferred from a multilocus coalescent approach. *Mol Phylogenet Evol* (2015) 84:240–4. doi: 10.1016/j.ympev.2015.01.004

22. Matthee CA, van Vuuren BJ, Bell D, Robinson TJ. A molecular supermatrix of the rabbits and hares (Leporidae) allows for the identification of five intercontinental exchanges during the Miocene. *Syst Biol* (2004) 53(3):433–47. doi: 10.1080/10635150490445715

23. Ge D, Wen Z, Xia L, Zhang Z, Erbajeva M, Huang C, et al. Evolutionary history of lagomorphs in response to global environmental change. *PloS One* (2013) 8(4):e59668. doi: 10.1371/journal.pone.0059668

24. Pinheiro A, Neves F, Lemos de Matos A, Abrantes J, van der Loo W, Mage R, et al. An overview of the lagomorph immune system and its genetic diversity. *Immunogenetics* (2016) 68(2):83–107. doi: 10.1007/s00251-015-0868-8

25. Esteves PJ, Abrantes J, Baldauf H-M, BenMohamed L, Chen Y, Christensen N, et al. The wide utility of rabbits as models of human diseases. *Exp Mol Med* (2018) 50(5):1–10. doi: 10.1038/s12276-018-0094-1

26. Soares J, Pinheiro A, Esteves PJ. The rabbit as an animal model to study innate immunity genes: Is it better than mice? *Front Immunol* (2022) 13:981815. doi: 10.3389/fimmu.2022.981815

27. Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SD. Automated phylogenetic detection of recombination using a genetic algorithm. *Mol Biol Evol* (2006) 23(10):1891–901. doi: 10.1093/molbev/msl051

28. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* (2021) 49(W1):W293–6. doi: 10.1093/nar/gkab301

29. Benton MJ, Donoghue PC. Paleontological evidence to date the tree of life. *Mol Biol Evol* (2007) 24(1):26–53. doi: 10.1093/molbev/msl150

30. Kriegs JO, Zemann A, Churakov G, Matzke A, Ohme M, Zischler H, et al. Retroposon insertions provide insights into deep lagomorph evolution. *Mol Biol Evol* (2010) 27(12):2678–81. doi: 10.1093/molbev/msq162

31. Sigrist CJA, Cerutti L, Hulo N, Gattiker A, Falquet L, Pagni M, et al. PROSITE: A documented database using patterns and profiles as motif descriptors. *Briefings Bioinf* (2002) 3(3):265–74. doi: 10.1093/bib/3.3.265

32. de Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, et al. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res* (2006) 34(Web Server issue):W362–365. doi: 10.1093/nar/gkl124

33. Sigrist CJ, de Castro E, Cerutti L, Cuche BA, Hulo N, Bridge A, et al. New and continuing developments at PROSITE. *Nucleic Acids Res* (2013) 41(Database issue): D344–347. doi: 10.1093/nar/gks1067

34. Britzen-Laurent N, Bauer M, Berton V, Fischer N, Syguda A, Reipschläger S, et al. Intracellular trafficking of guanylate-binding proteins is regulated by heterodimerization in a hierarchical manner. *PloS One* (2010) 5(12):e14246. doi: 10.1371/journal.pone.0014246

35. Shydlovskyi S, Zienert AY, Ince S, Dovengerds C, Hohendahl A, Dargazanli JM, et al. Nucleotide-dependent farnesyl switch orchestrates polymerization and membrane binding of human guanylate-binding protein. *Proc Natl Acad Sci United States America* (2017) 114(28):E5559–68. doi: 10.1073/pnas.1620959114

36. Sistemich L, Kutsch M, Hämisch B, Zhang P, Shydlovskyi S, Britzen-Laurent N, et al. The molecular mechanism of polymer formation of farnesylated human guanylate-binding protein 1. *J Mol Biol* (2020) 432(7):2164–85. doi: 10.1016/j.jmb.2020.02.009

37. Cui W, Braun E, Wang W, Tang J, Zheng Y, Slater B, et al. Structural basis for GTP-induced dimerization and antiviral function of guanylate-binding proteins. *Proc Natl Acad Sci United States America* (2021) 118(15):1–9. doi: 10.1073/pnas.2022269118

38. Sistemich L, Dimitrov Stanchev L, Kutsch M, Roux A, Günther Pomorski T, Herrmann C. Structural requirements for membrane binding of human guanylate-binding protein 1. *FEBS J* (2021) 288(13):4098–114. doi: 10.1111/febs.15703

39. Carneiro M, Rubin C-J, Di Palma F, Albert FW, Alföldi J, Barrio AM, et al. Rabbit genome analysis reveals a polygenic basis for phenotypic change during domestication. *Science* (2014) 345(6200):1074–9. doi: 10.1126/science.1253714

40. Tripal P, Bauer M, Naschberger E, Mörtinger T, Hohenadl C, Cornali E, et al. Unique features of different members of the human guanylate-binding protein family. *J Interferon Cytokine Res* (2007) 27(1):44–52. doi: 10.1089/jir.2007.0086

41. Braun E, Hotter D, Koepke L, Zech F, Groß R, Sparrer KMJ, et al. Guanylate-binding proteins 2 and 5 exert broad antiviral activity by inhibiting furin-mediated processing of viral envelope proteins. *Cell Rep* (2019) 27(7):2092–2104.e2010. doi: 10.1016/j.celrep.2019.04.063

42. Braun E, Sauter D. Furin-mediated protein processing in infectious diseases and cancer. *Clin Trans Immunol* (2019) 8(8):e1073. doi: 10.1002/cti2.1073

43. Tretina K, Park E-S, Maminska A, MacMicking JD. Interferon-induced guanylate-binding proteins: Guardians of host defense in health and disease. *J Exp Med* (2019) 216(3):482–500.

44. Kutsch M, Coers J. Human guanylate binding proteins: nanomachines orchestrating host defense. *FEBS J* (2020) 288(20):5826–49. doi: 10.1111/febs.15662

45. Zhang R, Li Z, Tang Y-D, Su C, Zheng C. When human guanylate-binding proteins meet viral infections. *J Biomed Sci* (2021) 28(1):17. doi: 10.1186/s12929-021-00716-8

46. Nei M, Gu X, Sitnikova T. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proc Natl Acad Sci United States America* (1997) 94(15):7799–806. doi: 10.1073/pnas.94.15.7799

47. Nei M, Rooney AP. Concerted and birth-and-death evolution of multigene families. *Annu Rev Genet* (2005) 39:121–52. doi: 10.1146/annurev.genet.39.073003.112240

48. Santos JC, Boucher D, Schneider LK, Demarco B, Dilucca M, Shkarina K, et al. Human GBP1 binds LPS to initiate assembly of a caspase-4 activating platform on cytosolic bacteria. *Nat Commun* (2020) 11(1):3276. doi: 10.1038/s41467-020-16889-z

49. Nantais DE, Schwemmle M, Stickney JT, Vestal DJ, Buss JE. Prenylation of an interferon-gamma-induced GTP-binding protein: the human guanylate binding protein, huGBP1. *J Leukoc Biol* (1996) 60(3):423–31. doi: 10.1002/jlb.60.3.423

50. Stickney JT, Buss JE. Murine guanylate-binding protein: incomplete geranylgeranyl isoprenoid modification of an interferon-γ–inducible guanosine triphosphate-binding protein. *Mol Biol Cell* (2000) 11(7):2191–200. doi: 10.1091/mbc.11.7.2191

51. Vestal DJ, Gorbacheva VY, Sen GC. Different subcellular localizations for the related interferon-induced GTPases, muGBP-1 and muGBP-2: implications for different functions? *J Interferon Cytokine Res* (2000) 20(11):991–1000. doi: 10.1089/10799900050198435

52. Modiano N, Lu YE, Cresswell P. Golgi targeting of human guanylate-binding protein-1 requires nucleotide binding, isoprenylation, and an IFN-gamma-inducible cofactor. *Proc Natl Acad Sci United States America* (2005) 102(24):8680–5. doi: 10.1073/pnas.0503227102

53. Lorenz C, Ince S, Zhang T, Cousin A, Batra-Safferling R, Nagel-Steger L, et al. Farnesylation of human guanylate-binding protein 1 as safety mechanism preventing structural rearrangements and uninduced dimerization. *FEBS J* (2020) 287(3):496–514. doi: 10.1111/febs.15015

54. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids symposium Ser* (1999) 41 (41):95–8.

55. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* (1994) 22(22):4673–80. doi: 10.1093/nar/22.22.4673

56. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* (2018) 35 (6):1547–9. doi: 10.1093/molbev/msy096

57. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* (1992) 8(3):275–82. doi: 10.1093/bioinformatics/8.3.275

58. Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* (2021) 30(1):70–82. doi: 10.1002/pro.3943

59. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nat Methods* (2022) 19(6):679–82. doi: 10.1038/s41592-022-01488-1

60. Matzek D, Baldauf H-M, Schieweck R, Popper B. Evaluation of a configurable, mobile and modular floor-pen system for group-housing of laboratory rabbits. *Animals* (2021) 11(4):977.

61. Spieker-Polet H, Yam P-C, Knight KL. Functional analysis of Iα Promoter regions of multiple igA heavy chain genes1. *J Immunol* (2002) 168(7):3360–8. doi: 10.4049/jimmunol.168.7.3360

62. Hague BF, Sawasdikosol S, Brown TJ, Lee K, Recker DP, Kindt TJ. CD4 and its role in infection of rabbit cell lines by human immunodeficiency virus type 1. *Proceedings of the National Academy of Sciences of the United States of America PNAS* (1992) 89(17):7963–67.

63. Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. *Bioinformatics* (2007) 23(10):1289–91. doi: 10.1093/bioinformatics/btm091

64. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3—new capabilities and interfaces. *Nucleic Acids Res* (2012) 40(15):e115–5. doi: 10.1093/nar/gks596

References

Ajoge, H.O. *et al.* (2023) 'Antiretroviral APOBEC3 cytidine deaminases alter HIV-1 provirus integration site profiles', *Nature Communications*, 14(1). Available at: https://doi.org/10.1038/s41467-022-35379-y.

Akira, S., Uematsu, S. and Takeuchi, O. (2006) 'Pathogen recognition and innate immunity', *Cell*, 124(4), pp. 783–801. Available at: https://doi.org/10.1016/j.cell.2006.02.015.

Allouch, A. *et al.* (2011) 'The TRIM family protein KAP1 inhibits HIV-1 integration', *Cell Host and Microbe*, 9(6), pp. 484–495. Available at: https://doi.org/10.1016/j.chom.2011.05.004.

Allouch, A. *et al.* (2013) 'P21-mediated RNR2 repression restricts HIV-1 replication in macrophages by inhibiting dNTP biosynthesis pathway', *Proceedings of the National Academy of Sciences of the United States of America*, 110(42). Available at: https://doi.org/10.1073/pnas.1306719110.

Bailey, C.C. *et al.* (2013) 'Interferon-induced transmembrane protein 3 is a type II transmembrane protein', *Journal of Biological Chemistry*, 288(45), pp. 32184–32193. Available at: https://doi.org/10.1074/jbc.M113.514356.

Bailey, C.C. *et al.* (2014) 'IFITM-family proteins: The cell's first line of antiviral defense', *Annual Review of Virology*, 1(1), pp. 261–283. Available at: https://doi.org/10.1146/annurev-virology-031413-085537.

Baldauf, H.M. *et al.* (2017) 'Vpx overcomes a SAMHD1-independent block to HIV reverse transcription that is specific to resting CD4 T cells', *Proceedings of the National Academy of Sciences of the United States of America*, 114(10), pp. 2729–2734. Available at: https://doi.org/10.1073/pnas.1613635114.

Barr, S.D., Smiley, J.R. and Bushman, F.D. (2008) 'The interferon response inhibits HIV particle production by induction of TRIM22', *PLoS Pathogens*, 4(2), pp. 1–11. Available at: https://doi.org/10.1371/journal.ppat.1000007.

Bergh, R. Van den *et al.* (2009) 'Do monocytes use the novel adipocytokine Visfatin/NAMPT/PBEF1 to flip the HIV coreceptor switch?', *Retrovirology*, 6(S2), p. P88. Available at: https://doi.org/10.1186/1742-4690-6-S2-P88.

Van den Bergh, R. *et al.* (2012) 'Monocytes Contribute to Differential Immune Pressure on R5 versus X4 HIV through the Adipocytokine Visfatin/NAMPT', *PLoS ONE*. Edited by T.B.H. Geijtenbeek, 7(4), p. e35074. Available at: https://doi.org/10.1371/journal.pone.0035074.

Braun, E. *et al.* (2019) 'Guanylate-Binding Proteins 2 and 5 Exert Broad Antiviral Activity by Inhibiting Furin-Mediated Processing of Viral Envelope Proteins', *Cell Reports*, 27(7), pp. 2092-2104.e10. Available at: https://doi.org/10.1016/j.celrep.2019.04.063.

Burdick, R.C. *et al.* (2020) 'HIV-1 uncoats in the nucleus near sites of integration', *Proceedings of the National Academy of Sciences of the United States of America*, 117(10), pp. 5486–5493. Available at: https://doi.org/10.1073/pnas.1920631117.

Burkhard, M. and Dean, G. (2005) 'Transmission and Immunopathogenesis of FIV in Cats as a Model for HIV', *Current HIV Research*, 1(1), pp. 15–29. Available at: https://doi.org/10.2174/1570162033352101.

Casola, C. and Betrán, E. (2017) 'The genomic impact of gene retrocopies: What have we learned from comparative genomics, population genomics, and transcriptomic analyses?', *Genome Biology and Evolution*, 9(6), pp. 1351–1373. Available at: https://doi.org/10.1093/gbe/evx081.

Cheetham, S.W., Faulkner, G.J. and Dinger, M.E. (2020) 'Overcoming challenges and dogmas to understand the functions of pseudogenes', *Nature Reviews Genetics*, 21(3), pp. 191–201. Available at: https://doi.org/10.1038/s41576-019-0196-1.

Chougui, G. *et al.* (2018) 'HIV-2/SIV viral protein X counteracts HUSH repressor complex', *Nature Microbiology*, 3(8), pp. 891–897. Available at: https://doi.org/10.1038/s41564-018-0179-6.

Cobos Jiménez, V. *et al.* (2015) 'G3BP1 restricts HIV-1 replication in macrophages and T-cells by sequestering viral RNA', *Virology*, 486, pp. 94–104. Available at: https://doi.org/10.1016/j.virol.2015.09.007.

Coffin, J. *et al.* (2021) 'ICTV Virus Taxonomy Profile: Retroviridae 2021', *Journal of General Virology*, 102(12), pp. 1–2. Available at: https://doi.org/10.1099/JGV.0.001712.

Colomer-Lluch, M. and Serra-Moreno, R. (2017) 'BCA2/Rabring7 Interferes with HIV-1 Proviral Transcription by Enhancing the SUMOylation of IκBα', *Journal of Virology*. Edited by G. Silvestri, 91(8), pp. 1–22. Available at: https://doi.org/10.1128/JVI.02098-16.

Cook, R.F., Leroux, C. and Issel, C.J. (2013) 'Equine infectious anemia and equine infectious anemia virus in 2013: A review', *Veterinary Microbiology*, 167(1–2), pp. 181–204. Available at: https://doi.org/10.1016/j.vetmic.2013.09.031.

Van Damme, N. *et al.* (2008) 'The Interferon-Induced Protein BST-2 Restricts HIV-1 Release and Is Downregulated from the Cell Surface by the Viral Vpu Protein', *Cell Host and Microbe*, 3(4), pp. 245–252. Available at: https://doi.org/10.1016/j.chom.2008.03.001.

Daugherty, M.D. and Malik, H.S. (2012) 'Rules of engagement: Molecular insights from host-virus arms races', *Annual Review of Genetics*, 46, pp. 677–700. Available at: https://doi.org/10.1146/annurev-genet-110711-155522.

Deeks, S.G. *et al.* (2015) 'HIV infection', *Nature Reviews Disease Primers*, 1(1), pp. 1–22. Available at: https://doi.org/10.1038/nrdp.2015.35.

Dharan, A. *et al.* (2020) 'Nuclear pore blockade reveals that HIV-1 completes reverse transcription and uncoating in the nucleus', *Nature Microbiology*, 5(9), pp. 1088–1095. Available at: https://doi.org/10.1038/s41564-020-0735-8.

Dharan, A. and Campbell, E.M. (2022) 'Teaching old dogmas new tricks: recent insights into the nuclear import of HIV-1', *Current Opinion in Virology*, 53, p. 101203. Available at: https://doi.org/10.1016/j.coviro.2022.101203.

Diamond, M.S. and Farzan, M. (2013) 'The broad-spectrum antiviral functions of IFIT and IFITM proteins', *Nature Reviews Immunology*, 13(1), pp. 46–57. Available at: https://doi.org/10.1038/nri3344.

Dietrich, I. *et al.* (2011) 'Feline Tetherin Efficiently Restricts Release of Feline Immunodeficiency Virus but Not Spreading of Infection', *Journal of Virology*, 85(12), pp. 5840–5852. Available at: https://doi.org/10.1128/jvi.00071-11.

Duggal, N.K. and Emerman, M. (2012) 'Evolutionary conflicts between viruses and restriction factors shape immunity', *Nature Reviews Immunology*, 12(10), pp. 687–695. Available at: https://doi.org/10.1038/nri3295.

Ebert, D. and Fields, P.D. (2020) 'Host–parasite co-evolution and its genomic signature', *Nature Reviews Genetics*, 21(12), pp. 754–768. Available at: https://doi.org/10.1038/s41576-020-0269-1.

Esnault, C. *et al.* (2008) 'Restriction by APOBEC3 proteins of endogenous retroviruses with an extracellular life cycle: Ex vivo effects and in vivo "traces" on the murine IAPE and human HERV-K elements', *Retrovirology*, 5, pp. 1–11. Available at: https://doi.org/10.1186/1742-4690-5-75.

Esteves, P.J. *et al.* (2018) 'The wide utility of rabbits as models of human diseases', *Experimental and Molecular Medicine*, 50(5), pp. 1–10. Available at: https://doi.org/10.1038/s12276-018-0094-1.

Fan, H. (1997) 'Leukemogenesis by Moloney murine leukemia virus: A multistep process', *Trends in Microbiology*, 5(2), pp. 74–82. Available at: https://doi.org/10.1016/S0966-842X(96)10076-7.

Forlani, G. *et al.* (2019) 'Restriction factors in human retrovirus infections and the unprecedented case of CIITA as link of intrinsic and adaptive immunity against HTLV-1', *Retrovirology*, 16(1), pp. 1–12. Available at: https://doi.org/10.1186/s12977-019-0498-6.

Francis, A.C. *et al.* (2020) 'HIV-1 replication complexes accumulate in nuclear speckles and integrate into speckle-associated genomic domains', *Nature Communications*, 11(1). Available at: https://doi.org/10.1038/s41467-020-17256-8.

Freed, E.O. (2015) 'HIV-1 assembly, release and maturation', *Nature Reviews Microbiology*, 13(8), pp. 484–496. Available at: https://doi.org/10.1038/nrmicro3490.

Friedlová, N. *et al.* (2022) 'IFITM protein regulation and functions: Far beyond the fight against viruses', *Frontiers in Immunology*, 13(November), pp. 1–24. Available at: https://doi.org/10.3389/fimmu.2022.1042368.

Friedman, R.L. *et al.* (1984) 'Transcriptional and posttranscriptional regulation of interferon-induced gene expression in human cells', *Cell*, 38(3), pp. 745–755. Available at: https://doi.org/10.1016/0092-8674(84)90270-8.

Ganser-Pornillos, B.K. and Pornillos, O. (2019) 'Restriction of HIV-1 and other retroviruses by TRIM5', *Nature Reviews Microbiology*, 17(9), pp. 546–556. Available at: https://doi.org/10.1038/s41579-019-0225-2.

Gao, G., Guo, X. and Goff, S.P. (2002) 'Inhibition of retroviral RNA production by ZAP, a CCCH-type zinc finger protein', *Science*, 297(5587), pp. 1703–1706. Available at: https://doi.org/10.1126/science.1074276.

Gómez-Herranz, M., Taylor, J. and Sloan, R.D. (2023) 'IFITM proteins: Understanding their diverse roles in viral infection, cancer, and immunity', *Journal of Biological Chemistry*, 299(1), p. 102741. Available at: https://doi.org/10.1016/j.jbc.2022.102741.

Goujon, C. *et al.* (2013) 'Human MX2 is an interferon-induced post-entry inhibitor of HIV-1 infection', *Nature*, 502(7472), pp. 559–562. Available at: https://doi.org/10.1038/nature12542.

Gramberg, T. *et al.* (2013) 'Restriction of diverse retroviruses by SAMHD1', *Retrovirology*, 10(1), pp. 1–12. Available at: https://doi.org/10.1186/1742-4690-10-26.

Graziano, F. *et al.* (2018) 'Reversible Human Immunodeficiency Virus Type-1 Latency in Primary Human Monocyte-Derived Macrophages Induced by Sustained M1 Polarization', *Scientific Reports*, 8(1), pp. 1–13. Available at: https://doi.org/10.1038/s41598-018-32451-w.

Guo, G. *et al.* (2021) 'Human Schlafen 11 exploits codon preference discrimination to attenuate viral protein synthesis of prototype foamy virus (PFV)', *Virology*, 555(September 2020), pp. 78–88. Available at: https://doi.org/10.1016/j.virol.2020.12.015.

Hanagata, N. *et al.* (2011) 'Characterization of the osteoblast-specific transmembrane protein IFITM5 and analysis of IFITM5-deficient mice', *Journal of Bone and Mineral Metabolism*, 29(3), pp. 279–290. Available at: https://doi.org/10.1007/s00774-010-0221-0.

Harris, R.S. and Liddament, M.T. (2004) 'Retroviral restriction by APOBEC proteins', *Nature Reviews Immunology*, 4(11), pp. 868–877. Available at: https://doi.org/10.1038/nri1489.

Hatziioannou, T. and Evans, D.T. (2012) 'Animal models for HIV/AIDS research', *Nature Reviews Microbiology*, 10(12), pp. 852–867. Available at: https://doi.org/10.1038/nrmicro2911.

Hayward, A. (2017) 'Origin of the retroviruses: when, where, and how?', *Current Opinion in Virology*, 25, pp. 23–27. Available at: https://doi.org/10.1016/j.coviro.2017.06.006.

Hedjazi, G. *et al.* (2022) *Alterations of bone material properties in growing Ifitm5/BRIL p.S42 knock-in mice, a new model for atypical type VI osteogenesis imperfecta*, *Bone*. Available at: https://doi.org/10.1016/j.bone.2022.116451.

Henriet, S. *et al.* (2009) 'Tumultuous Relationship between the Human Immunodeficiency Virus Type 1 Viral Infectivity Factor (Vif) and the Human APOBEC-3G and APOBEC-3F Restriction Factors', *Microbiology and Molecular Biology Reviews*, 73(2), pp. 211–232. Available at: https://doi.org/10.1128/mmbr.00040-08.

Hessell, A.J. and Haigwood, N.L. (2015) 'Animal models in HIV-1 protection and therapy', *Current Opinion in HIV and AIDS*, 10(3), pp. 170–176. Available at: https://doi.org/10.1097/COH.0000000000000152.

Hickford, D. *et al.* (2012) 'Evolution of vertebrate interferon inducible transmembrane proteins', *BMC Genomics*, 13(1). Available at: https://doi.org/10.1186/1471-2164-13-155.

Hofacre, A. and Fan, H. (2010) 'Jaagsiekte sheep retrovirus biology and oncogenesis', *Viruses*, 2(12), pp. 2618–2648. Available at: https://doi.org/10.3390/v2122618.

Hulo, C. *et al.* (2011) 'ViralZone: A knowledge resource to understand virus diversity', *Nucleic Acids Research*, 39(SUPPL. 1), pp. 576–582. Available at: https://doi.org/10.1093/nar/gkq901.

Jasinska, A.J., Apetrei, C. and Pandrea, I. (2023) 'Walk on the wild side: SIV infection in African non-human primate hosts—from the field to the laboratory', *Frontiers in Immunology*, 13(January), pp. 1–20. Available at: https://doi.org/10.3389/fimmu.2022.1060985.

Jolly, C., Booth, N.J. and Neil, S.J.D. (2010) 'Cell-Cell Spread of Human Immunodeficiency Virus Type 1 Overcomes Tetherin/BST-2-Mediated Restriction in T cells', *Journal of Virology*, 84(23), pp. 12185–12199. Available at: https://doi.org/10.1128/jvi.01447-10.

Jouvenet, N. *et al.* (2009) 'Broad-Spectrum Inhibition of Retroviral and Filoviral Particle Release by Tetherin', *Journal of Virology*, 83(4), pp. 1837–1844. Available at: https://doi.org/10.1128/jvi.02211-08.

Juliarena, M.A. *et al.* (2017) 'Bovine leukemia virus: Current perspectives', *Virus Adaptation and Treatment*, 9, pp. 13–26. Available at: https://doi.org/10.2147/VAAT.S113947.

Jung, Y.T. *et al.* (2002) ' Characterization of a Polytropic Murine Leukemia Virus Proviral Sequence Associated with the Virus Resistance Gene Rmcf of DBA/2 Mice ', *Journal of Virology*, 76(16), pp. 8218–8224. Available at: https://doi.org/10.1128/jvi.76.16.8218-8224.2002.

Kaessmann, H., Vinckenbosch, N. and Long, M. (2009) 'RNA-based gene duplication: Mechanistic and evolutionary insights', *Nature Reviews Genetics*, 10(1), pp. 19–31. Available at: https://doi.org/10.1038/nrg2487.

Kane, M. *et al.* (2013) 'MX2 is an interferon-induced inhibitor of HIV-1 infection', *Nature*, 502(7472), pp. 563–566. Available at: https://doi.org/10.1038/nature12653.

Kayesh, M.E.H. *et al.* (2021) 'Tree shrew as an emerging small animal model for human viral infection: A recent overview', *Viruses*, 13(8), pp. 1–13. Available at: https://doi.org/10.3390/v13081641.

Klatt, N.R., Silvestri, G. and Hirsch, V. (2012) 'Nonpathogenic Simian Immunodeficiency Virus Infections', *Cold Spring Harbor Perspectives in Medicine*, 2(1), pp. a007153–a007153. Available at: https://doi.org/10.1101/cshperspect.a007153.

Krapp, C. *et al.* (2016) 'Guanylate Binding Protein (GBP) 5 Is an Interferon-Inducible Inhibitor of HIV-1 Infectivity', *Cell Host and Microbe*, 19(4), pp. 504–514. Available at: https://doi.org/10.1016/j.chom.2016.02.019.

Kriesel, F., Schelle, L. and Baldauf, H. (2020) 'Same same but different - Antiviral factors interfering with the infectivity of HIV particles', *Microbes and Infection*, 22(9), pp. 416–422. Available at: https://doi.org/10.1016/j.micinf.2020.05.009.

Kuang, Z., Seo, E.J. and Leis, J. (2011) 'Mechanism of Inhibition of Retrovirus Release from Cells by Interferon-Induced Gene ISG15', *Journal of Virology*, 85(14), pp. 7153–7161. Available at: https://doi.org/10.1128/jvi.02610-10.

Kulaga, H. *et al.* (1988) 'Infection of rabbit T-cell and macrophage lines with human immunodeficiency virus', *Proceedings of the National Academy of Sciences of the United States of America*, 85(12), pp. 4455–4459. Available at: https://doi.org/10.1073/pnas.85.12.4455.

Kumar, N., Chahroudi, A. and Silvestri, G. (2016) 'Animal models to achieve an HIV cure', *Current Opinion in HIV and AIDS*, 11(4), pp. 432–441. Available at: https://doi.org/10.1097/COH.0000000000000290.

Kutsch, M. and Coers, J. (2021) 'Human guanylate binding proteins: nanomachines orchestrating host defense', *The FEBS Journal*, 288(20), pp. 5826–5849. Available at: https://doi.org/10.1111/febs.15662.

Laguette, N. *et al.* (2011) 'SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx', *Nature*, 474(7353), pp. 654–657. Available at: https://doi.org/10.1038/nature10117.

Lahouassa, H. *et al.* (2012) 'SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates', *Nature Immunology*, 13(3), pp. 223–228. Available at: https://doi.org/10.1038/ni.2236.

Lefkowitz, E.J. *et al.* (2018) 'Virus taxonomy: The database of the International Committee on Taxonomy of Viruses (ICTV)', *Nucleic Acids Research*, 46(D1), pp. D708–D717. Available at: https://doi.org/10.1093/nar/gkx932.

Leroux, C., Cadore, J.-L. and Montelaro, R.C. (2004) 'Equine Infectious Anemia Virus (EIAV): what has HIV?s country cousin got to tell us?', *Veterinary Research*, 35(4), pp. 485–512. Available at: https://doi.org/10.1051/vetres:2004020.

Li, K. *et al.* (2013) 'IFITM Proteins Restrict Viral Membrane Hemifusion', *PLoS Pathogens*, 9(1). Available at: https://doi.org/10.1371/journal.ppat.1003124.

Li, M. *et al.* (2012) 'Codon-usage-based inhibition of HIV protein synthesis by human schlafen 11', *Nature*, 491(7422), pp. 125–128. Available at: https://doi.org/10.1038/nature11433.

Liao, Y. *et al.* (2019) 'Functional involvement of interferon-inducible transmembrane proteins in antiviral immunity', *Frontiers in Microbiology*, 10(MAY), pp. 1–9. Available at: https://doi.org/10.3389/fmicb.2019.01097.

Lim, E.S. *et al.* (2012) 'The ability of primate lentiviruses to degrade the monocyte restriction factor SAMHD1 preceded the birth of the viral accessory protein Vpx', *Cell Host and Microbe*, 11(2), pp. 194–204. Available at: https://doi.org/10.1016/j.chom.2012.01.004.

Lin, Y.Z. *et al.* (2016) 'Equine schlafen 11 restricts the production of equine infectious anemia virus via a codon usage-dependent mechanism', *Virology*, 495, pp. 112–121. Available at: https://doi.org/10.1016/j.virol.2016.04.024.

Little, T.J. *et al.* (2010) 'The coevolution of virulence: Tolerance in perspective', *PLoS Pathogens*, 6(9). Available at: https://doi.org/10.1371/journal.ppat.1001006.

Liu, D. (2015) 'Feline Immunodeficiency Virus', *Molecular Detection of Animal Viral Pathogens*, pp. 191–196. Available at: https://doi.org/10.1136/inpract.11.3.87.

Liu, S.-Y. *et al.* (2013) 'Interferon-Inducible Cholesterol-25-Hydroxylase Broadly Inhibits Viral Entry by Production of 25-Hydroxycholesterol', *Immunity*, 38(1), pp. 92–105. Available at: https://doi.org/10.1016/j.immuni.2012.11.005.

Liu, Z. *et al.* (2013) 'The interferon-inducible MxB protein inhibits HIV-1 infection', *Cell Host and Microbe*, 14(4), pp. 398–410. Available at: https://doi.org/10.1016/j.chom.2013.08.015.

Lodermeyer, V. *et al.* (2013) '90K, an interferon-stimulated gene product, reduces the infectivity of HIV-1', *Retrovirology*, 10(1), pp. 1–18. Available at: https://doi.org/10.1186/1742-4690-10-111.

Lodermeyer, V. *et al.* (2018) 'The Antiviral Activity of the Cellular Glycoprotein LGALS3BP/90K Is Species Specific', *Journal of Virology*. Edited by G. Silvestri, 92(14). Available at: https://doi.org/10.1128/JVI.00226-18.

Lu, J. *et al.* (2011) 'The IFITM Proteins Inhibit HIV-1 Infection', *Journal of Virology*, 85(5), pp. 2126–2137. Available at: https://doi.org/10.1128/jvi.01531-10.

Lubow, J. *et al.* (2020) 'Mannose receptor is an HIV restriction factor counteracted by Vpr in macrophages', *eLife*, 9, pp. 1–31. Available at: https://doi.org/10.7554/eLife.51035.

Lun, C.M. *et al.* (2021) 'Mechanism of viral glycoprotein targeting by membrane-associated ring-ch proteins', *mBio*, 12(2), pp. 1–26. Available at: https://doi.org/10.1128/mBio.00219-21.

Luo, M.-T. *et al.* (2021) 'Tree Shrew Cells Transduced with Human CD4 and CCR5 Support Early Steps of HIV-1 Replication, but Viral Infectivity Is Restricted by APOBEC3', *Journal of Virology*, 95(16). Available at: https://doi.org/10.1128/jvi.00020-21.

Mantovani, A. and Garlanda, C. (2023) 'Humoral Innate Immunity and Acute-Phase Proteins', *New England Journal of Medicine*, 388(5), pp. 439–452. Available at: https://doi.org/10.1056/nejmra2206346.

Marshall, J.S. *et al.* (2018) 'An introduction to immunology and immunopathology', *Allergy, Asthma and Clinical Immunology*, 14(s2), pp. 1–10. Available at: https://doi.org/10.1186/s13223-018-0278-1.

Miyagi, E. *et al.* (2009) 'Vpu enhances HIV-1 virus release in the absence of Bst-2 cell surface down-modulation and intracellular depletion', *Proceedings of the National Academy of Sciences of the United States of America*, 106(8), pp. 2868–2873. Available at: https://doi.org/10.1073/pnas.0813223106.

Miyakawa, K. *et al.* (2009) 'BCA2/rabring7 promotes tetherin-dependent HIV-1 restriction', *PLoS Pathogens*, 5(12), pp. 21–25. Available at: https://doi.org/10.1371/journal.ppat.1000700.

Miyazato, P. *et al.* (2019) 'HTLV-1 contains a high CG dinucleotide content and is susceptible to the host antiviral protein ZAP', *Retrovirology*, 16(1), pp. 1–11. Available at: https://doi.org/10.1186/s12977-019-0500-3.

Moretti, J. and Blander, J.M. (2017) 'Cell-autonomous stress responses in innate immunity', *Journal of Leukocyte Biology*, 101(1), pp. 77–86. Available at: https://doi.org/10.1189/jlb.2mr0416-201r.

Muller, T.G. *et al.* (2022) 'Nuclear Capsid Uncoating and Reverse Transcription of HIV-1', *Annual Review of Virology*, 9, pp. 261–284. Available at: https://doi.org/10.1146/annurev-virology-020922-110929.

Müller, T.G. *et al.* (2021) 'Hiv-1 uncoating by release of viral cdna from capsid-like structures in the nucleus of infected cells', *eLife*, 10, pp. 1–32. Available at: https://doi.org/10.7554/ELIFE.64776.

Münk, C. *et al.* (1997) 'Amphotropic murine leukemia viruses induce spongiform encephalomyelopathy', *Proceedings of the National Academy of Sciences of the United States of America*, 94(11), pp. 5837–5842. Available at: https://doi.org/10.1073/pnas.94.11.5837.

Nei, M. and Rooney, A.P. (2005) 'Concerted and birth-and-death evolution of multigene families', *Annual review of genetics*, 39, pp. 121–152. Available at: https://doi.org/10.1146/annurev.genet.39.073003.112240.

Neil, S.J.D., Zang, T. and Bieniasz, P.D. (2008) 'Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu', *Nature*, 451(7177), pp. 425–430. Available at: https://doi.org/10.1038/nature06553.

Nityanandam, R. and Serra-Moreno, R. (2014) 'BCA2/Rabring7 Targets HIV-1 Gag for Lysosomal Degradation in a Tetherin-Independent Manner', *PLoS Pathogens*, 10(5). Available at: https://doi.org/10.1371/journal.ppat.1004151.

Okumura, A. *et al.* (2006) 'Innate antiviral response targets HIV-1 release by the induction of ubiquitin-like protein ISG15', *Proceedings of the National Academy of Sciences of the United States of America*, 103(5), pp. 1440–1445. Available at: https://doi.org/10.1073/pnas.0510518103.

Paludan, S.R. *et al.* (2021) 'Constitutive immune mechanisms: mediators of host defence and immune regulation', *Nature Reviews Immunology*, 21(3), pp. 137–150. Available at: https://doi.org/10.1038/s41577-020-0391-5.

Payne, L.N. and Nair, V. (2012) 'The long view: 40 years of avian leukosis research', *Avian Pathology*, 41(1), pp. 11–19. Available at: https://doi.org/10.1080/03079457.2011.646237.

Pincetic, A. *et al.* (2010) 'The Interferon-Induced Gene ISG15 Blocks Retrovirus Release from Cells Late in the Budding Process', *Journal of Virology*, 84(9), pp. 4725–4736. Available at: https://doi.org/10.1128/jvi.02478-09.

Pincus, T., Hartley, J.W. and Rowe, W.P. (1971) 'A MAJOR GENETIC LOCUS AFFECTING RESISTANCE TO INFECTION WITH MURINE LEUKEMIA VIRUSES', *The Journal of Experimental Medicine*, 133(6), pp. 1219–1233. Available at: https://doi.org/10.1084/jem.133.6.1219.

Policicchio, B.B., Pandrea, I. and Apetrei, C. (2016) 'Animal models for HIV cure research', *Frontiers in Immunology*, 7(12), pp. 1–15. Available at: https://doi.org/10.3389/fimmu.2016.00012.

Pornillos, O. and Ganser-Pornillos, B.K. (2019) 'Maturation of retroviruses', *Current Opinion in Virology*, 36, pp. 47–55. Available at: https://doi.org/10.1016/j.coviro.2019.05.004.

Proietti, F.A. *et al.* (2005) 'Global epidemiology of HTLV-I infection and associated diseases', *Oncogene*, 24(39), pp. 6058–6068. Available at: https://doi.org/10.1038/sj.onc.1208968.

Rahman, K. and Compton, A.A. (2021) 'The Indirect Antiviral Potential of Long Noncoding RNAs Encoded by IFITM Pseudogenes', *Journal of Virology*, 95(21), pp. 2–7. Available at: https://doi.org/10.1128/jvi.00680-21.

Ramdas, P. *et al.* (2020) 'From Entry to Egress: Strategic Exploitation of the Cellular Processes by HIV-1', *Frontiers in Microbiology*, 11(559792), pp. 1–18. Available at: https://doi.org/10.3389/fmicb.2020.559792.

Ramezani, S. *et al.* (2022) 'HTLV, a multi organ oncovirus', *Microbial Pathogenesis*, 169(105622). Available at: https://doi.org/10.1016/j.micpath.2022.105622.

Randow, F., MacMicking, J.D. and James, L.C. (2013) 'Cellular Self-Defense: How Cell-Autonomous Immunity Protects Against Pathogens', *Science*, 340(6133), pp. 701–706. Available at: https://doi.org/10.1126/science.1233028.

Rein, A. (2011) 'Murine Leukemia Viruses: Objects and Organisms', *Advances in Virology*, 2011, pp. 1–14. Available at: https://doi.org/10.1155/2011/403419.

Ren, L. *et al.* (2020) 'Current Progress on Host Antiviral Factor IFITMs', *Frontiers in Immunology*, 11(November), pp. 1–8. Available at: https://doi.org/10.3389/fimmu.2020.543444.

Rheinemann, L. *et al.* (2022) 'Recurrent evolution of an inhibitor of ESCRT-dependent virus budding and LINE-1 retrotransposition in primates', *Current Biology*, 32(7), pp. 1511-1522.e6. Available at: https://doi.org/10.1016/j.cub.2022.02.018.

Riera Romo, M., Pérez-Martínez, D. and Castillo Ferrer, C. (2016) 'Innate immunity in vertebrates: An overview', *Immunology*, 148(2), pp. 125–139. Available at: https://doi.org/10.1111/imm.12597.

Rosa, A. *et al.* (2015) 'HIV-1 Nef promotes infection by excluding SERINC5 from virion incorporation', *Nature*, 526(7572), pp. 212–217. Available at: https://doi.org/10.1038/nature15399.

Ross, S.R. (2010) 'Mouse mammary tumor virus molecular biology and oncogenesis', *Viruses*, 2(9), pp. 2000–2012. Available at: https://doi.org/10.3390/v2092000.

Rovnak, J. and Quackenbush, S.L. (2010) 'Walleye dermal sarcoma virus: Molecular biology and oncogenesis', *Viruses*, 2(9), pp. 1984–1999. Available at: https://doi.org/10.3390/v2091984.

Rowe, H.M. *et al.* (2010) 'KAP1 controls endogenous retroviruses in embryonic stem cells', *Nature*, 463(7278), pp. 237–240. Available at: https://doi.org/10.1038/nature08674.

Salas-Briceno, K., Zhao, W. and Ross, S.R. (2020) 'Mouse APOBEC3 restriction of retroviruses', *Viruses*, 12(11), pp. 1–12. Available at: https://doi.org/10.3390/v12111217.

Schaller, T., Hué, S. and Towers, G.J. (2007) 'An Active TRIM5 Protein in Rabbits Indicates a Common Antiviral Ancestor for Mammalian TRIM5 Proteins', *Journal of Virology*, 81(21), pp. 11713–11721. Available at: https://doi.org/10.1128/jvi.01468-07.

Schelle, L. *et al.* (2023) 'Functional cross-species conservation of guanylate-binding proteins in innate immunity', *Medical Microbiology and Immunology*, 212(2), pp. 141–152. Available at: https://doi.org/10.1007/s00430-022-00736-7.

Schneider, W.M., Chevillotte, M.D. and Rice, C.M. (2014) 'Interferon-stimulated genes: A complex web of host defenses', *Annual Review of Immunology*, 32, pp. 513–545. Available at: https://doi.org/10.1146/annurev-immunol-032713-120231.

Sebastian, S. and Luban, J. (2005) 'TRIM5α selectively binds a restriction-sensitive retroviral capsid', *Retrovirology*, 2, pp. 1–3. Available at: https://doi.org/10.1186/1742-4690-2-40.

Selyutina, A. *et al.* (2020) 'Nuclear Import of the HIV-1 Core Precedes Reverse Transcription and Uncoating', *Cell Reports*, 32(13), p. 108201. Available at: https://doi.org/10.1016/j.celrep.2020.108201.

Shi, B. *et al.* (2018) 'Inhibition of HIV early replication by the p53 and its downstream gene p21', *Virology Journal*, 15(1), pp. 1–13. Available at: https://doi.org/10.1186/s12985-018-0959-x.

Singh, R. *et al.* (2011) 'Association of TRIM22 with the Type 1 Interferon Response and Viral Control during Primary HIV-1 Infection', *Journal of Virology*, 85(1), pp. 208–216. Available at: https://doi.org/10.1128/jvi.01810-10.

Smith, S.E. *et al.* (2019) 'Interferon-Induced Transmembrane Protein 1 Restricts Replication of Viruses That Enter Cells via the Plasma Membrane', *Journal of Virology*. Edited by T.S. Dermody, 93(6), pp. e02003-18. Available at: https://doi.org/10.1128/JVI.02003-18.

Soares, J., Pinheiro, A. and Esteves, P.J. (2022) 'The rabbit as an animal model to study innate immunity genes: Is it better than mice?', *Frontiers in Immunology*, 13(September), pp. 1–7. Available at: https://doi.org/10.3389/fimmu.2022.981815.

Spencer, H.J. and Groupé, V. (1962a) 'Pathogenesis of virus-induced rous sarcoma. i. distribution of virus and tumor foci in chicks and turkeys', *Journal of the National Cancer Institute*, 29(3), pp. 397–419. Available at: https://doi.org/10.1093/jnci/29.3.397.

Spencer, H.J. and Groupé, V. (1962b) 'Pathogenesis of virus-induced rous sarcoma. ii. dynamics of tumor development and viral growth patterns in chicks', *Journal of the National Cancer Institute*, 29(3), pp. 421–441. Available at: https://doi.org/10.1093/jnci/29.3.421.

Stabell, A.C. *et al.* (2016) 'Non-human Primate Schlafen11 Inhibits Production of Both Host and Viral Proteins', *PLoS Pathogens*, 12(12), pp. 1–19. Available at: https://doi.org/10.1371/journal.ppat.1006066.

Staszak, K. and Makałowska, I. (2021) 'Cancer, retrogenes, and evolution', *Life*, 11(1), pp. 1–16. Available at: https://doi.org/10.3390/life11010072.

Stavrou, S. *et al.* (2014) 'Different Modes of Retrovirus Restriction by Human APOBEC3A and APOBEC3G In Vivo', *PLoS Pathogens*, 10(5). Available at: https://doi.org/10.1371/journal.ppat.1004145.

Stremlau, M. *et al.* (2004) 'The cytoplasmic body component TRIM5α restricts HIV-1 infection in Old World monkeys', *Nature*, 427(6977), pp. 848–853. Available at: https://doi.org/10.1038/nature02343.

Sykes, J.E. and Hartmann, K. (2014) 'Feline Leukemia Virus Infection', in *Canine and Feline Infectious Diseases*. Elsevier, pp. 224–238. Available at: https://doi.org/10.1016/B978-1-4377-0795-3.00022-3.

Tada, T. *et al.* (2015) 'March8 inhibits HIV-1 infection by reducing virion incorporation of envelope glycoproteins', *Nature Medicine*, 21(12), pp. 1502–1507. Available at: https://doi.org/10.1038/nm.3956.

Takeda, A. and Matano, T. (2007) 'Inhibition of infectious murine leukemia virus production by Fv-4 env gene products exerting dominant negative effect on viral envelope glycoprotein', *Microbes and Infection*, 9(14–15), pp. 1590–1596. Available at: https://doi.org/10.1016/j.micinf.2007.09.012.

Tartour, K. *et al.* (2014) 'IFITM proteins are incorporated onto HIV-1 virion particles and negatively imprint their infectivity', *Retrovirology*, 11(1), pp. 1–14. Available at: https://doi.org/10.1186/s12977-014-0103-y.

Taylor, G.M., Gao, Y. and Sanders, D.A. (2001) 'Fv-4: Identification of the Defect in Env and the Mechanism of Resistance to Ecotropic Murine Leukemia Virus', *Journal of Virology*, 75(22), pp. 11244–11248. Available at: https://doi.org/10.1128/jvi.75.22.11244-11248.2001.

Tervo, H.-M. and Keppler, O.T. (2010) 'High Natural Permissivity of Primary Rabbit Cells for HIV-1, with a Virion Infectivity Defect in Macrophages as the Final Replication Barrier', *Journal of Virology*, 84(23), pp. 12300–12314. Available at: https://doi.org/10.1128/jvi.01607-10.

Timilsina, U. *et al.* (2020) 'Serinc5 potently restricts retrovirus infection in vivo', *mBio*, 11(4), pp. 1–18. Available at: https://doi.org/10.1128/mBio.00588-20.

Tretina, K. *et al.* (2019) 'Interferon-induced guanylate-binding proteins: Guardians of host defense in health and disease', *Journal of Experimental Medicine*, 216(3), pp. 482–500. Available at: https://doi.org/10.1084/jem.20182031.

Troskie, R.L., Faulkner, G.J. and Cheetham, S.W. (2021) 'Processed pseudogenes: A substrate for evolutionary innovation: Retrotransposition contributes to genome evolution by propagating pseudogene sequences with rich regulatory potential throughout the genome', *BioEssays*, 43(11), pp. 1–11. Available at: https://doi.org/10.1002/bies.202100186.

Turrini, F. *et al.* (2015) 'HIV-1 transcriptional silencing caused by TRIM22 inhibition of Sp1 binding to the viral promoter', *Retrovirology*, 12(1), pp. 1–8. Available at: https://doi.org/10.1186/s12977-015-0230-0.

Turvey, S.E. and Broide, D.H. (2010) 'Innate immunity', *Journal of Allergy and Clinical Immunology*, 125(2), pp. S24–S32. Available at: https://doi.org/10.1016/j.jaci.2009.07.016.

Umthong, S. *et al.* (2021) 'Elucidating the antiviral mechanism of different march factors', *mBio*, 12(2), pp. 1–22. Available at: https://doi.org/10.1128/mBio.03264-20.

Uriu, K. *et al.* (2021) 'The battle between retroviruses and APOBEC3 genes: Its past and present', *Viruses*, 13(1), pp. 1–11. Available at: https://doi.org/10.3390/v13010124.

Usami, Y., Wu, Y. and Göttlinger, H.G. (2015) 'SERINC3 and SERINC5 restrict HIV-1 infectivity and are counteracted by Nef', *Nature*, 526(7572), pp. 218–223. Available at: https://doi.org/10.1038/nature15400.

Vigerust, D.J., Egan, B.S. and Shepherd, V.L. (2005) 'HIV-1 Nef mediates post-translational down-regulation and redistribution of the mannose receptor', *Journal of Leukocyte Biology*, 77(4), pp. 522–534. Available at: https://doi.org/10.1189/jlb.0804454.

Votteler, J. and Schubert, U. (2008) 'Human Immunodeficiency Viruses: Molecular Biology', *Encyclopedia of Virology*, (2005), pp. 517–525. Available at: https://doi.org/10.1016/B978-012374410-4.00428-3.

Waight, E. *et al.* (2022) 'Animal models for studies of HIV-1 brain reservoirs', *Journal of Leukocyte Biology*, 112(5), pp. 1285–1295. Available at: https://doi.org/10.1002/JLB.5VMR0322-161R.

Weichseldorfer, M. *et al.* (2020) 'Use of Humanized Mouse Models for Studying HIV-1 Infection, Pathogenesis and Persistence', *Journal of AIDS and HIV Treatment*, 2(1), pp. 23–29. Available at: https://doi.org/10.33696/aids.2.003.

Wein, T. and Sorek, R. (2022) 'Bacterial origins of human cell-autonomous innate immune mechanisms', *Nature Reviews Immunology*, 22(10), pp. 629–638. Available at: https://doi.org/10.1038/s41577-022-00705-4.

Weiss, R.A. and Vogt, P.K. (2011) '100 Years of Rous Sarcoma Virus', *Journal of Experimental Medicine*, 208(12), pp. 2351–2355. Available at: https://doi.org/10.1084/jem.20112160.

White, T.E. *et al.* (2013) 'Contribution of SAM and HD domains to retroviral restriction mediated by human SAMHD1', *Virology*, 436(1), pp. 81–90. Available at: https://doi.org/10.1016/j.virol.2012.10.029.

Willems, L. *et al.* (2000) 'Genetic determinants of bovine leukemia virus pathogenesis', *AIDS Research and Human Retroviruses*, 16(16), pp. 1787–1795. Available at: https://doi.org/10.1089/08892220050193326.

Wilson, S.J. *et al.* (2012) 'Inhibition of HIV-1 particle assembly by 2′,3′-cyclic-nucleotide 3′-Phosphodiesterase', *Cell Host and Microbe*, 12(4), pp. 585–597. Available at: https://doi.org/10.1016/j.chom.2012.08.012.

Wolf, D. and Goff, S.P. (2007) 'TRIM28 Mediates Primer Binding Site-Targeted Silencing of Murine Leukemia Virus in Embryonic Cells', *Cell*, 131(1), pp. 46–57. Available at: https://doi.org/10.1016/j.cell.2007.07.026.

Wong, M.E., Jaworowski, A. and Hearps, A.C. (2019) 'The HIV reservoir in monocytes and macrophages', *Frontiers in Immunology*, 10(JUN). Available at: https://doi.org/10.3389/fimmu.2019.01435.

Wu, T., Yan, Y. and Kozak, C.A. (2005) ' Rmcf2 , a Xenotropic Provirus in the Asian Mouse Species Mus castaneus , Blocks Infection by Polytropic Mouse Gammaretroviruses ', *Journal of Virology*, 79(15), pp. 9677–9684. Available at: https://doi.org/10.1128/jvi.79.15.9677-9684.2005.

Xue, G. *et al.* (2023) 'The HIV-1 capsid core is an opportunistic nuclear import receptor', *Nature Communications*, 14(1), pp. 1–16. Available at: https://doi.org/10.1038/s41467-023-39146-5.

Yánez, D.C., Ross, S. and Crompton, T. (2020) 'The IFITM protein family in adaptive immunity', *Immunology*, 159(4), pp. 365–372. Available at: https://doi.org/10.1111/imm.13163.

Yang, L. *et al.* (2020) 'Retrocopying expands the functional repertoire of APOBEC3 antiviral proteins in primates', *eLife*, 9, pp. 1–18. Available at: https://doi.org/10.7554/eLife.58436.

Yap, M.W. *et al.* (2014) 'Evolution of the Retroviral Restriction Gene Fv1: Inhibition of Non-MLV Retroviruses', *PLoS Pathogens*, 10(3). Available at: https://doi.org/10.1371/journal.ppat.1003968.

Yu, J. *et al.* (2015) 'IFITM Proteins Restrict HIV-1 Infection by Antagonizing the Envelope Glycoprotein', *Cell Reports*, 13(1), pp. 145–156. Available at: https://doi.org/10.1016/j.celrep.2015.08.055.

Yuan, P. *et al.* (2021) 'Trim28 acts as restriction factor of prototype foamy virus replication by modulating H3K9me3 marks and destabilizing the viral transactivator Tas', *Retrovirology*, 18(1), pp. 1–18. Available at: https://doi.org/10.1186/s12977-021-00584-y.

Yurkovetskiy, L. *et al.* (2018) 'Primate immunodeficiency virus proteins Vpx and Vpr counteract transcriptional repression of proviruses by the HUSH complex', *Nature Microbiology*, 3(12), pp. 1354–1361. Available at: https://doi.org/10.1038/s41564-018-0256-x.

Zhang, J., Scadden, D.T. and Crumpacker, C.S. (2007) 'Primitive hematopoietic cells resist HIV-1 infection via p21 Waf1/Cip1/Sdi1', *Journal of Clinical Investigation*, 117(2), pp. 473–481. Available at: https://doi.org/10.1172/JCI28971.

Zhang, R. *et al.* (2021) 'When human guanylate-binding proteins meet viral infections', *Journal of Biomedical Science*, 28(1), pp. 1–7. Available at: https://doi.org/10.1186/s12929-021-00716-8.

Zhang, Y. *et al.* (2020) 'MARCH8 inhibits viral infection by two different mechanisms', *eLife*, 9, pp. 1–14. Available at: https://doi.org/10.7554/ELIFE.57763.

Zhang, Y., Lu, J. and Liu, X. (2018) 'MARCH2 is upregulated in HIV-1 infection and inhibits HIV-1 production through envelope protein translocation or degradation', *Virology*, 518(February), pp. 293–300. Available at: https://doi.org/10.1016/j.virol.2018.02.003.

Zhang, Z. *et al.* (2012) 'Evolutionary Dynamics of the Interferon-Induced Transmembrane Gene Family in Vertebrates', *PLoS ONE*, 7(11), pp. 1–13. Available at: https://doi.org/10.1371/journal.pone.0049265.

Zhao, X. *et al.* (2019) 'IFITM genes, variants, and their roles in the control and pathogenesis of viral infections', *Frontiers in Microbiology*, 10(JAN), pp. 1–12. Available at: https://doi.org/10.3389/fmicb.2018.03228.

Zhu, Y. *et al.* (2011) 'Zinc-finger antiviral protein inhibits HIV-1 infection by selectively targeting multiply spliced viral mRNAs for degradation', *Proceedings of the National Academy of Sciences of the United States of America*, 108(38), pp. 15834–15839. Available at: https://doi.org/10.1073/pnas.1101676108.

Zila, V. *et al.* (2021) 'Cone-shaped HIV-1 capsids are transported through intact nuclear pores', *Cell*, 184(4), pp. 1032-1046.e18. Available at: https://doi.org/10.1016/j.cell.2021.01.025.

Appendix A: GBP Review

The review article in this Appendix provides important information concerning GBPs and shows them in the context of their current literature.

Review R2)

Functional cross-species conservation of guanylate-binding proteins in innate immunity

Luca Schelle[1], João Vasco Côrte-Real[1,2,3,4], Pedro José Esteves[2,3,4,5] Joana Abrantes[2,3,4], Hanna-Mari Baldauf[1]

[1]Max von Pettenkofer Institute and Gene Center, Virology, National Reference Center for Retroviruses, Faculty of Medicine, LMU München, Munich, Germany

[2]CIBIO-InBIO, Research Center in Biodiversity and Genetic Resources, University of Porto, 4485-661 Vairão, Portugal

[3]Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, 4099-002 Porto, Portugal

[4]BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, 4485-661, Vairão, Portugal

[5]CITS - Center of Investigation in Health Technologies, CESPU, 4585-116 Gandra, Portugal

Luca Schelle and João Vasco Côrte-Real contributed equally.

This review article was published 2023 in Medical Microbiology and Immunology and is available under: doi: 10.1007/s00430-022-00736-7

Medical Microbiology and Immunology (2023) 212:141–152
https://doi.org/10.1007/s00430-022-00736-7

REVIEW

Functional cross-species conservation of guanylate-binding proteins in innate immunity

Luca Schelle[1] · João Vasco Côrte-Real[1,2,3,4] · Pedro José Esteves[2,3,4,5] · Joana Abrantes[2,3,4] · Hanna-Mari Baldauf[1]

Received: 16 February 2022 / Accepted: 25 March 2022 / Published online: 13 April 2022

Abstract

Guanylate binding proteins (GBPs) represent an evolutionary ancient protein family widely distributed among eukaryotes. They are interferon (IFN)-inducible guanosine triphosphatases that belong to the dynamin superfamily. GBPs are known to have a major role in the cell-autonomous innate immune response against bacterial, parasitic and viral infections and are also involved in inflammasome activation. Evolutionary studies depicted that *GBPs* present a pattern of gain and loss of genes in each family with several genes pseudogenized and some genes more divergent, indicative for the birth-and-death evolution process. Most species harbor large *GBP* gene clusters encoding multiple paralogs. Previous functional studies mainly focused on mouse and human GBPs, but more data are becoming available, broadening the understanding of this multifunctional protein family. In this review, we will provide new insights and give a broad overview about GBP evolution, conservation and their roles in all studied species, including plants, invertebrates and vertebrates, revealing how far the described features of GBPs can be transferred to other species.

Keywords Guanylate binding protein · Evolution · Innate immunity · Antiviral proteins · Cross-species conservation · Plants · Invertebrates · Mammals

Introduction

GBPs are members of the dynamin superfamily (protein family) and the IFN-inducible guanosine triphosphatases. Of note, IFN inducibility is not true for GBPs in plants [1, 2] (Fig. 1a). The GBP proteins share common features and functions as outlined below:

Structure

The information on GBPs' structure is scarce. Indeed, until now, out of seven human GBP paralogs (hGBP1-7) only structural data for human GBP1 (hGBP1) exist [3], which has been recently extended to hGBP2/5 [4]. GBPs comprise three main domains: the large GTPase (LG) domain at the N-terminus connected by a hinge region (N-terminal part in α6 and C-terminal part in α7) to the middle domain (MD) and the GTPase effector domain (GED) at the C-terminus (Fig. 1b). The LG domain is a globular domain including five motifs: P-loop (G1), switch I (G2), switch II (G3), (N/T) KxD motif (G4) and the guanine cap (G5). These motifs are involved in GTP binding/orientation, Mg^{2+} cofactor finding and GTP/GDP hydrolysis [2, 4, 5]. The MD is an α-helical

Communicated by Asisa Volz.

Luca Schelle and João Vasco Côrte-Real contributed equally.

This article is part of the Special Issue on Immunobiology of Viral Infections.

✉ Hanna-Mari Baldauf
baldauf@mvp.lmu.de

1 Faculty of Medicine, Max Von Pettenkofer Institute and Gene Center, Virology, National Reference Center for Retroviruses, LMU München, Feodor-Lynen-Str. 23, 81377 Munich, Germany

2 CIBIO-InBIO, Research Center in Biodiversity and Genetic Resources, University of Porto, 4485-661 Vairão, Portugal

3 Department of Biology, Faculty of Sciences, University of Porto, 4169-007 Porto, Portugal

4 BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, 4485-661 Vairão, Portugal

5 CITS-Center of Investigation in Health Technologies, CESPU, 4585-116 Gandra, Portugal

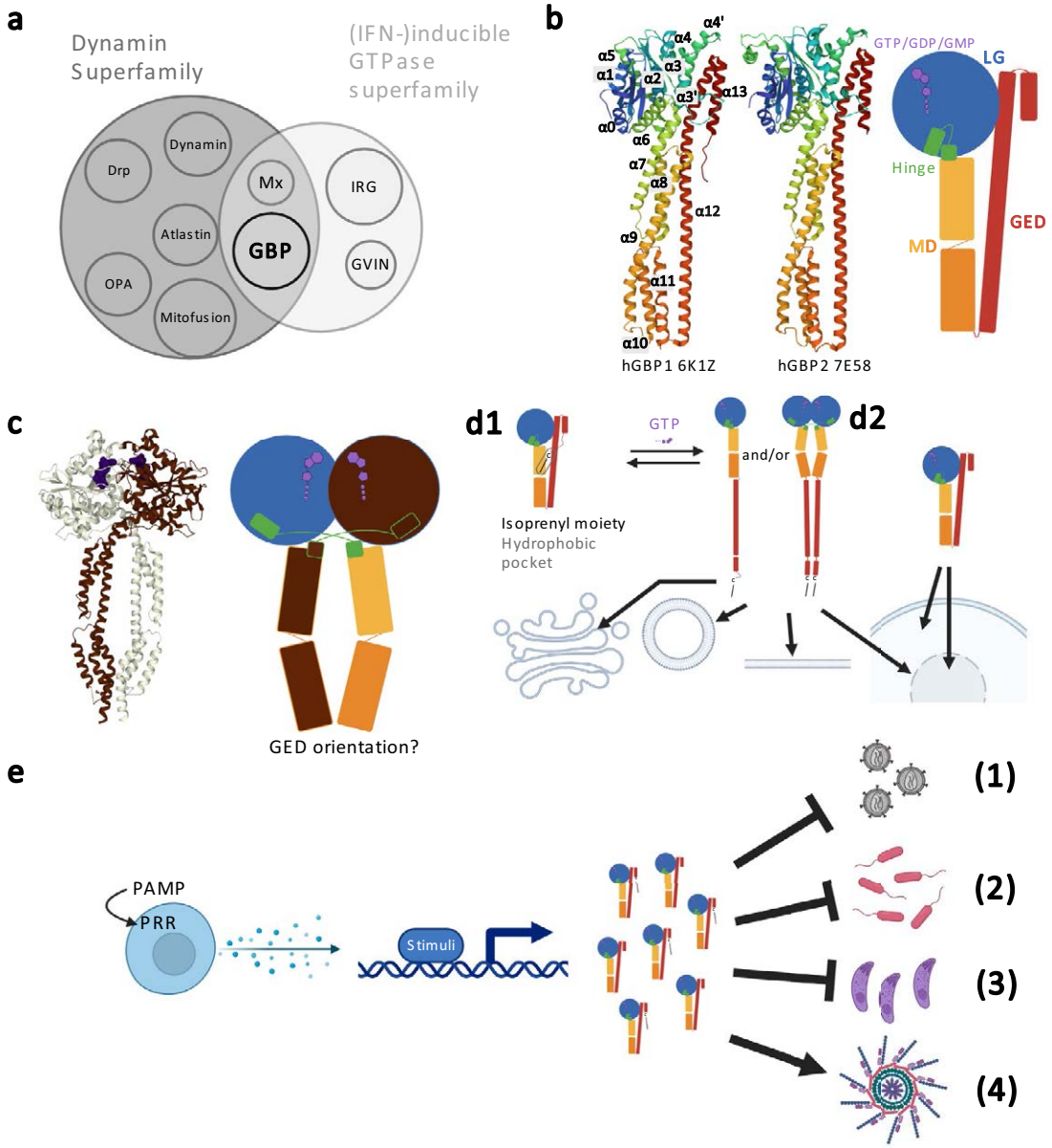

Fig. 1 Structure and function of GBPs. **a** Depicted is the relationship of GBPs within the dynamin superfamily proteins (protein family) and (IFN)-inducible Guanosine Triphosphatases. Since this functional classification is not true for plant GBPs, IFN is put in brackets [1, 2]. **b** Depicted is the structure of GBPs: GBP comprises three main domains: the N-terminal large GTPase (LG) domain connected by a hinge region to the middle domain (MD) and the GTPase effector domain (GED) C-terminal (PDB accession numbers: 6K1Z, 7E58), the α helices are labeled [4, 10, 75, 76]. **c** The proposed model of GBP dimerization (PDB accession number: 7E5A) is given [4, 75, 76]. **d** Depicted is the proposed model for GBP localization: upon GTP/dimerization, GBPs isoprenylated at their CaaX motif anchor to membranes via released isoprenyl moiety and open state; and localize therefore to different cellular organelles (vesicle-like, plasma membrane, perinuclear membrane, Golgi) (D1). GBPs without isoprenylation motif or in a closed state homogeneously localize in the cytosol and few in the nucleus (D2). **e** Depicted are the proposed functions of GBP: GBPs are part of the cell-autonomous innate immune response against various bacterial, parasitic and viral infections and involved in inflammasome activation. (1) viruses; (2) bacteria; (3) parasites; (4) inflammasome activation. Figure was created with BioRender.com

elongated domain (α7-11) comprising two three-helix bundles (α9 is shared). The GED is an α-helical elongated domain (α12-13) which, in nucleotide free state, is folded onto LG and MD.

Dimerization and polymerization

Recently, it has been described that hGBPs probably share a conserved dimerization mode between paralogs [4].

Upon GTP binding and LG:LG interface building, the GBP structurally rearranges to an open state driven by GTPase hydrolysis cycles. Kinetically delayed, the MD domain rearranges beneath the LG domain of the second GBP. Hereby, the hinge regions cross each other and form a closed dimeric state, which is further stabilized by the MD interface (Fig. 1c) [4, 6–9]. For hGBP1, farnesylation and GTP-dependent polymerization have been observed, but the exact function remains unclear [9].

Based on current knowledge it may be hypothesized that the conserved closed dimeric state represents the actual "active" form of specific GBPs' innate immunity-related functions but not all functions have to be solely related to dimerization [4].

Localization and membrane anchoring

hGBP1/2/5 harbor a CaaX motif at the C-terminus of the GED, which serves as a signal for in vivo isoprenylation (GBP1: farnesylated; GBP2/5: geranylgeranylated) and membrane anchoring. In a closed monomeric state hGBP1/2/5 localize homogenously distributed in the cytoplasm. Further, the isoprenyl moiety is buried in a hydrophobic pocket between GED (α12) and MD (α9) [2, 4, 5, 10]. Favored by the described GTP binding/hydrolysis and intra-dimeric interactions, the buried isoprenyl moiety is released from the hydrophobic pocket leading to a rearrangement into an open state. Subsequently, the released isoprenyl moiety is the determinant for membrane anchoring and, consequently, for the localization to the membranes at the cytosolic face of cellular compartments (Fig. 1d) (hGBP1: vesicle-like, plasma membrane; hGBP2: perinuclear membrane; hGBP5: Golgi). [4, 6, 8, 10]. Whereas the non-isoprenylated hGBP3/4 stay homogeneously localized in the cytosol or sometimes localized in the nucleus (Fig. 1d) [6, 11], hGBP4/6 can also be found to colocalize with vesicle-like structures without being isoprenylated [12]. It has also been described that homo- and heterodimerization influence localization [6, 11] but details are not yet clear.

GBP functions and roles in innate immunity

The expression of *GBPs* is triggered by inflammatory signals. The most potent stimuli for expression are interferons (IFN) due to IFN-stimulated response elements in the 5′ cis regulatory region of the *hGBP* genes. GBPs are among the most upregulated genes upon IFNγ stimulation. Especially hGBP1/5 expression is upregulated by up to two to three orders of magnitudes [13]. GBPs can be further stimulated by interleukins (ILs) and tumor necrosis factors (TNFs), but to a much lesser extent (reviewed in [13]).

The IFN-inducibility hints to some functions of GBPs. They are part of the cell-autonomous innate immune response against various pathogens and, in this context, are involved in canonical and non-canonical inflammasome activation. They respond to various intracellular bacteria, mostly gram negative, but also gram positive, as well as parasites (e.g., *Shigella flexneri*, *Salmonella enterica*, *Salmonella typhimurium*, *Legionella pneumophila*, *Francisella novicida*, *Chlamydia trachomatis*, *Listeria monocytogenes*, *Mycobacterium bovis*, *Leishmania donovani* and *Toxoplasma gondii*) (reviewed in [13, 14]). Moreover, GBPs inhibit viral infections such as vesicular stomatitis virus (VSV), classical swine fever virus (CSFV), murine norovirus-1 (MNV-1), Newcastle disease virus (NDV), encephalomyocarditis virus (EMCV), dengue virus (DENV), herpes simplex virus type 1 (HSV-1), Kaposi's sarcoma-associated herpesvirus (KSHV), hepatitis E virus (HEV), hepatitis C virus (HCV), influenza A virus (IAV), human immunodeficiency virus (HIV) and respiratory syncytial virus (RSV) (Fig. 1e) (reviewed in [14, 15]).

Taken together, GBPs have been considered as major players in the host innate immunity by providing defense against a broad range of invading pathogens.

GBP evolution and conservation

The origin and evolution of *GBPs* have been analyzed only recently with most of the evolutionary history of *GBPs* still unclear [16–19]. *GBPs* originated from a common ancestor and belong to the multigene family of the large dynamin superfamily [20]. *GBPs* can be found in a broad range of organisms from plants to humans [18]. The presence of *GBPs* in plants species like *Arabidopsis thaliana*, *Oryza sativa* and *Solanum lycopersicum* indicates that *GBPs* are active in organisms that do not present migratory immune cells and an IFN-inducible immune system [18].

In mammals, *GBP* genes are usually organized in tandem on the same chromosome [19, 20]; however, in some rodents, like *Mus* and *Rattus norvegicus*, the *Gbps* are located on two gene clusters on different chromosomes [16]. In addition, in zebrafish and frogs, *gbp* genes are found in three small genomic islands [13]. Plants also have a variation regarding the number of *GBPL* (GBP-like) genes present in their genome, for example, *Oryza sativa* has three orthologs, while in *Arabidopsis thaliana* and *Zea mays* seven *GBPL* are encoded in their genome [18]. Altogether, this suggests that independent duplication events contributed to *GBP* diversity across plant and animal kingdoms [18]. Moreover, since *GBPs* are a multigene family that belongs to the immune system, it follows the birth-and-death process of evolution [21]. This results in some genes being either deleted or maintained in the genome. When maintained, the genes can acquire a new function (neofunctionalization), split functions (subfunctionalization) or even lose

function and become pseudogenes [17, 22]. For example, *GBP3* gene appears to have emerged only in Simiiformes through a duplication of *GBP1* and gained a new function being responsible for the regulation of caspase-4 activation (Table 1) [23]. As for *GBP7*, it most likely emerged from a duplication event of *GBP4* and seems to be only present in primates (Table 1) [17].

GBP4 and *GBP5* seem to have been deleted from the genomes of Old-World monkeys and the lack of *GBP5* orthologs might explain the HIV-2 transmission susceptibility in these primates since *GBP5* inhibits HIV-2 infection [13, 17].

Some *GBP* orthologs are not present in different species, while others might be exclusive to specific orders. According to phylogenetic analyses it appears that primate *GBP1, GBP3* and *GBP7* are absent from muroid genomes (Table 1) [16]. This further indicates that the nomenclature of muroids *Gbps* has been incorrect and functional studies of these *Gbps* might have led to misleading results [16]. Following an evolutionary study in muroids, *Gbp2*, *Gbp5* and *Gbp6* have been found to be orthologs to their primate counterparts [16].

Gbp2 is found in every family of muroids and duplication events occurred in all genera except in *Rattus*. *Gbp5* presents only one copy in each species of muroids, similar to primates. Maintenance of *Gbp2* and *Gbp5* in the muroid genomes supports the importance of these two genes for the host immune system [24, 25].

Phylogenetic analyses in Muroidea and Cricetidae indicate the presence of four *Gbps* that are exclusive to these taxa (*Gbpa, b, c* and *d*) (Table 1) [16]. The *Gbpa* and *Gbpb* groups are mainly composed of *Gbps* previously classified in public databases (NCBI and Ensembl) as *Gbp1* [16].

Table 1 General overview of GBP genes in Primates and Muroids

	Primates		Muroids		
	New world monkeys and great apes	Old world monkeys	*Muridae*	*Cricetidae*	*Mus musculus*
GBP1	+	+	−	−	−
GBP2	+	+	+	+	+
GBP3	+/ψ	+	−	−	−
GBP4	+	−	−	−	−
GBP5	+	−	+	+	+
GBP6	+	+	+	+	+
GBP7	+/φ	+	−	−	−
Gbpa	−	−	+/ω	+/ω	−
Gbpb	−	−	+/ω	+/ω	−
Gbpc	−	−	+/ω	+/ω	−
Gbpd	−	−	+/ω	+/ω	+

+, present; −, not present; ψ, exclusive to Simiiformes; φ, exclusive to Primates; ω, exclusive to Muroids

Phylogenetically, they are not similar to *hGBP1*. Interestingly, these genes are not present in *Mus musculus* [16]. The function of these genes has yet to be determined, but the study of the sequences and the 3D structure of the proteins may provide hints on their function. *Gbpc* is only present in three species, being absent in *Mus musculus*, but its function is also not known. Considering the *Gbpd* group, three main groups emerged and are present in all species of muroids indicating a possible duplication in the common ancestor of Muridae and Cricetidae (Table 1) [16]. The *Mus musculus* classified as *Gbpd1* [16], previously annotated in NCBI and Ensembl as *Gbp7,* appears to be a cellular host dependency factor for IAV replication [26].

Gbp6 cluster is present in most Muridae and Cricetidae species, and in *Mus musculus* and *M. caroli*, an expansion of this gene has observed, with *Mus musculus* presenting six copies and *Mus caroli* four. This expansion might be explained as a compensation mechanism due to the lack of *Gbpa*, *b* and *c* in these two species [16].

The evolutionary history of the *GBP* multigene family is complex and dynamic with duplication (*Gbp2* and *Gbp6* in several species), deletion (*Gbpa*, *b* and *c* in *Mus musculus*; Table 1) and neofunctionalization (*GBP3* in primates) of genes, in line with the proposed birth-and-death mode of evolution [17]. In each mammalian family, the different evolutionary histories open new research opportunities to study the evolution and function of *GBPs*, which should be conducted in a more holistic approach.

GBP functions in plants, invertebrates and vertebrates

GBPs in plants

GBP-like proteins seem to be widely distributed as they even exist in plants. Plants solely rely on innate immune mechanisms to resist against phytopathogens (reviewed in [27, 28]). GBPs are poorly characterized in plants, but first results have been obtained in recent years. Indeed, tomato (*Solanum lycopersicum*) GBP homolog, SIGBP1, has been reported to be involved in fruit tissue differentiation by maintaining cells in a non-proliferative state [29] (Fig. 2A). First comparisons of the modeled structure of *Arabidopsis* GBP-like (AtGBPL) to hGBP1 crystal structure revealed a similar architecture. AtGBPL1/3 seem to comprise an intrinsically disordered region (IDR) at the C-terminus instead of an isoprenylation motif [18]. Functional studies with AtGBPL1/2/3 have revealed the roles of AtGBP1 (negative allosteric regulator of AtGBP3) and AtGBP3 in host defense. Indeed, they confer resistance to phytopatogens such as *Pseudomonas syringae* pv. *maculicola* (Psm), *Pseudomonas syringae* pv. Tomato (Pst) and *Hyaloperonospora*

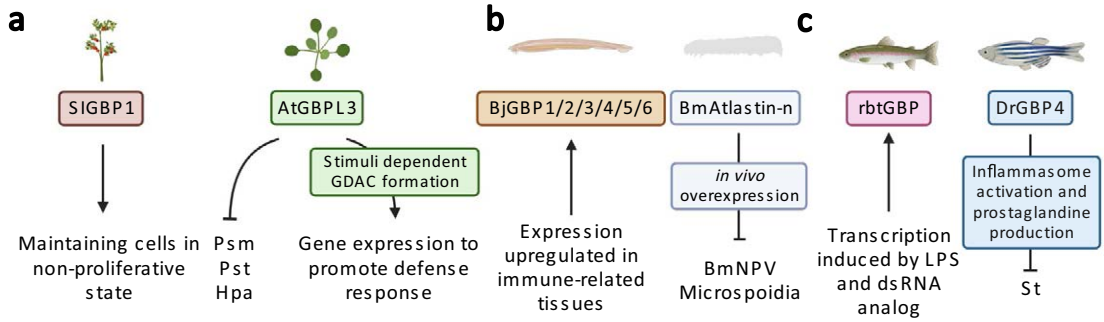

Fig. 2 Plant, invertebrate and teleost GBPs in innate immunity. **a** Plants: AtGBPL3 confers plant defense against Psm, Pst and Hpa. Further, stimuli-dependent formation of GDACs reprogram host gene expression to promote defense response. SIGBP1 maintains cells in a non-proliferative state. **b** Invertebrates: BjGBPs expression is upregulated in immune-related tissues. BmAtlastin-n inhibits in vitro and in vivo replication of BmNPV and microsporidia. **c** Teleosts: The transcription of rbtGBP is induced by LPS and dsRNA analogs. DrGBP4 supports clearance of *St* infections via inflammasome activation and prostaglandine production. Abbreviations: *Arabidopsis thaliana* (At), *Pseudomonas syringae* pv. *maculicola* (Psm), *Pseudomonas syringae* pv. Tomato (Pst), *Hyaloperonospora arabidopsidis* (Hpa), GBPL defense-activated condensates (GDAC), *Solanum lycopersicum* (Sl), *Branchiostoma japonicum* (Bj), *Bombyx mori* (Bm), double-stranded (ds), nucleopolyhedrovirus (NPV), rainbow trout (rbt), lipopolysaccharide (LPS), *Danio rerio* (Dr), *Salmonella typhimurium* (St), mouse (m), *Listeria monocytogenes* (Lm), *Mycobacterium bovis* (Mb), vesicular stomatitis virus (VSV), encephalomyocarditis virus (EMCV). Figure was created with BioRender.com

arabidopsidis (Hpa). Upon salicylic acid, pipecolic acid or phytopathogen activation, AtGBP3 condensates to unique membraneless organelles, termed GBPL defense-activated condensates (GDACs), within the nucleus, binding defense-gene promotors and recruiting transcriptional coactivators. This, in turn, reprograms the host gene expression to promote host defense responses (Fig. 2a). GDACs have also been observed in tomato and maize, which could hint for a conserved mechanism in plants [15]. Since phytohormone salicylic acid biosynthesis is also promoted by plant viruses (reviewed in [30]), it seems possible that AtGBPLs also might be involved in antiviral response, but this hypothesis needs to be proven.

In summary, GBP-dependent innate immunity processes are present in plants and animals and, thus, probably exist already over a longer period of time.

GBPs in invertebrates

The function/presence of GBPs in invertebrates is still unclear. Indeed, in silico analyses have revealed that non-vertebrate species harbor *GBP*-like genes, but not all of them seem to be completely lacking them [13, 20]. If this is due to a low genome coverage or, in fact, if these genes are not present still needs further clarification. In amphioxus (*Branchiostoma japonicum*), expression of *GBPs* is upregulated in immune-related tissues [20] (Fig. 2b), which could indicate their involvement in innate host defense.

Recently, the BmAtlastin-n protein of silkworm (*Bombyx mori*) has been suggested to be part of the GBP family [31] due to the lack of the typical atlasin transmembrane domain [32] and similarity in the GTPase domain [31]. Transgenic silk worms overexpressing BmAtlastin-n have shown in vitro and in vivo inhibition of viral reproduction capacity of *Bombyx mori* nucleopolyhedrovirus (BmNPV), a virus causing nuclear polyhedrosis [32]. The mechanism of viral inhibition is elusive, but it seems to correlate with the reduction of VP39 (capsid protein from late baculovirus gene) expression levels (mRNA and protein) [32]. Furthermore, it also enhances in vivo resistance against the obligate intracellular parasite microsporidia. Therefore, BmAtlastin-n seems to protect from intracellular infections caused by more than one pathogen (Fig. 2b), similar to other GBPs.

Why some invertebrates harbor GBPs in their genome and others seem to have lost them remains an open question requiring further investigations. Since atlastins and GBPs are closely related, it raises the question if in invertebrates without GBP homologs atlastins may have adopted some of their defense functions or if their common ancestor already possessed anti-pathogenic functions.

Gbps in teleosts

Studies regarding Gbps in teleosts are scarce. The first characterization of Gbps in fish has been in 2006 by Robertsen and colleagues [33], while mammalian GBPs have been described since 1983 [34]. The Gbp found in rainbow trouts (rbtGBP) appears to have a similar structure as hGBP1 with similar domains and a CaaX motif at the C-terminus, responsible for isoprenylation. Moreover, the most conserved region is the N-terminal surrounding GTP-binding region (amino acid 6–278) [33], while the C-terminal region is 43 amino acids longer compared to the human counterpart [33]. rbtGbp shares 41 to 47% amino acid sequence identity with

mammalian GBPs. Interestingly, the region encompassing the GTP-binding motifs shares 67% identity with mammals. However, the C-terminus has only 37% identity with the mammalian GBPs [33]. The transcription level of rbtGBP is upregulated by lipopolysaccharide (LPS) and polyinosinic polycytidylic acid (poly I:C, double-stranded RNA analog) [43] (Fig. 2c). This may hint for an involvement of rbtGBP in innate immunity against bacteria and RNA viruses. Zebrafish Gbp is similar in length to the rbtGbp, but lacks a CaaX motif at the end of the C-terminus [33]. Nevertheless, DrGbp may play a role in the innate immunity against bacterial infections since DrGbp4 is involved in inflammasome activation and clearance of *Salmonella typhimurium* (St) infections [35].

In *Danio rerio*, eight Gbps have been found, with two Gbps being studied until now, Gbp1 and Gbp4. The nomenclature of *gbps* in fish is probably inaccurate since they do not cluster with their human counterparts, similar to the observations in muroids [26]. DrGbp1 contains an N-terminal GTPase domain and a helical C-terminal domain similar to mammalian GBPs [36]. DrGbp4 has a similar architecture as DrGBP1 with an additional C-terminal caspase recognition domain (CARD) and shares 53% identity with hGBP5 [35]. DrGbp4 is an IFNγ-induced GTPase, similar to mammalian GBPs. It is expressed in neutrophils, but in macrophages expression levels were hardly detected [35]. Tyrkalska and colleagues have demonstrated the paramount role of Gbp4 in bacterial clearance, being crucial for the biosynthesis of prostaglandins via an inflammasome-dependent pathway to clear *St* bacterial infection [35]. The GTPase activity of Gbp4 is crucial for caspase-1 activity, inflammasome activation and resistance to infection by *St* bacterial infection [35]. Indeed, Gbp4-deficient fish have a negatively affected caspase-1 activity and display increased susceptibility to *St* infections compared to fish with wildtype Gbp4. Interestingly, when Gbp4-deficient fish are trans-complemented with mouse *Gbp5*, *St* susceptibility decreases and caspase-1 defects are rescued [35]. Additionally, DrGbp4 regulates the expression of WD repeat domain 90 (WDR90), which is a component of the NOD-like receptor with CARD domain 4 inflammasome and is responsible for the conformational change needed for its activation [37] (Fig. 2c). Altogether, in fish, Gbps appear to have also an important role in the innate immune system, especially for bacterial infection. However, more studies are needed to further understand the functions of Gbps in teleost.

GBPs in mammals

Several studies have already been performed to understand the functions of GBPs in humans and, at some extent, in rodents and few further mammals; however, in general, the function of the majority of the mammalian GBPs remains unclear.

Since we would like to emphasize in this review the roles of non-human GBPs, we only shortly point out the antiviral activity of hGBPs. Needless to say their activity against bacteria and parasites are not less important, they have been recently reviewed in detail in [13, 14]. hGBP1/2/3/5 are known to be involved in restriction of viruses, employing thereby various mechanisms and targeting different steps in their life cycle. Yet, the underlying mechanisms remain elusive for specific viruses [14, 15]. hGBP1 employs several mechanisms to restrict viruses (Fig. 3a). For KSHV, the transport of the viral capsid to the nucleus is hampered by disruption of the actin filaments by hGBP1 [38]. HEV is inhibited through the relocation of the capsid protein by hGBP1 to the lysosome [39]. For HCV, the observed interaction with RNA-dependent RNA polymerase NS5B could be a possible explanation for the viral restriction [40]. In the case of IAV, the NS1 virulence factor is antagonized by hGBP1 [41]. For other viruses (e.g., VSV, DENV) the mode of action for their inhibition by GBP1 remains unknown [42, 43]. hGBP1 may employ similar mechanisms as mentioned above to inhibit the other viruses but also other mechanisms are conceivable. hGBP3 has only now been identified to play a role in IAV infection by inhibiting the viral polymerase complex [44] (Fig. 3b). GBP2/5 interfere with the host protease furin, which impairs HIV glycoprotein maturation resulting in a decreased infectivity of released viral particles [12, 45]. This has been also observed for Zika virus (ZIKV), measles virus (MEV) and lentiviral particles pseudotyped with various envelope glycoproteins (avian IAV, murine leukemia virus (MLV), Marburg virus (MARV) and human endogenous retrovirus K (HERV-K)) [12, 45, 46]. GBP5 further restricts the replication of RSV by reducing intracellular levels of the viral small hydrophobic protein [47]. Thus, GBP5 is generally involved in innate immunity as it can induce enhanced production of IFN and proinflammatory signals [48] (Fig. 3c).

Five pig (p) GBPs are described in literature. Based on NCBI, *Sus scrofa* has 7 *GBPs* in one gene cluster on chromosome 4 (accession numbers: NM_001128473.1, NM_001128474.1, XM_005663706.3, XM_021090310.1, XM_013997408.2, XM_021090315.1, XM_005663708.3). Only pGBP1/2 have been characterized on protein level. They share a conserved N-terminal GTPase domain and a C-terminal CaaX motive similar to other mammalian GBPs [49]. Pig GBP research is limited to pathogens especially affecting the global swine industry: the respiratory syndrome virus (PRRSV) and classical swine fever virus (CSFV) [50]. CSFV replication is potently inhibited by pGBP1 via its GTPase activity. pGBP1 mainly acts in the early phase of viral replication by inhibiting the translation efficiency of the internal ribosome entry site (IRES). Notably, CSFV NS5A

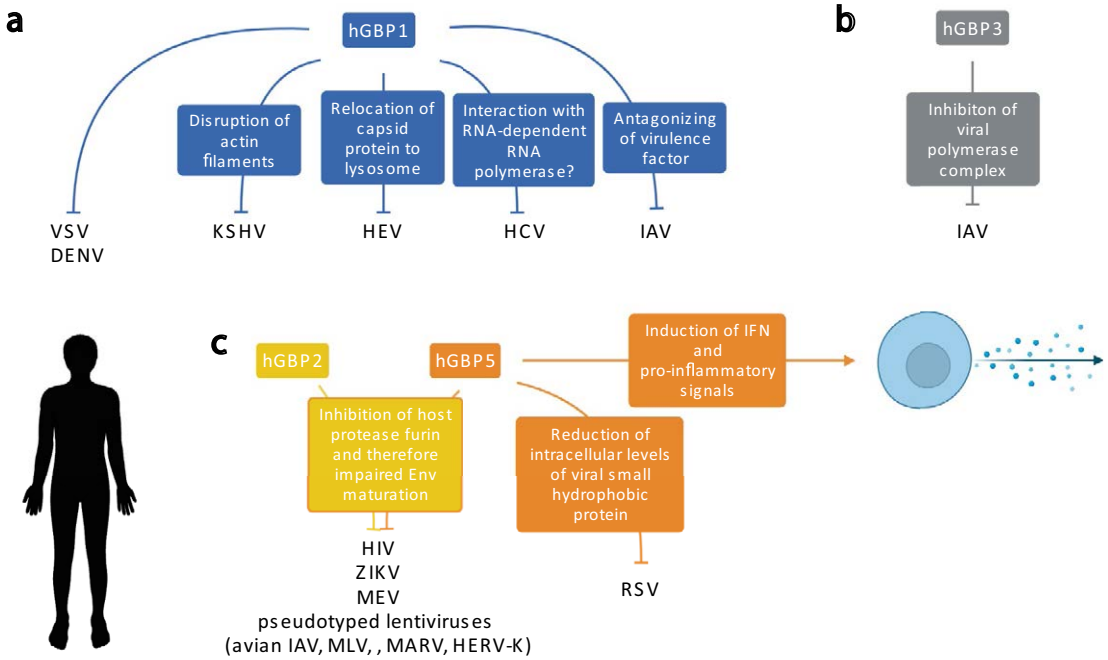

Fig. 3 Antiviral activity and underlying mechanisms of hGBPs. **a** hGBP1: Restriction of VSV and DENV by unknown mechanism. Restriction of KSHV by disruption of actin filaments, resulting in impaired transport. Restriction of HEV by relocation of the capsid protein to the lysosome. Interaction with RNA-dependent RNA polymerase (NS5B) of HCV might explain restriction. For IAV, the NS1 virulence factor is antagonized by hGBP1. **b** hGBP3: Inhibition of the viral polymerase complex of IAV. **c** hGBP2/5: Interference with host protease furin, which impairs HIV glycoprotein maturation, resulting in a decreased infectivity of released viral particles. Same has been described for Zika virus (ZIKV), measles virus (MEV) and lentiviral particles pseudotyped with various envelope glycoproteins (avian IAV, murine leukemia virus (MLV), Marburg virus (MARV) and human endogenous retrovirus K (HERV-K) Env glycoproteins. hGBP5 further restricts the replication of RSV by reducing intracellular levels of the viral small hydrophobic protein. hGBP5 has also a more general role in innate immunity as it can induce enhanced production of IFN and proinflammatory signals. Abbreviations: human (h), vesicular stomatitis virus (VSV), dengue virus (DENV), Kaposi's sarcoma-associated herpesvirus (KSHV), hepatitis E virus (HEV), hepatitis C virus (HCV), influenza A virus (IAV), human immunodeficiency virus (HIV), murine leukemia virus (MLV), Zika virus (ZIKV), measles virus (MEV), Marburg virus (MARV), human endogenous retrovirus K (HERV-K), respiratory syncytial virus (RSV), Interferon (IFN). Figure was created with BioRender.com

protein counteracts pGBP1's antiviral activity by inhibition of the GTPase activity [50]. For PRRSV, a quantitative trait locus (QTL) on *Sus scrofa* chromosome (SSC) 4 has been identified being beneficial for controlling infection. The characterization of this QTL revealed that it contains inter alia pGBP1/2/4/5/6 and that the QTL is associated with resistance to PRRSV infection. Furthermore, pGBP1/5/6 lead to a reduction of PRRSV viral loads in vivo in pigs [51–54]. Yet, the underlying mechanisms remain elusive.

Tupaia has 5 copies of *GBPs* in one gene cluster similar to humans, while most rodents present two gene clusters [19, 55]. Also similar to human and mouse GBPs, the coding region of *Tupaia GBPs* (*tGBPs*) ranges from 1733 to 1884 bp and the molecular weight of the proteins is between 67 to 72kD [55]. Most of the conserved motifs are present, particularly in the N-terminus where the GTPase domain is located. As expected, the C-terminus shares low sequence identity among the different groups. Phylogenetically, the sequences of *tGBP* genes are clustered with the h*GBP* genes, which indicates that the *Tupaia* genes are human orthologs [55]. Only in *tGBP1*, *tGBP2* and *tGBP5* a CaaX motif is present as in humans and mice [13, 56, 57]. This motif allows isoprenylation and consequently the anchorage to membranous organelles, enabling the destruction of pathogen-containing vacuoles, mainly bacterial pathogens, which exposes the pathogen to the host [15, 58–60].

When acute signaling is absent, hGBPs are expressed at low to medium levels in immune cells, lung, liver, kidney, brain and skin [13, 61]. tGBPs are also ubiquitously expressed at low levels in heart, spleen, kidneys, intestines, liver, lung and brain [55]. Human, mouse and *Tupaia GBPs* are strongly induced by IFN [19, 55, 62, 63] and *Tupaia* mRNA levels of *GBPs* are increased after RNA virus infections of primary renal cells such as Newcastle disease virus (NDV) and encephalomyocarditis virus (EMCV), and DNA virus type 1 herpes simplex virus (HSV-1) [55] (Fig. 4b).

As outlined above, hGBP1 is the most studied GBP, it has been described to have antiviral activity against a broad

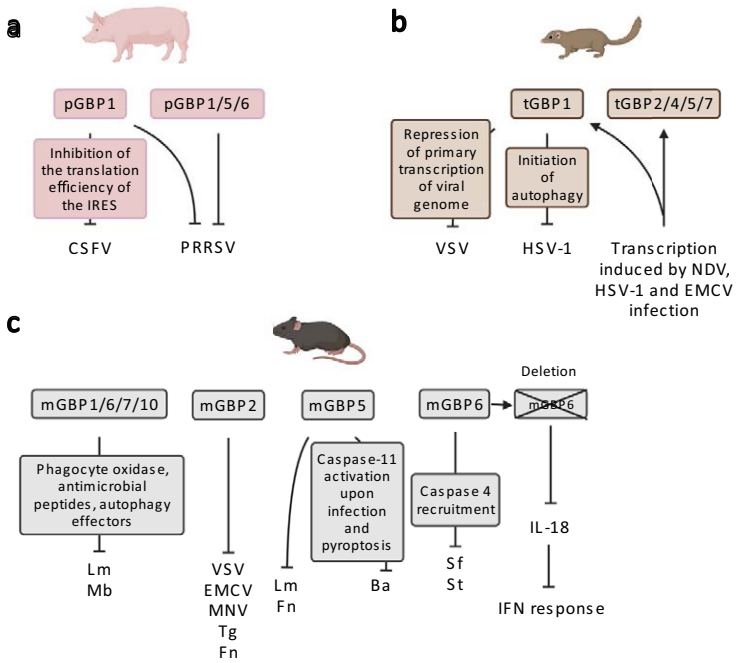

Fig. 4 Non-human mammalian GBPs in innate immunity. **a** Pig: pGBP1 inhibits CSFV by inhibition of the translation efficiency of the IRES. pGBP1/5/6 reduce viral loads of PRRSV in vivo in pigs via unknown mechanisms. **b** Tupaia: The transcription of tGBP1/2/4/5/7 is upregulated upon VSV, HSV-1 and NDV infection. tGBP1 restricts VSV by repression of primary transcription of the viral genome. tGBP1 further restricts HSV-1 via initiation of autophagy. **c** Mouse: mGBP1/6/7/10 restrict Lm and Mb combined via phagocyte oxidase, antimicrobial peptides, and autophagy effectors. mGBP2 displays restriction towards various viral, bacterial and parasitic pathogens (VSV, EMCV, MNV, Tg, Fn). mGBP5 restricts Lm and Fn. mGBP5 further inhibits Ba through Caspase-11 activation and pyroptosis. mGBP6 restricts Sf and *St* via Caspase-4 recruitment. Deletion of mGBP6 leads to reduced IFN response. Abbreviations: pig (p), respiratory syndrome virus (PRRSV), classical swine fever virus (CSFV), internal ribosome entry site (IRES), *Tupaia* (t), vesicular stomatitis virus (VSV), Newcastle disease virus (NDV), type 1 herpes simplex virus (HSV-1), encephalomyocarditis virus (EMCV), mouse (m) *Listeria monocytogenes* (Lm), *Mycobacterium bovis* (Mb), murine norovirus (MNV), *Toxoplasma gondii* (Tg), *Francisella novicida* (Fn), *Bacillus abortus* (Ba), *Shigella flexneri* (Sf), *Salmonella typhimurium* (St), interleukin (IL), interferon (IFN). Figure was created with BioRender.com

range of viruses [38, 40, 42, 64]. In *Tupaia*, tGBP1 is the only GBP from the five tGBPs that displays antiviral activity against VSV and HSV-1. It significantly represses the primary transcription of VSV viral genomes, but only presents a rather moderate effect against HSV-1 [55]. For VSV-G, tGBP1 restricts the viral genomic transcription in the cytoplasm by competitively binding to the VSV-N subunit [55]. The moderate HSV-1 inhibition by tGBP1 is tSTING-dependent, promoting tSTING-mediated autophagy, but the mechanism remains unclear. The authors speculated that autophagy could clear pathogens and DNA from the cytoplasm [65].

All tGBPs are upregulated through different viral infections, which suggests they may play a role in antiviral immunity (Fig. 4B). Yet, it is unclear how they inhibit viral replication, infectivity and proliferation [55]. The other four tGBPs need to be further investigated as *Tupaia* is becoming a recognized animal model to study human diseases (e.g., metabolic, brain aging, neurological, psychiatric and cancer) due to its closer relationship to humans than rodents [55] and also to its susceptibility to a wide range of human pathogens (HCV, HSV and SARS-CoV-2) [55, 64, 66, 67].

Murine GBP functions are the second most studied after human GBPs. As previously described, they are important for the host defense against pathogens and inflammasome activation. mGBP2 antiviral activity has been first described in 2005, revealing inhibition of VSV and EMCV replication [68]. EMCV replication inhibition requires GTPase activity of mGBP2, unlike the inhibition of VSV replication [68]. Murine norovirus (MNV) replication is inhibited when *mGBP2* is expressed in mouse macrophages. The N-terminus of mGBP2 is crucial for anti-MNV activity since only GBP2 mutants that express the G domain and the GM domain inhibit viral replication at RNA and protein level, M domain alone and the remaining domains did not present anti-MNV activities [69]. hGBP2 and hGBP5 have been described to exert a broad antiviral activity against Zika virus, measles, HIV-1 and influenza A virus by reducing their replication and also impairing furin-mediated processing of envelope glycoproteins leading to a decrease in

infectivity [12, 45]. Despite the phylogenetic analyses and the conserved function of GBPs, the antiviral functions of mGBP2 and mGBP5 are yet to be fully disclosed and further studies are needed.

Additional studies demonstrate that *mGbp2* knockout increases susceptibility to infections with *Toxoplasma gondii* and *Francisella novicida*; yet mGBP2 did protect against infections with *Listeria monocytogenes* [24, 70]. mGBP5 also provides host defense against bacterial infections such as *L. monocytogenes* and *F. novicida* [24, 25]. In mouse macrophages, mGBP5 mediates caspase-11 activation and pyroptosis upon *Bacillus abortus* infection; knockdown of *mGbp5* decreased IL-1β concentrations and, expectedly, bacterial count in macrophages is increased [71, 72].

For the newly classified *Gbp6*, previously designated *Gbp4* in *Mus musculus* [16], Wandel and colleagues demonstrated its importance in caspase-4 recruitment, with the depletion of *Gbp4* in cells leading to the inability of processing and releasing IL-18 during *Shigella flexneri* and *Salmonella typhimurium* infection [23], confirming that GBPs are crucial for inflammasome activation and bacterial clearance. Most studies have focused on the individual function of each mGBP; however, the combined function of GBPs is starting to be addressed. Indeed, *en bloc* knockout of *mGBPs* located on chromosome 3 leads to reduced release of IL-18 and IL-1β via canonical NLRP3 and AIM2 inflammasomes, which is needed for IFN-γ production and host defense against bacteria, ultimately increasing susceptibility of infection [24, 73]. Moreover, it has been described that mGBP1, mGBP6, mGBP7 and mGBP10 are paramount to hamper virulent strains of *L. monocytogenes* and *M. bovis* in mouse involving phagocyte oxidase, antimicrobial peptides and autophagy effectors [63]. Silencing m*Gbps* with siRNAs has indicated that the protective effects of mGBPs operate in a collaborative way, since the combination of siRNAs decreased the killing ability via IFN-γ [63] (Fig. 4c).

Curiously, the expression of all *Gbps* located on chromosome 3 have displayed a beneficial interaction which limited acute inflammatory bone loss since $Gbp^{Chr3-/-}$ mouse cells exhibit increased bone loss compared to wildtype [74].

Concluding remarks

GBPs exist in a variety of eukaryotic organisms ranging from plants to animal kingdoms. Despite playing an important role in the innate immunity, the evolutionary history of *GBPs* as a multigene family is not yet fully disclosed. The immune system is continuously challenged by a broad range of intracellular pathogens, which leads to a complex evolution of the innate immunity genes. In each family, the number of *GBPs* varies, presenting several events of duplication, pseudogenization and deletion. Human and mouse GBPs have been characterized in more detail, but mostly restricted to GBPs 1/2/5. Yet, even for those, many functions remain undetermined as GBPs seem to be involved in a complicated cellular network. In this review, we provide insights on the maintenance of GBPs basal functions, like resistance to pathogens (viral, bacterial and parasitic); however, the detailed mechanisms and networks among species have not yet been sufficiently characterized. Therefore, studies on GBPs including more species may be beneficial to further understand the complex GBP network and their functions. It will be also crucial to understand the differences within the *GBP* gene clusters even in closely related species.

Acknowledgements This work was supported by Fundação para a Ciência e Tecnologia (FCT) through the PhD fellowship of J.V.C.R. (DFA/BD/4965/2020), the Assistant Researcher grant of J.A. (CEECIND/00078/2017) and the Principal Researcher grant of P.J.E. (CEECIND/01495/2020). H-M.B. acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG) (BA-6820/1-1). H-M.B. and J.A. acknowledge the project-related personal exchange (PPP) program of the FCT/German Academic Exchange Service (DAAD) (57518622).

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest The authors report no conflict of interests.

References

1. MacMicking JD (2004) GTP-binding proteins: structure-function relationships. Trends Immunol 25(11):601–609. https://doi.org/10.1016/j.it.2004.08.010
2. Jimah JR, Hinshaw JE (2019) Structural insights into the mechanism of dynamin superfamily proteins. Trends Cell Biol 29(3):257–273. https://doi.org/10.1016/j.tcb.2018.11.003
3. Prakash B, Praefcke GJK, Renault L, Wittinghofer A, Herrmann C (2000) Structure of human guanylate-binding protein 1 representing a unique class of GTP-binding proteins. Nature 403(6769):567–571. https://doi.org/10.1038/35000617
4. Cui W, Braun E, Wang W, Tang J, Zheng Y, Slater B et al (2021) Structural basis for GTP-induced dimerization and antiviral function of guanylate-binding proteins. Proc Natl Acad Sci USA 118(15):1–9. https://doi.org/10.1073/pnas.2022269118

5. Wittinghofer A, Vetter IR (2011) Structure-function relationships of the G domain, a canonical switch motif. Annu Rev Biochem 80:943–971. https://doi.org/10.1146/annurev-biochem-062708-134043 (**Epub 2011/06/17**)
6. Britzen-Laurent N, Bauer M, Berton V, Fischer N, Syguda A, Reipschläger S et al (2010) Intracellular trafficking of guanylate-binding proteins is regulated by heterodimerization in a hierarchical manner. PLoS ONE 5(12):e14246. https://doi.org/10.1371/journal.pone.0014246
7. Shydlovskyi S, Zienert AY, Ince S, Dovengerds C, Hohendahl A, Dargazanli JM et al (2017) Nucleotide-dependent farnesyl switch orchestrates polymerization and membrane binding of human guanylate-binding protein. Proc Natl Acad Sci USA 114(28):E5559–E5568. https://doi.org/10.1073/pnas.1620959114
8. Sistemich L, Dimitrov Stanchev L, Kutsch M, Roux A, Günther Pomorski T, Herrmann C (2021) Structural requirements for membrane binding of human guanylate-binding protein 1. FEBS J 288(13):4098–4114. https://doi.org/10.1111/febs.15703
9. Sistemich L, Kutsch M, Hämisch B, Zhang P, Shydlovskyi S, Britzen-Laurent N et al (2020) The molecular mechanism of polymer formation of farnesylated human guanylate-binding protein 1. J Mol Biol 432(7):2164–2185. https://doi.org/10.1016/j.jmb.2020.02.009
10. Ji C, Du S, Li P, Zhu Q, Yang X, Long C et al (2019) Structural mechanism for guanylate-binding proteins (GBPs) targeting by the Shigella E3 ligase IpaH9.8. Plos Pathog 15(6):e1007876. https://doi.org/10.1371/journal.ppat.1007876
11. Tripal P, Bauer M, Naschberger E, Mörtinger T, Hohenadl C, Cornali E et al (2007) Unique features of different members of the human guanylate-binding protein family. J Interferon Cytokine Res 27(1):44–52. https://doi.org/10.1089/jir.2007.0086
12. Braun E, Hotter D, Koepke L, Zech F, Groß R, Sparrer KMJ et al (2019) Guanylate-binding proteins 2 and 5 exert broad antiviral activity by inhibiting furin-mediated processing of viral envelope proteins. Cell Rep 27(7):2092–104.e10. https://doi.org/10.1016/j.celrep.2019.04.063
13. Tretina K, Park E-S, Maminska A, MacMicking JD (2019) Interferon-induced guanylate-binding proteins: guardians of host defense in health and disease. J Exp Med 216(3):482–500
14. Kutsch M, Coers J (2020) Human guanylate binding proteins: nanomachines orchestrating host defense. FEBS J. https://doi.org/10.1111/febs.15662
15. Zhang R, Li Z, Tang Y-D, Su C, Zheng C (2021) When human guanylate-binding proteins meet viral infections. J Biomed Sci 28(1):17. https://doi.org/10.1186/s12929-021-00716-8
16. Côrte-Real JV, Baldauf H-M, Melo-Ferreira J, Abrantes J, Esteves PJ (2022) Evolution of guanylate binding protein (GBP) genes in muroid rodents (Muridae and Cricetidae) reveals an outstanding pattern of gain and loss. Front Immunol. https://doi.org/10.3389/fimmu.2022.752186
17. Côrte-Real JV, Baldauf HM, Abrantes J, Esteves PJ (2021) 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 132:79–81. https://doi.org/10.1016/j.molimm.2021.01.025 (**Epub 2021/02/08**)
18. Huang S, Meng Q, Maminska A, MacMicking JD (2019) Cell-autonomous immunity by IFN-induced GBPs in animals and plants. Curr Opin Immunol 60:71–80. https://doi.org/10.1016/j.coi.2019.04.017 (**Epub 2019/06/09**)
19. Olszewski MA, Gray J, Vestal DJ (2006) In silico genomic analysis of the human and murine guanylate-binding protein (GBP) gene clusters. J Interferon Cytokine Res 26(5):328–352. https://doi.org/10.1089/jir.2006.26.328 (**Epub 2006/05/13**)
20. Li G, Zhang J, Sun Y, Wang H, Wang Y (2009) The evolutionarily dynamic IFN-inducible GTPase proteins play conserved immune functions in vertebrates and cephalochordates. Mol Biol Evol 26(7):1619–1630. https://doi.org/10.1093/molbev/msp074 (**Epub 2009/04/17**)
21. 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 94(15):7799–7806. https://doi.org/10.1073/pnas.94.15.7799
22. Nei M, Rooney AP (2005) Concerted and birth-and-death evolution of multigene families. Annu Rev Genet 39:121–152. https://doi.org/10.1146/annurev.genet.39.073003.112240
23. Wandel MP, Kim B-H, Park E-S, Boyle KB, Nayak K, Lagrange B et al (2020) Guanylate-binding proteins convert cytosolic bacteria into caspase-4 signaling platforms. Nat Immunol. https://doi.org/10.1038/s41590-020-0697-2
24. Meunier E, Wallet P, Dreier RF, Costanzo S, Anton L, Rühl S et al (2015) Guanylate-binding proteins promote activation of the AIM2 inflammasome during infection with *Francisella novicida*. Nat Immunol 16(5):476–484. https://doi.org/10.1038/ni.3119 (**Epub 2015/03/17**)
25. Shenoy AR, Wellington DA, Kumar P, Kassa H, Booth CJ, Cresswell P et al (2012) GBP5 promotes NLRP3 inflammasome assembly and immunity in mammals. Science 336(6080):481–485. https://doi.org/10.1126/science.1217141 (**Epub 2012/03/31**)
26. Feng M, Zhang Q, Wu W, Chen L, Gu S, Ye Y et al (2021) Inducible guanylate-binding protein 7 facilitates influenza A virus replication by suppressing innate immunity via NF-κB and JAK-STAT signaling pathways. J Virol 95(6):e02038-e2120. https://doi.org/10.1128/JVI.02038-20
27. Bentham AR, De la Concepcion JC, Mukhi N, Zdrzałek R, Draeger M, Gorenkin D et al (2020) A molecular roadmap to the plant immune system. J Biol Chem 295(44):14916–14935. https://doi.org/10.1074/jbc.REV120.010852 (**Epub 2020/08/21**)
28. Dangl JL, Horvath DM, Staskawicz BJ (2013) Pivoting the plant immune system from dissection to deployment. Science 341(6147):746–751. https://doi.org/10.1126/science.1236011 (**Epub 2013/08/21**)
29. Musseau C, Jorly J, Gadin S, Sørensen I, Deborde C, Bernillon S et al (2020) The tomato guanylate-binding protein SlGBP1 enables fruit tissue differentiation by maintaining endopolyploid cells in a non-proliferative state. Plant Cell 32(10):3188–3205. https://doi.org/10.1105/tpc.20.00245
30. Murphy AM, Zhou T, Carr JP (2020) An update on salicylic acid biosynthesis, its induction and potential exploitation by plant viruses. Curr Opin Virol 42:8–17. https://doi.org/10.1016/j.coviro.2020.02.008 (**Epub 2020/04/25**)
31. Dong Z, Zheng N, Hu C, Huang X, Chen P, Wu Q et al (2021) Genetic bioengineering of overexpressed guanylate binding protein family *BmAtlastin-n* enhances silkworm resistance to *Nosema bombycis*. Int J Biol Macromol 172:223–230. https://doi.org/10.1016/j.ijbiomac.2021.01.021 (**Epub 2021/01/17**)
32. Liu T-h, Dong X-l, Pan C-x, Du G-y, Wu Y-f, Yang J-g et al (2016) A newly discovered member of the Atlastin family, *BmAtlastin-n*, has an antiviral effect against BmNPV in *Bombyx mori*. Sci Rep 6(1):28946. https://doi.org/10.1038/srep28946
33. Robertsen B, Zou J, Secombes C, Leong J-A (2006) Molecular and expression analysis of an interferon-gamma-inducible guanylate-binding protein from rainbow trout (*Oncorhynchus mykiss*). Dev Comp Immunol 30(11):1023–1033. https://doi.org/10.1016/j.dci.2006.01.003
34. Cheng YS, Colonno RJ, Yin FH (1983) Interferon induction of fibroblast proteins with guanylate binding activity. J Biol Chem 258(12):7746–7750 (**Epub 1983/06/25**)
35. Tyrkalska SD, Candel S, Angosto D, Gómez-Abellán V, Martín-Sánchez F, García-Moreno D et al (2016) Neutrophils mediate *Salmonella* Typhimurium clearance through the GBP4 inflammasome-dependent production of prostaglandins. Nat

Commun 7:12077. https://doi.org/10.1038/ncomms12077 (**Epub 2016/07/02**)

36. Jin T, Huang M, Smith P, Jiang J, Xiao TS (2013) Structure of the caspase-recruitment domain from a zebrafish guanylate-binding protein. Acta Crystallogr Sect F Struct Biol Cryst Commun 69(Pt 8):855–860. https://doi.org/10.1107/S1744309113015558 (**Epub 2013/07/27**)
37. Valera-Pérez A, Tyrkalska SD, Viana C, Rojas-Fernández A, Pelegrín P, García-Moreno D et al (2019) WDR90 is a new component of the NLRC4 inflammasome involved in Salmonella Typhimurium resistance. Dev Comp Immunol 100:103428. https://doi.org/10.1016/j.dci.2019.103428
38. Zou Z, Meng Z, Ma C, Liang D, Sun R, Lan K (2017) Guanylate-binding protein 1 inhibits nuclear delivery of Kaposi's sarcoma-associated herpesvirus virions by disrupting formation of actin filament. J Virol 91(16):e00632-e717
39. Glitscher M, Himmelsbach K, Woytinek K, Schollmeier A, Johne R, Praefcke GJK et al (2021) Identification of the interferon-inducible GTPase GBP1 as major restriction factor for the hepatitis E virus. J Virol. https://doi.org/10.1128/jvi.01564-20 (**Epub 2021/01/22**)
40. Itsui Y, Sakamoto N, Kakinuma S, Nakagawa M, Sekine-Osajima Y, Tasaka-Fujita M et al (2009) Antiviral effects of the interferon-induced protein guanylate binding protein 1 and its interaction with the hepatitis C virus NS5B protein. Hepatology 50(6):1727–1737. https://doi.org/10.1002/hep.23195
41. Zhu Z, Shi Z, Yan W, Wei J, Shao D, Deng X et al (2013) Nonstructural protein 1 of influenza A virus interacts with human guanylate-binding protein 1 to antagonize antiviral activity. PLoS One 8(2):e55920. https://doi.org/10.1371/journal.pone.0055920 (**Epub 2013/02/14**)
42. Anderson SL, Carton JM, Lou J, Xing L, Rubin BY (1999) Interferon-induced guanylate binding protein-1 (GBP-1) mediates an antiviral effect against vesicular stomatitis virus and encephalomyocarditis virus. Virology 256(1):8–14. https://doi.org/10.1006/viro.1999.9614 (**Epub 1999/03/24**)
43. Pan W, Zuo X, Feng T, Shi X, Dai J (2012) Guanylate-binding protein 1 participates in cellular antiviral response to dengue virus. Virol J 9:292. https://doi.org/10.1186/1743-422x-9-292 (**Epub 2012/11/29**)
44. Nordmann A, Wixler L, Boergeling Y, Wixler V, Ludwig S (2012) A new splice variant of the human guanylate-binding protein 3 mediates anti-influenza activity through inhibition of viral transcription and replication. FASEB J 26(3):1290–1300. https://doi.org/10.1096/fj.11-189886
45. Krapp C, Hotter D, Gawanbacht A, McLaren Paul J, Kluge Silvia F, Stürzel Christina M et al (2016) Guanylate binding protein (GBP) 5 is an interferon-inducible inhibitor of HIV-1 infectivity. Cell Host Microbe 19(4):504–514. https://doi.org/10.1016/j.chom.2016.02.019
46. Srinivasachar Badarinarayan S, Shcherbakova I, Langer S, Koepke L, Preising A, Hotter D et al (2020) HIV-1 infection activates endogenous retroviral promoters regulating antiviral gene expression. Nucleic Acids Res 48(19):10890–10908. https://doi.org/10.1093/nar/gkaa832
47. Li Z, Qu X, Liu X, Huan C, Wang H, Zhao Z et al (2020) GBP5 is an interferon-induced inhibitor of respiratory syncytial virus. J Virol 94(21):e01407-e1420. https://doi.org/10.1128/JVI.01407-20
48. Feng J, Cao Z, Wang L, Wan Y, Peng N, Wang Q et al (2017) Inducible GBP5 mediates the antiviral response via interferon-related pathways during influenza A virus infection. J Innate Immun 9(4):419–435
49. Ma G, Huang J, Sun N, Liu X, Zhu M, Wu Z et al (2008) Molecular characterization of the porcine GBP1 and GBP2 genes. Mol Immunol 45(10):2797–2807. https://doi.org/10.1016/j.molimm.2008.02.007
50. Li LF, Yu J, Li Y, Wang J, Li S, Zhang L et al (2016) Guanylate-binding protein 1, an interferon-induced GTPase, exerts an antiviral activity against classical swine fever virus depending on its GTPase activity. J Virol 90(9):4412–4426. https://doi.org/10.1128/JVI.02718-15 (**Epub 20160414**)
51. Kommadath A, Bao H, Choi I, Reecy JM, Koltes JE, Fritz-Waters E et al (2017) Genetic architecture of gene expression underlying variation in host response to porcine reproductive and respiratory syndrome virus infection. Sci Rep 7:46203. https://doi.org/10.1038/srep46203 (**Epub 20170410**)
52. Khatun A, Nazki S, Jeong CG, Gu S, Mattoo SUS, Lee SI et al (2020) Effect of polymorphisms in porcine guanylate-binding proteins on host resistance to PRRSV infection in experimentally challenged pigs. Vet Res 51(1):14. https://doi.org/10.1186/s13567-020-00745-5 (**Epub 20200219**)
53. Koltes JE, Fritz-Waters E, Eisley CJ, Choi I, Bao H, Kommadath A et al (2015) Identification of a putative quantitative trait nucleotide in guanylate binding protein 5 for host response to PRRS virus infection. BMC Genom 16:412. https://doi.org/10.1186/s12864-015-1635-9 (**Epub 20150528**)
54. Gol S, Estany J, Fraile LJ, Pena RN (2015) Expression profiling of the GBP1 gene as a candidate gene for porcine reproductive and respiratory syndrome resistance. Anim Genet 46(6):599–606. https://doi.org/10.1111/age.12347 (**Epub 20150911**)
55. Gu T, Yu D, Fan Y, Wu Y, Yao YL, Xu L et al (2019) Molecular identification and antiviral function of the guanylate-binding protein (GBP) genes in the Chinese tree shrew (*Tupaia belangeri chinesis*). Dev Comp Immunol 96:27–36. https://doi.org/10.1016/j.dci.2019.02.014 (**Epub 2019/03/01**)
56. Wehner M, Herrmann C (2010) Biochemical properties of the human guanylate binding protein 5 and a tumor-specific truncated splice variant. FEBS J 277(7):1597–1605. https://doi.org/10.1111/j.1742-4658.2010.07586.x (**Epub 2010/02/26**)
57. Nantais DE, Schwemmle M, Stickney JT, Vestal DJ, Buss JE (1996) Prenylation of an interferon-gamma-induced GTP-binding protein: the human guanylate binding protein, huGBP1. J Leukoc Biol 60(3):423–431. https://doi.org/10.1002/jlb.60.3.423 (**Epub 1996/09/01**)
58. Modiano N, Lu YE, Cresswell P (2005) Golgi targeting of human guanylate-binding protein-1 requires nucleotide binding, isoprenylation, and an IFN-gamma-inducible cofactor. Proc Natl Acad Sci USA 102(24):8680–8685. https://doi.org/10.1073/pnas.0503227102 (**Epub 2005/06/03**)
59. Ngo CC, Man SM (2017) Mechanisms and functions of guanylate-binding proteins and related interferon-inducible GTPases: roles in intracellular lysis of pathogens. Cell Microbiol 19(12):e12791. https://doi.org/10.1111/cmi.12791
60. Santos JC, Broz P (2018) Sensing of invading pathogens by GBPs: at the crossroads between cell-autonomous and innate immunity. J Leukoc Biol 104(4):729–735. https://doi.org/10.1002/jlb.4mr0118-038r (**Epub 2018/07/19**)
61. Mostafavi S, Yoshida H, Moodley D, LeBoité H, Rothamel K, Raj T et al (2016) Parsing the interferon transcriptional network and its disease associations. Cell 164(3):564–578. https://doi.org/10.1016/j.cell.2015.12.032 (**Epub 2016/01/30**)
62. Degrandi D, Konermann C, Beuter-Gunia C, Kresse A, Würthner J, Kurig S et al (2007) Extensive characterization of IFN-induced GTPases mGBP1 to mGBP10 involved in host defense. J Immunol 179(11):7729–7740. https://doi.org/10.4049/jimmunol.179.11.7729
63. Kim BH, Shenoy AR, Kumar P, Das R, Tiwari S, MacMicking JD (2011) A family of IFN-γ-inducible 65-kD GTPases protects against bacterial infection. Science 332(6030):717–721. https://doi.org/10.1126/science.1201711

64. Gu T, Yu D, Xu L, Yao YL, Zheng X, Yao YG (2021) Tupaia guanylate-binding protein 1 interacts with vesicular stomatitis virus phosphoprotein and represses primary transcription of the viral genome. Cytokine 138:155388. https://doi.org/10.1016/j.cyto.2020.155388 (**Epub 2020/12/04**)
65. Gu T, Yu D, Xu L, Yao Y-L, Yao Y-G (2021) Tupaia GBP1 interacts with STING to initiate autophagy and restrict herpes simplex virus type 1 infection. J Immunol 207(11):2673. https://doi.org/10.4049/jimmunol.2100325
66. Li R, Zanin M, Xia X, Yang Z (2018) The tree shrew as a model for infectious diseases research. J Thorac Dis 10(Suppl 19):S2272–S2279. https://doi.org/10.21037/jtd.2017.12.121 (**Epub 2018/08/18**)
67. Xu L, Yu DD, Ma YH, Yao YL, Luo RH, Feng XL et al (2020) COVID-19-like symptoms observed in Chinese tree shrews infected with SARS-CoV-2. Zool Res 41(5):517–526. https://doi.org/10.24272/j.issn.2095-8137.2020.053 (**Epub 2020/07/24**)
68. Carter CC, Gorbacheva VY, Vestal DJ (2005) Inhibition of VSV and EMCV replication by the interferon-induced GTPase, mGBP-2: differential requirement for wild-type GTP binding domain. Arch Virol 150(6):1213–1220. https://doi.org/10.1007/s00705-004-0489-2 (**Epub 2005/02/18**)
69. Yu P, Li Y, Li Y, Miao Z, Peppelenbosch MP, Pan Q (2020) Guanylate-binding protein 2 orchestrates innate immune responses against murine norovirus and is antagonized by the viral protein NS7. J Biol Chem 295(23):8036–8047. https://doi.org/10.1074/jbc.RA120.013544 (**Epub 2020/04/30**)
70. Degrandi D, Kravets E, Konermann C, Beuter-Gunia C, Klümpers V, Lahme S et al (2013) Murine guanylate binding protein 2 (mGBP2) controls *Toxoplasma gondii* replication. Proc Natl Acad Sci USA 110(1):294–299. https://doi.org/10.1073/pnas.1205635110 (**Epub 2012/12/19**)
71. Cerqueira DM, Gomes MTR, Silva ALN, Rungue M, Assis NRG, Guimarães ES et al (2018) Guanylate-binding protein 5 licenses caspase-11 for Gasdermin-D mediated host resistance to *Brucella abortus* infection. PLoS Pathog 14(12):e1007519-e. https://doi.org/10.1371/journal.ppat.1007519
72. Corsetti PP, de Almeida LA, Gonçalves ANA, Gomes MTR, Guimarães ES, Marques JT et al (2018) miR-181a-5p regulates TNF-α and miR-21a-5p influences gualynate-binding protein 5 and IL-10 expression in macrophages affecting host control of *Brucella abortus* infection. Front Immunol 9:1331. https://doi.org/10.3389/fimmu.2018.01331 (**Epub 2018/06/27**)
73. Man SM, Karki R, Malireddi RK, Neale G, Vogel P, Yamamoto M et al (2015) The transcription factor IRF1 and guanylate-binding proteins target activation of the AIM2 inflammasome by *Francisella* infection. Nat Immunol 16(5):467–475. https://doi.org/10.1038/ni.3118 (**Epub 2015/03/17**)
74. Place DE, Malireddi RKS, Kim J, Vogel P, Yamamoto M, Kanneganti T-D (2021) Osteoclast fusion and bone loss are restricted by interferon inducible guanylate binding proteins. Nat Commun 12(1):496. https://doi.org/10.1038/s41467-020-20807-8
75. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H et al (2000) The protein data bank. Nucleic Acids Res 28(1):235–242. https://doi.org/10.1093/nar/28.1.235 (**Epub 1999/12/11**)
76. Sehnal D, Bittrich S, Deshpande M, Svobodová R, Berka K, Bazgier V et al (2021) Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. Nucleic Acids Res 49(W1):W431–W437. https://doi.org/10.1093/nar/gkab314

Appendix B: Acknowledgements and Curriculum Vitae

Appendix B1: Acknowledgements

I would like to hereby thank all involved people, without you this work would not have been possible.

In particular, I would like to thank my doctoral supervisor Hanna-Mari Baldauf for her constant support, challenge, discussion, mentorship, and her kind and optimistic manner, which helped me to grow not only scientifically but also on a personal level.

I would like to thank João Vasco Côrte-Real for our great and successful teamwork that resulted in two papers included and more to come.

I would like to thank Joana Abrantes and Pedro José Esteves for my great time in Portugal and constant input to our work.

I would like to thank my interns Caroline Kliem, Margarita Shnipova and Bhavna Menon for the pleasant internships and results.

I would like to thank all the other collaborators for the great teamwork that resulted in the papers included in this thesis: Augusto del Pozo Ben, Sharmeen Fayyaz, Moritz Petersen, Rishikesh Lotke, Dana Matzek, Lena Pfaff, Ana Pinheiro, João Pedro Marques, José Melo-Ferreira, Bastian Popper and Daniel Sauter.

I would like to thank my TAC Christine Josenhans and Barbara Adler for the fruitful discussions and input.

I would like to thank all (former) members of AG Baldauf for the great time. Especially, Ramya Nair and Alejandro Salinas for their introduction to the lab and constant support and discussion.

I would like to thank all members of AG Keppler for the nice working environment.

I would also like to thank my former co-workers and mentors, that supported me throughout my academic career.

These thanks are not only meant on a professional basis but also a personal level having a good time and ongoing friendships.

I would like to thank my family and friends whether still here with me or not. I would like to thank and highlight Liv for supporting me on a daily basis through my doctorate and especially my parents Susanne and Manfred who greatly supported me during my whole life.