

RESEARCH ARTICLE

# Not so unique to Primates: The independent adaptive evolution of TRIM5 in Lagomorpha lineage

Ana Águeda-Pinto<sup>1,2</sup>, Ana Lemos de Matos<sup>3</sup>, Ana Pinheiro<sup>1,2</sup>, Fabiana Neves<sup>1</sup>, Patrícia de Sousa-Pereira<sup>1,2</sup>, Pedro J. Esteves<sup>1,2,4\*</sup>

**1** CIBIO/InBio—Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, Vairão, Portugal, **2** Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, Porto, Portugal, **3** Center for Immunotherapy, Vaccines, and Virotherapy (CIVV), The Biodesign Institute, Arizona State University, Tempe, Arizona, United States of America, **4** CITS—Centro de Investigação em Tecnologias da Saúde, IPSN, CESPU, Gandra, Portugal

\* [pjesteves@cibio.up.pt](mailto:pjesteves@cibio.up.pt)



**OPEN ACCESS**

**Citation:** Águeda-Pinto A, Lemos de Matos A, Pinheiro A, Neves F, de Sousa-Pereira P, Esteves PJ (2019) Not so unique to Primates: The independent adaptive evolution of TRIM5 in Lagomorpha lineage. PLoS ONE 14(12): e0226202. <https://doi.org/10.1371/journal.pone.0226202>

**Editor:** Christine A. Kozak, National Institute of Allergy and Infectious Diseases, UNITED STATES

**Received:** May 14, 2019

**Accepted:** November 21, 2019

**Published:** December 12, 2019

**Copyright:** © 2019 Águeda-Pinto et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was supported by the project NORTE-01-0145-FEDER-000007, as a result of Norte Portugal Regional Operational Programme (NORTE2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF) and by the Foundation for Science and Technology (SFRH/BD/

## Abstract

The plethora of restriction factors with the ability to inhibit the replication of retroviruses have been widely studied and genetic hallmarks of evolutionary selective pressures in Primates have been well documented. One example is the tripartite motif-containing protein 5 alpha (TRIM5 $\alpha$ ), a cytoplasmic factor that restricts retroviral infection in a species-specific fashion. In Lagomorphs, similarly to what has been observed in Primates, the specificity of TRIM5 restriction has been assigned to the PRYSPRY domain. In this study, we present the first insight of an intra-genus variability within the Lagomorpha TRIM5 PRYSPRY domain. Remarkably, and considering just the 32 residue-long v1 region of this domain, the deduced amino acid sequences of Daurian pika (*Ochotona dauurica*) and steppe pika (*O. pusilla*) evidenced a high divergence when compared to the remaining *Ochotona* species, presenting values of 44% and 66% of amino acid differences, respectively. The same evolutionary pattern was also observed when comparing the v1 region of two *Sylvilagus* species members (47% divergence). However, and unexpectedly, the PRYSPRY domain of *Lepus* species exhibited a great conservation. Our results show a high level of variation in the PRYSPRY domain of Lagomorpha species that belong to the same genus. This suggests that, throughout evolution, the Lagomorpha TRIM5 should have been influenced by constant selective pressures, likely as a result of multiple different retroviral infections.

## Introduction

Long periods of co-evolution between retroviruses and their hosts have resulted in the emergence of numerous host defense mechanisms important for an antiviral response, as well as the selection of diverse viral countermeasures [1, 2]. The host attempts to circumvent the viral infection firstly by the activation of innate immune responses initiated by pattern-recognition receptors (PRRs), and secondly by subsequent inducible activation of multiple protective signaling pathways [3, 4]. Restriction factors, sometimes induced by interferon (IFN) signaling,

128752/2017 to AAP, SFRH/BPD/117451/2016 to AP, PD/BD/52602/2014 to PSP and IF/00376/2015 to PJE). The funders had no role in the design of the study, collection, analysis and interpretation of data, and in writing the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

are widely expressed components of the innate host immune system with the ability to inhibit the replication of viruses during their life cycle in host cells [5, 6]. As a result of the selective pressure (positive selection) exerted by viral infections, restriction factors can undergo a rapid evolution of their coding sequences, allowing the host to keep pace with the viral adaptations against these host factors [1, 7]. The study of human innate immune factors responsible for detecting and fighting viral infections has been investigated by comprehensive surveys of the evolutionary relationships between orthologous restriction factor genes in Primates. One such example is the tripartite motif-containing protein 5 alpha (TRIM5 $\alpha$ ), a cytoplasmic factor that restricts retroviral infection in a species-specific manner [8, 9]. TRIM5 $\alpha$  belongs to the TRIM family, characterized by the presence of a RING domain, one or two B-box domains, and a coiled-coil domain [8, 10]. The specificity of Primate TRIM5 $\alpha$  restriction has been assigned to the “variable loops” (v1, v2, v3 and v4) of the C-terminal PRYSPRY domain (also known as B30.2 domain), as the sequence identity between species is low and strong evidence of positive selection during Primate evolution has been detected [9, 11].

Given the significance of the Primate TRIM5 $\alpha$  evolutionary history, several orthologs have been described for other mammalian species [12–15]. For example, in Lagomorpha a differentiation in the PRYSPRY domain was observed and, even more revealingly, in genera and species with an evolutionary divergence timeline similar to Primates evolution [16]. The order Lagomorpha is divided into two families which diverged around 30–35 million years ago (Mya): Ochotonidae and Leporidae [17]. While Ochotonidae only contains one extant genus, *Ochotona*, the Leporidae family includes 11 genera that comprises *Lepus*, *Sylvilagus* and *Oryctolagus* [18]. TRIM5 of genera *Oryctolagus*, *Sylvilagus* and *Lepus* exhibit an overall similarity between 89% and 91%, but for the variable loop 1 (v1) of the PRYSPRY domain similarity values decrease to 50% [16]. When including *Ochotona*, for the stretch of 30 amino acids in v1 region, 11 have undergone positive selective pressure during Lagomorpha evolution [16]. Such evidence of gene adaptation in Lagomorpha TRIM5 variable loops is indeed very comparable to the selective pressures imposed by exposure to retroviral capsid (CA) in primate TRIM5 genes.

Due to the absence of known infecting exogenous retroviruses in Lagomorpha species, the discovery of the rabbit endogenous lentivirus type K (RELK) in the genome of several Leporidae genera raised the hypothesis that this retroviral relic could have played a role in the selection of the leporid TRIM5 [19, 20]. Nevertheless, the 12 million year old RELK is apparently absent in members of the other Lagomorpha family, the Ochotonidae [20]. Recently, the ability of the Lagomorpha TRIM5 to restrict a wide range of retroviruses, and particularly viral vectors containing the RELK CA, have been examined and supported the previous suggestion of RELK exerting selective pressures on leporid TRIM5 [21, 22].

The coincident evolutionary divergence aspects between Lagomorpha and Primates orders, and the sequence variation in PRYSPRY domain between different genera of Lagomorpha [16,18], propelled us to conduct a more exhaustive study of the PRYSPRY domain throughout Lagomorpha evolution. While in the extensively studied Primates the PRYSPRY domain sequence and length variations occurred mainly between genera (implying high divergence times) [23–25], we observed an array of intra-genera differences in Lagomorpha. These results provide evidence of different and contemporary selection events throughout Lagomorpha speciation.

## Results

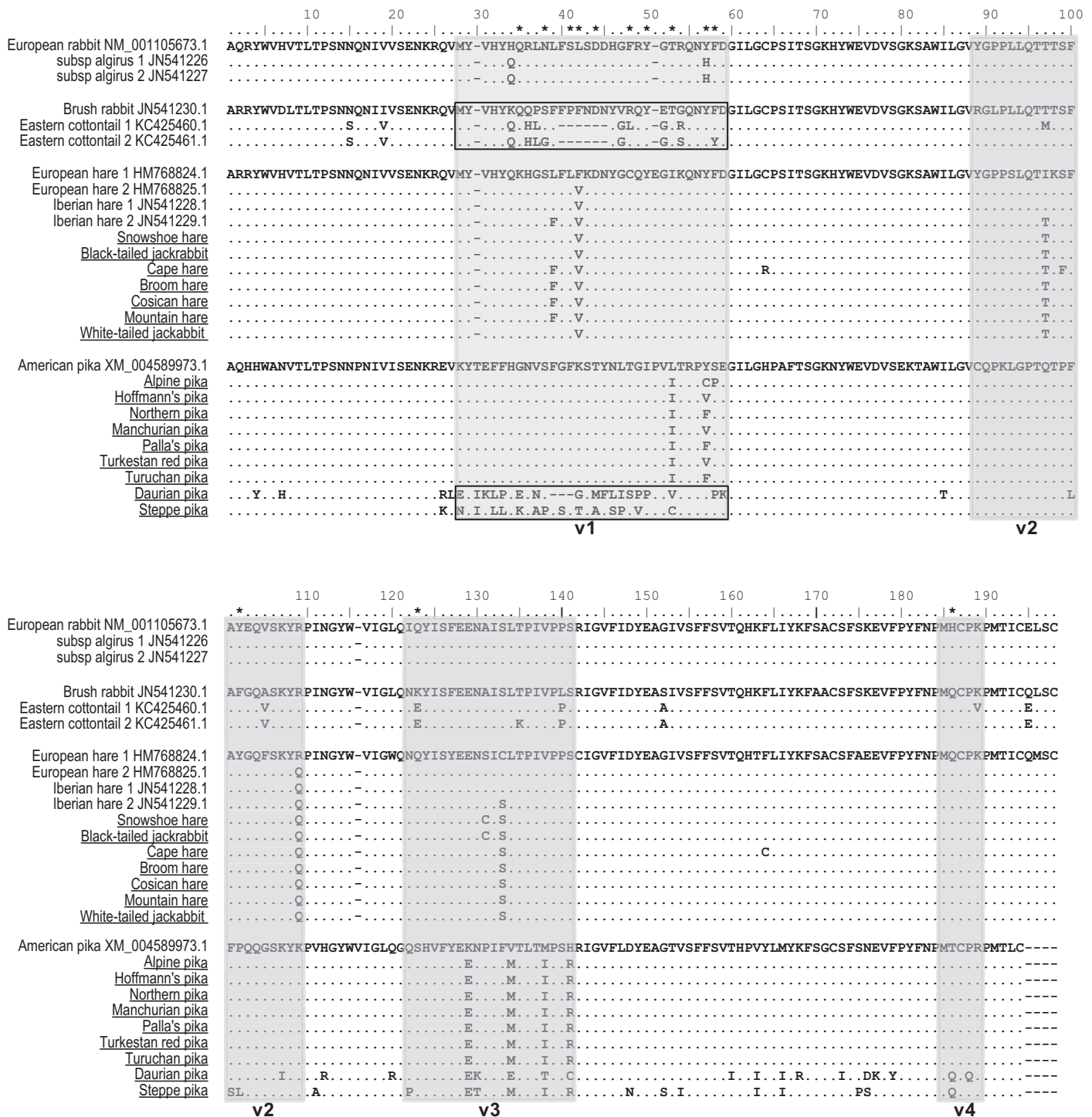
It is known that the PRYSPRY domain is primarily responsible for the direct binding of TRIM5 $\alpha$  protein to retrovirus capsids, leading to structural damage of the incoming viral core

and disruption of virus infection [26]. The sequence of *TRIM5* gene has been already obtained in a limited number of Lagomorpha species, namely the European rabbit (*Oryctolagus cuniculus*), the European brown hare (*Lepus europaeus*), the Iberian hare (*L. granatensis*), the brush rabbit (*Sylvilagus bachmani*) and the eastern cottontail (*S. floridanus*) [16, 22]. Therefore, in this study, we performed a more extensive intra-genus study of the orthologous Leporidae and Ochotonidae PRYSPRY domains, increasing the number of sequences from additional members of *Lepus* and *Ochotona* genera. Here, we present the PRYSPRY domain sequences of seven other *Lepus* species, and also provide the first insight into the variability of *TRIM5* PRYSPRY domain of nine pika (*Ochotona*) species, all of them representative of a wide geographic distribution. In Fig 1 all the deduced PRYSPRY sequences from *Lepus* and *Ochotona* species, along with other Lagomorpha *TRIM5* sequences that have been previously published [16, 22] are represented. The phylogenetic tree generated from the Lagomorpha sequences was supported by high bootstrap values (Fig 2) and the phylogenetic relationships match the species tree [18]; however, Daurian pika and steppe pika present longer branches comparatively to the remaining species (Fig 2).

### Identification of positively selected sites in Lagomorpha PRYSPRY domain

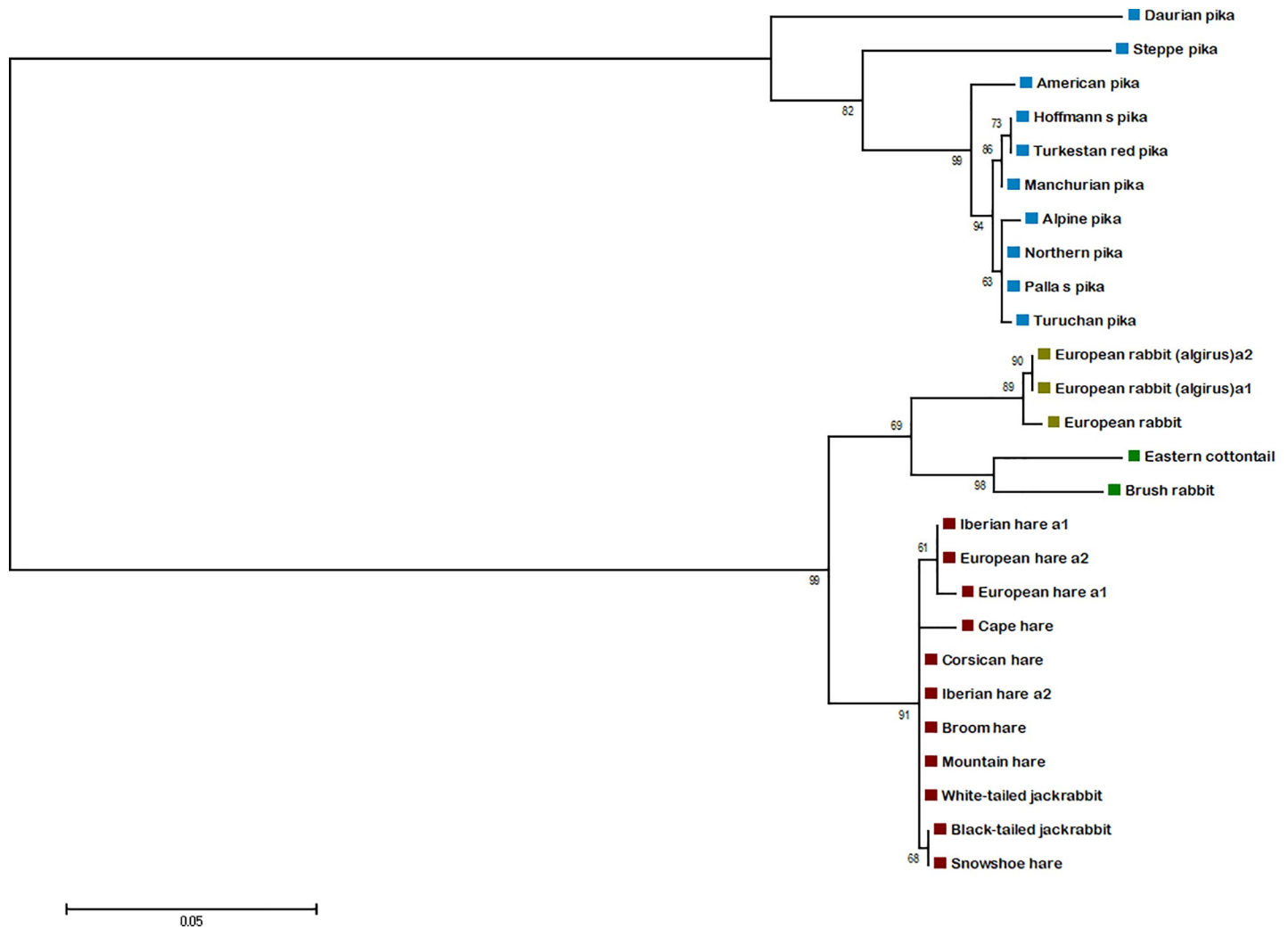
To look for evidence of potential selection pressures acting on the PRYSPRY domain, we used the dataset of Lagomorpha sequences mentioned above and implemented a maximum-likelihood (ML) approach, by using PAML and Datamonkey softwares (see [methods](#) for more information). For most protein-coding genes, the rate between nonsynonymous and synonymous substitutions (dN/dS) is a measure of natural selection, with positive selection (dN/dS > 1) acting against the common genotype [28–30]. Comparison of the *TRIM5 $\alpha$*  gene dN/dS ratio of several Primates has previously revealed strong evidence of positive selection acting in different species groups (i.e. New World Monkeys, Old World Monkeys and Homioids), more specifically in the PRYSPRY domain [9]. Similar results were also observed for three leporid genera, where among the 25 positively-selected codons identified for the *TRIM5* gene, 20 were located in the PRYSPRY domain [16]. In this study, we were able to locate 13 sites that reflect strong positive selection pressure in the Lagomorpha's PRYSPRY domain (S1 Appendix). Among the 13 individual sites detected, only 5 overlap with the previous work [16], a discrepancy resulting from the addition of a greater number of species to *Lepus* and *Ochotona* genera. As seen in Fig 1, residues identified as being under positive selection are within the variable loops v1-v4 (residues under positive selection are marked with asterisk "\*" in the alignment). Most importantly, 10 of the 13 residues fall within the v1 variable region, known to be a key determinant domain for virus specificity of *TRIM5 $\alpha$*  restriction activity in Primates [31, 32]. In addition to the strong positive selection found in the v1 region of PRYSPRY domain, this domain has undergone an unusual number of indels. In fact, a deletion of eight predicted amino acid residues has occurred in the eastern cottontail genome, whereas the orthologous v1 region of the European rabbit and *Lepus* genus *TRIM5* sequences encode a protein two and one amino acids, respectively, shorter than the American pika sequence.

Despite the several modifications that can be observed between different Lagomorpha genera, major differences can also be observed within each genus. Indeed, the deduced amino acid differences between the two *Sylvilagus* species were very dramatic: for example, when comparing the v1 region between the eastern cottontail rabbit and the brush rabbit, the two species registered a remarkable 47% divergence, resulting from the substitution of eight amino acid residues and a deletion of six residues (considered a significant "difference" between sequences) in the eastern cottontail rabbit genome (Table 1). Moreover, when the number of



**Fig 1. Lagomorpha PRYSPRY sequences.** PRYSPRY sequences from 23 Lagomorpha genomes, two belonging to the *Oryctolagus* genus, two from *Sylvilagus* genus, nine belonging to the *Lepus* genus and ten to the *Ochotona* genus. Additional members of the *Lepus* and *Ochotona* genera used in this study are underlined. Variable loops from the PRYSPRY domain (v1, v2, v3 and v4) are defined by grey boxes. High variable residues are shown in black boxes. Positively-selected codon positions identified at least by three ML methods are indicated by an asterisk (\*).

<https://doi.org/10.1371/journal.pone.0226202.g001>



**Fig 2. Phylogenetic analysis of PRYSPRY domain of Lagomorpha.** Nucleotide sequences of the PRYSPRY domain of Lagomorpha species were used to construct a Maximum Likelihood tree using PhyML v3.0 [27]. The support of the resulting nodes was estimated using 1000 bootstrap replicates. The bootstraps values are indicated on the branches. Blue squares: *Ochotona* species; yellow squares: *Oryctolagus* species; green squares: *Sylvilagus* species; red squares: *Lepus* species.

<https://doi.org/10.1371/journal.pone.0226202.g002>

non-synonymous substitutions that gave rise to non-synonymous sites was plotted in a sliding window, this value was undoubtedly higher in the v1 region when compared with the remaining regions of the PRYSPRY domain (Fig 3). Considering the *Ochotona* genus, this domain is quite conserved amongst the different species, with the clear exception of the Daurian pika and the steppe pika (Fig 1). The PRYSPRY variable regions of these two pika species evidenced a high divergence when compared to the remaining *Ochotona*, with special incidence in the 32 residue-long v1 region, showing values of 44% (steppe pika) and 66% (Daurian pika, including the 3 amino acid residues deletion) of amino acid differences (Table 1). Interestingly, the divergence between these two species in this particular region was also extremely high (62%). For the same species, the estimated dN/dS values of the PRYSPRY v1 region showed evidence of positive selection acting on this domain (Table 1). Also, the sliding window in Fig 3 shows that in the v1 region the number of non-synonymous sites was higher than for the remaining regions of the PRYSPRY domain.

**Table 1. Pairwise estimation of non-synonymous to synonymous substitution ratios (dN/dS) for the PRYSPRY v1 region by using Nei-Gojobori (Jukes-Cantor correction) method [33].**

PRYSPRY v1 regions <sup>a</sup> for comparison	% aa difference <sup>b</sup>	d <sub>N</sub> /d <sub>S</sub> <sup>c</sup>
Brush rabbit vs European rabbit	50 (15/30)	1,96
Eastern cottontail 2 vs European rabbit	50 (15/30)	3,82
Brush rabbit vs eastern cottontail 2	47 (14/30)	0.40
Daurian pika vs American pika	66 (21/32)	5,07
Steppe pika vs American pika	44 (14/32)	3,81
Daurian pika vs steppe pika	62 (20/32)	3,85

<sup>a</sup> Lagomorpha species PRYSPRY v1 region is indicated in Fig 1.

<sup>b</sup> The percentages of amino acid residues that differ between the PRYSPRY v1 region of the indicated species are shown. The percentages were calculated as follows: number of different residues/total number of residues compared x 100%. Indels in one sequence were counted as differences and also contributed to the total number of residues compared.

<sup>c</sup> Pairwise estimation of non-synonymous to synonymous substitution ratios (d<sub>N</sub>/d<sub>S</sub>) using Nei-Gojobori (Jukes-Cantor correction) method.

<https://doi.org/10.1371/journal.pone.0226202.t001>

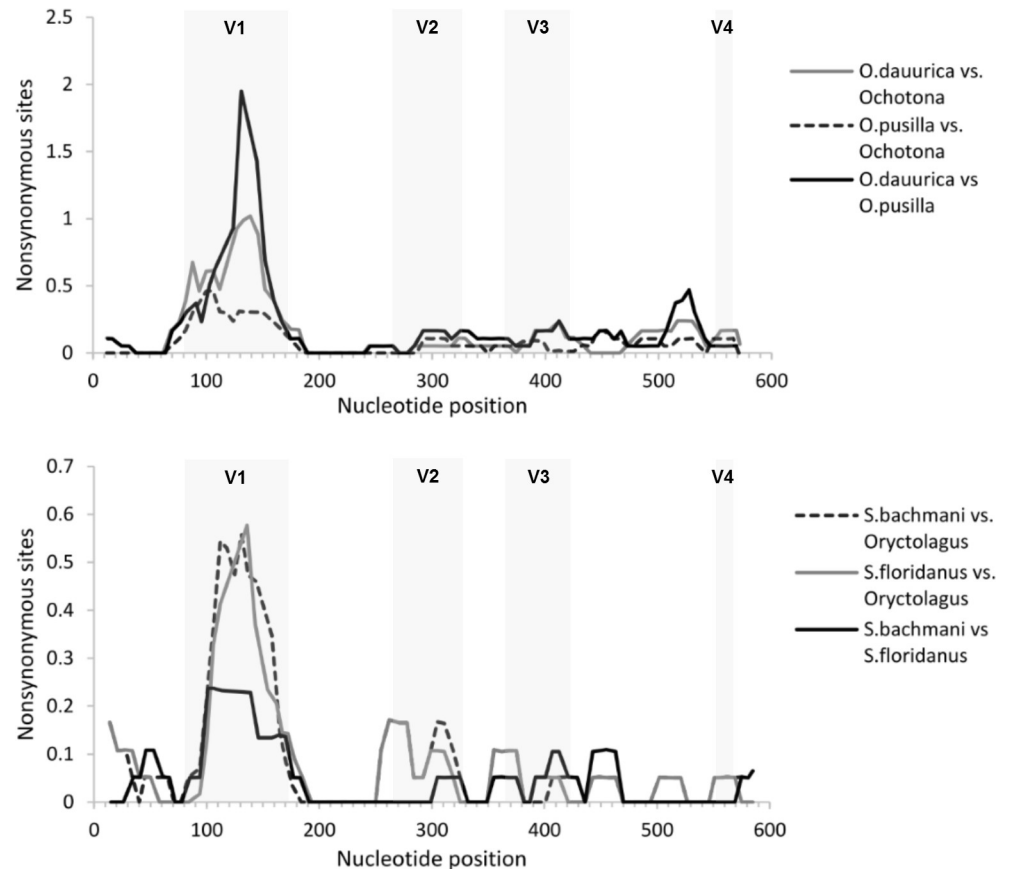
## Evolution of TRIM5α PRYSPRY domain throughout Primate lineages

To trace whether Primates also manifest a high intra-genus variability within their orthologous PRYSPRY domains, we collected and aligned the PRYSPRY domain coding sequences from 54 Primate species (S2 Appendix), representing ~45 million years of the evolutionary history of Primates [34]. Only the sequences of genera with two or more species were collected, allowing us to query TRIM5α Primates intra-genera differences (S3 Appendix). The translated nucleotide sequences were aligned against *Homo sapiens* PRYSPRY domain and are represented in S2 Appendix.

New World Primates (Platyrrhini) diverged from a common ancestor with Catarrhini about 45 Mya [34]. We were able to obtain 24 sequences of Platyrrhini species, representative of 9 different genera: *Callithrix* (three species), *Mico* (three species), *Leontopithecus* (two species), *Saguinus* (five species), *Aotus* (two species), *Saimiri* (three species), *Alouatta* (two species), *Pitheca* (two species) and *Callicebus* (two species). From the obtained alignment (S2 Appendix), and according to previous reports [23], it is possible to observe that the PRYSPRY domain from Platyrrhini species is very different from the human PRYSPRY domain. In fact, the Platyrrhini species v1 loop is nine amino acid shorter than the v1 loop of *H. sapiens*. Of particular note, the two species of *Alouatta* sp. possess an indel of 62 amino acids in the PRYSPRY domain, a feature that is not observed in any other Platyrrhini species.

Catarrhini group includes Cercopithecoidea (Old World Monkeys) and Hominoidea (human, great apes and gibbons). Cercopithecoidea comprises two big subfamilies, Colobinae and Cercopithecinae, with divergence times of ~18 Mya [34]. For Colobinae, we were only able to collect two PRYSPRY sequences, both for *Rhinopithecus* genus. Regarding Cercopithecinae, 19 sequences were obtained, belonging to five different genera: *Cercopithecus* (four species), *Chlorocebus* (four species), *Cercocebus* (two species), *Papio* (three species) and *Macaca* (six species). Interestingly, all *Chlorocebus* species have a longer PRYSPRY domain, as a result of a tandem duplication (S2 Appendix). For Hominoidea, which radiation is approximately 13 Million years old [34], sequences from *Pan* (two species), *Pongo* (two species), *Hylobates* (three species) and *Nomascus* (two species) were obtained.

Consistent with previous reports [23, 35], high variation in the PRYSPRY domain among different families of Primates was observed. However, as readily observed in the aligned



**Fig 3. Sliding-window analysis to detect nucleotide differences between PRYSPRY domains from different Lagomorphs.** Representation of the non-synonymous per non-synonymous sites between *Sylvilagus bachmani* vs *Oryctolagus* genus, *S. florianus* vs *Oryctolagus* genus and *S. bachmani* vs *S. florianus*; *Ochotona daurica* vs *Ochotonidae* family, *O. pusilla* vs *Ochotonidae* family and *O. daurica* vs *O. pusilla*.

<https://doi.org/10.1371/journal.pone.0226202.g003>

sequences (S2 Appendix), and particularly in the highlighted PRYSPRY v1 region, the divergence between Primate species in the same genus was nearly nonexistent, suggesting that much of the evolutionary pressures on this restriction factor were exerted prior to intra-genus speciation.

## Discussion

Under normal circumstances, the host TRIM5 $\alpha$  protein binds to the retrovirus capsid (CA) protein via its PRYSPRY domain, blocking the retrovirus replication early in its life cycle [36]. It is now established that species variation on the PRYSPRY domain led to differences in their ability to restrict retroviruses and, therefore, PRYSPRY sequence determines which retrovirus a specific TRIM5 $\alpha$  will restrict. For example, human TRIM5 $\alpha$  is not effective against human immunodeficiency virus (HIV-1), while TRIM5 $\alpha$  proteins encoded by rhesus macaque are able to efficiently restrict this lentivirus [23, 36]. In the light of the extensively described role of TRIM5 $\alpha$  PRYSPRY domain in Primates [37–40], previous studies have also supported the same role of this domain in Lagomorpha species. In fact, active TRIM5 proteins were identified and described in several lagomorphs, including the European rabbit, European brown hare, cottontail rabbit and American pika [21, 22]. Evolutionary analysis also showed that TRIM5 $\alpha$  has evolved under positive selection in Primates and that this selection has been

directed to the PRYSPRY domain [9, 21]. Accordingly, our results reported here reveal that the Lagomorpha PRYSPRY domain is under strong positive selection, with 10 of the 13 positive sites falling in the v1 variable loop (Fig 1). Moreover, our observations were reinforced by the sliding-window analysis of PRYSPRY nucleotide divergence, especially when considering *Ochotona* and *Sylvilagus* species (Fig 3).

Our results reveal that the nine studied *Lepus* species exhibit a very conserved PRYSPRY domain (Fig 1), a relevant finding since *Lepus* is the most geographically distributed genus of all Lagomorpha and some variability on the domain was expected [18]. Fossil records along with phylogenetic approaches suggest that North America is the *Lepus* region of origin [18]. The global development of temperate grasslands (7 to 5 Mya) and the formation of the west Antarctic ice sheet (6.5 Mya) enabled the development of land bridges and consequent rapid expansion and radiation of the *Lepus* genus through Eurasia and into Africa [18]. However, even when using species with different dispersal times and representative of American, European, Asian and African continents, *Lepus* spp. PRYSPRY domain apparently did not undergo strong selective pressures.

A different evolutionary history is reflected by the *Sylvilagus* TRIM5 PRYSPRY domain (Fig 1). The two *Sylvilagus* species included in this study, the brush rabbit and the eastern cottontail, have a divergence time of ~5 Mya and, yet, they present a dramatic divergence on the PRYSPRY v1 region. Interestingly, it appears that the divergence within this domain causes slightly different restriction phenotypes when compared to the TRIM5 PRYSPRY domains of the remaining Leporids. Murine leukemia viruses (MLV, gammaretroviruses) are generally insensitive to Leporid's TRIM5 action [21, 22]. However, it is striking that the TRIM5 of the eastern cottontail rabbit is able to restrict two strains of MLV (N-MLV and B-MLV) and partially restrict a third one, Mo-MLV [22]. No known exogenous infecting retroviruses have been described specifically for *Sylvilagus* genus, yet such striking evidence of acting selection on the genus' TRIM5 probably reflects the action of species-specific exogenous retroviruses still to be found or possibly recently extinct.

Similar to the pattern observed for *Sylvilagus*, the PRYSPRY domain of the ten *Ochotona* species included in this study is also a striking example of ongoing diversification. Eight of the studied species have a conserved TRIM5 PRYSPRY domain (Fig 2). However, the striking evidence of diversification in Daurian pika and steppe pika PRYSPRY domain, not only on the v1 region but throughout the full domain, highly supports the relatively recent existence of species-specific infecting retroviruses.

Recent studies identified the v1 loop as being the most important region of the PRYSPRY domain for the retroviral CA interaction. In humans, it was described that v1 loop has critical residues that allow different conformational changes responsible for the adaptability of this protein to the varying curvatures of retrovirus CA [31, 41, 42]. Studies on Rhesus monkey PRYSPRY domain showed that v2-v3 loops are located in the SPRY subdomain, while v1 loop residues locate in the PRY subdomain [26]. These findings prompted Yang and collaborators (2012) to suggest that the acquisition of PRY domain by the more ancient SPRY domain might be correlated with the emergence of the viral capsid-sensing capacity in vertebrates [26]. The maintenance of several polymorphisms on the variable loops that are responsible for the control of antiviral specificity suggests that different selective pressures may be acting on the PRYSPRY domain of different Lagomorpha species.

To the best of our knowledge, the observed variation within the PRYSPRY domain between closely related species of the same genus was never reported before. Primates have divergence times similar to those of lagomorphs [18, 34]. However, when considering the PRYSPRY domain, we did not find high variations in the amino acid sequences within the different Primates genera (S2 Appendix). Both the *Sylvilagus* and *Macaca* genera diversification began at



around the same divergence time (~5 Mya) [18, 34]. However, in contradiction to what was observed for *Sylvilagus* spp., the six *Macaca* species did not present marked amino acid variation in the PRYSPRY domain, with the only difference between species being a two amino acid deletion in *M. fascicularis*, *M. nigra* and *M. thibetana*. The same comparison can also be made between *Ochotona* and *Saguinus* genera, since the diversification of both began at ~8 Mya [17, 34]. In contradiction to what was observed in the Daurian pika and the steppe pika, none of the six *Saguinus* species displayed marked differences in the amino acid sequence of the PRYSPRY domain, especially considering the v1 loop.

So, what is driving the evolution of PRYSPRY domain in Lagomorpha? Initial studies suggested that leporid retroviruses like RELIK, the first reported endogenous lentivirus ever [43], may have imposed positive selection on TRIM5 orthologs. RELIK was first identified in the genome of the European rabbit and, subsequently, in the genome of other leporid genera, including *Lepus*, *Sylvilagus* and *Bunolagus*, which places its origins around 12 Mya [20]. Indeed, the role of endogenous retroviruses, such as RELIK, in driving the evolution of TRIM5 might be determinant, for example, in *Sylvilagus* species; yet, it is apparently less important for the evolution of *Lepus* TRIM5, since the nine *Lepus* species exhibited a great conservation of the PRYSPRY domain (Fig 1).

Retroviruses have been infecting vertebrate hosts for millions of years by integrating their proviral DNA copies as permanent insertions of host genomes. After becoming fixed in a population, endogenous retroviruses (ERVs) can be used as “fossils”, providing a remarkable record of virus-host interactions [44, 45]. In the past years, taking advantage of the European rabbit (oryCun2.0) and American pika (ochPri3.0) genomes assembly, several efforts have been made to detect the presence of other retroviruses, rather than RELIK, in lagomorph's genomes. For example, a study focused on retroviral diversity across wide samples of vertebrates showed that in European rabbit and American pika the ERVs abundance is dominated by two major groups: Gamma-like ERVs and Beta-like ERVs [46, 47]. A Pika-BERV (pika endogenous betaretrovirus) with an endogenization event calculated ~3–6 Mya is, interestingly, only present in a few *Ochotona* species, including Hoffmann's pika, Manchurian pika, Alpine pika, Turuchan pika and American pika [48], all part of the alpine group that diverged around 3–6 Mya from the remaining pika groups [17]. In fact, the presence of Pika-BERV in only five *Ochotona* species, representative of a 3–6 million years divergence, reinforces our hypothesis that a primitive retroviral infection could have shaped the PRYSPRY domain of the Daurian pika and the steppe pika. However, taking in consideration that a large number of mammalian retrovirus remain to be identified [46, 47] and that maybe past retroviruses were not endogenized in the Lagomorpha genomes, it is impossible to predict what shaped the TRIM5 evolution in some Lagomorpha species.

Overall, our results fuel the hypothesis that the Lagomorpha TRIM5 evolution might have been impacted by different unknown ancient retroviruses, either endogenous or exogenous. With this work, we demonstrate the multitude of evolutionary independent episodes of TRIM5 variation in different Lagomorpha genera, through the exhibition of species-specific length and sequence variation in the PRYSPRY domain. Such findings were not observed in TRIM5 Primates genera, supporting the uniqueness of TRIM5 Lagomorpha evolution at the species level.

## Materials and methods

### Lagomorpha genomic DNA (gDNA) sources and tissue extraction

Tissues from *Lepus* and *Ochotona* specimens were used to extract genomic DNA (gDNA). Snowshoe hare (*Lepus americanus*), black-tailed jackrabbit (*L. californicus*), Cape hare (*L. capensis*), broom hare (*L. castroviejoii*), Corsican hare (*L. corsicanus*), mountain hare (*L.*

*timidus*) and white-tailed jackrabbit (*L. townsendii*) samples were supplied by CIBIO/InBIO, Vairão, Portugal. Samples from alpine pika (*Ochotona alpina*), Daurian pika (*O. dauurica*), Hoffmann's pika (*O. hoffmanni*), northern pika (*O. hyperborea*), Manchurian pika (*O. mantchurica*), Palla's pika (*O. pallasi*), steppe pika (*O. pusilla*), Turkestan red pika (*O. rutila*) and Turuchan pika (*O. turuchanensis*) were kindly provided by the Zoological Museum of Moscow State Lomonosov University, Moscow, Russia. gDNA was extracted using the E.Z.N.A.® Tissue DNA Kit (Omega Bio-Tek, Norcross, GA, USA) according to manufacturer's instructions.

### Lagomorpha TRIM5 PRYSPRY domain amplification and sequencing

Primers were designed according the TRIM5 gene from *Oryctolagus cuniculus* chromosome 1 [GenBank: NC\_013669] (Forward 5' -CAAATTCATGAGCTGAAAAGGA-3' and Reverse 5' AAGAGATGTACCCAGGGTAAGAG-3'), and from *Ochotona princeps* unplaced genomic scaffold00040 [GenBank: NW\_004535475] (Forward 5' -CAGAGGAAACCATTTGAAGCT-3' and Reverse 5' -CTAGCAAAGCGTCATGGGT-3'). The primers designed according the rabbit TRIM5 were used in hare samples whereas the pika based primers were used in pika samples. The approximately 1.2Kb PCR product corresponds to all length of exon 6 and exon 7 (covering the entire PRYSPRY region).

Phusion® High-Fidelity DNA Polymerase (Finnzymes, Espoo, Finland) was used in the PCR amplification, the conditions included an initial denaturation (98°C for 3min), 40 cycles of denaturation (98°C for 30s), annealing (60°C for 20s) and extension (72°C for 45s) and a final extension (72°C for 5min) for both hare and pika DNA. Amplicons sequencing was performed with the ABI PRISM BigDye Terminator v3.1 Cycle Sequencing Kit and according to manufacturer's protocol; reactions were cleaned with Sephadex™ (GE Healthcare Life Sciences, UK) and applied on an ABI PRISM 310 Genetic Analyser (Life Technologies, Applied Biosystems, Carlsbad, CA, USA). PCR products were sequenced in both directions and, particularly for pika samples, an internal primer was also used (5' - ACTGGGAGGTGGATGTGTCT-3').

Virtual transcripts of *Lepus* and *Ochotona* TRIM5 PRYSPRY domains were created by splicing together the exons reads (~ 600 bp) and have been deposited in the GenBank database under the following accession numbers: #MN605824 (*Ochotona hoffmanni*), #MN605825 (*Ochotona hyperborean*), #MN605826 (*Ochotona mantchurica*), #MN605827 (*Ochotona pallasi*), #MN605828 (*Ochotona rutila*), #MN605829 (*Ochotona turuchanensis*), #MN605830 (*Ochotona alpina*), #MN605831 (*Ochotona pusilla*), #MN605832 (*Ochotona dauurica*), #MN605833 (*Lepus capensis*), #MN605834 (*Lepus castroviejoi*), #MN605835 (*Lepus corsicanus*), #MN605836 (*Lepus timidus*), #MN605837 (*Lepus townsendii*), #MN605838 (*Lepus americanus*), #MN605839 (*Lepus californicus*). *Lepus* and *Ochotona* samples have been used in previous publications [49–52]

### Sequence and Phylogenetic analyses

To complete the Lagomorpha dataset, besides the PCR-amplified *Lepus* and *Ochotona* PRYSPRY domain sequences, the TRIM5 nucleotide sequences of other leporid genera, including *Oryctolagus* and *Sylvilagus*, were obtained from NCBI database (<http://www.ncbi.nlm.nih.gov>). The PRYSPRY domain nucleotide sequences of Lagomorpha species were aligned in BioEdit Sequence Alignment Editor [53] using Clustal W [54], followed by manual corrections when necessary. Maximum likelihood (ML) phylogenetic reconstruction of Lagomorpha PRYSPRY domains was performed using PhyML v3.0 [27]. TIM3+G was identified as the best-fitting nucleotide substitution model, according to the Akaike information criterion (AIC) implemented in jModelTest v2.1.1 [55]. The support of the resulting nodes was estimated using 1000 bootstrap replicates.

## Molecular evolutionary analyses

The nucleotide sequences alignment for TRIM5 PRYSPRY domain was firstly tested for recombination, as this biological process can mislead molecular evolutionary analyses [56]. Coding sequences were scanned for recombination by using six methods (GENECONV, BootScan, MaxChi, Chimaera, SiScan and 3Seq) available in the RDP software, version 4.95 [57]. Nevertheless, no significant breakpoints were identified in the alignments.

To look for signatures of natural selection operating in the alignment, we used codeml of the PAML v4.9 package [58] and compared site-based models to determine if a model that allows positive selection (alternative model, M8) is a better fit to the data than a neutral model (null model, M7). The analysis was run with an initial  $\omega$  ratio value of 1, and conducted with the F3×4 model of codon frequencies. Likelihood ratio test (LRT) was performed with two degrees of freedom to compare the fit of the two models by using the likelihood scores of the null neutral and alternative selection models. A Bayes empirical Bayes (BEB) approach was employed to detect codons with a posterior probability >95% of being under selection [59]. Five other methods, using HyPhy software implemented in the Datamonkey Web server [60, 61], were also applied to detect codons under selection: the Single Likelihood Ancestor Counting (SLAC) model, the Fixed Effect Likelihood (FEL) method, the Random Effect Likelihood (REL) method [62] and the recently described Mixed Effects Model of Evolution (MEME) [63] and Fast Unbiased Bayesian Approximation (FUBAR) [64] methods. To avoid a high false-positive rate [62], codons with p-values <0.1 for FEL and MEME models, Bayes Factor >50 for REL model and a posterior probability >0.90 for FUBAR were accepted as candidates for selection (S1 Appendix). For a more conservative approach, and as used previously [52, 65], only residues identified as being under positive selection in more than two ML methods were considered.

Sliding-window analysis for the Lagomorpha PRYSPRY domain was carried out using DnaSP v5.10 [66], for which a window length of 20 and a step size of 5 nucleotides were defined. The following non-synonymous substitutions per non-synonymous sites analyses were performed: *Sylvilagus bachmani* vs *Oryctolagus cuniculus*, *S. floridanus* vs *O. cuniculus* and *S. bachmani* vs *S. floridanus*; *O. dauurica* vs Ochotonidae family, *O. pusilla* vs Ochotonidae family and *O. dauurica* vs *O. pusilla*. A pairwise estimation of non-synonymous to synonymous substitution ratios (dN/dS) for the PRYSPRY v1 region of these same comparisons was performed on MEGA7 [33] using Nei-Gojobori (Jukes-Cantor correction) method.

## Primates PRYSPRY sequences

Using *Homo sapiens* TRIM5 $\alpha$  PRYSPRY domain as template, we performed a BLAST search in GenBank (NCBI, <http://BLAST.ncbi.nlm.nih.gov/>) and Ensembl (<http://www.ensembl.org/Multi/blastview>) databases to obtain all the available Primate TRIM5 $\alpha$  PRYSPRY sequences. The PRYSPRY sequences collected belong to Hominoidea, Cercopithecoidea and Platyrrhini groups. Sequences were aligned with Clustal W [54], implemented in BioEdit Sequence Alignment Editor [53], followed by visual inspection (S2 Appendix). The complete list of all Primate PRYSPRY sequences used in this study, together with accession numbers, is given in S3 Appendix.

## Supporting information

**S1 Appendix. Positive selection analyses for TRIM5 PRYSPRY domain of Lagomorphs.** (DOCX)

**S2 Appendix. Amino acid alignment of the TRIM5 PRYSPRY domain of Hominoidea, Cercopithecoidea and Platyrrhini species.** Amino acid sequences are grouped according to

species genus. Variable loop “v1” from PRYSPRY domain is represented (grey box).  
Dots = identity with the PRYSPRY sequence from *Homo sapiens*.  
(DOCX)

### S3 Appendix. List of the sequences of the TRIM5 PRYSPRY domain, available from NCBI and Ensemble databases and used in this study.

(DOCX)

## Author Contributions

**Conceptualization:** Ana Lemos de Matos, Pedro J. Esteves.

**Data curation:** Ana Águeda-Pinto.

**Formal analysis:** Ana Águeda-Pinto, Ana Lemos de Matos, Pedro J. Esteves.

**Funding acquisition:** Pedro J. Esteves.

**Investigation:** Ana Águeda-Pinto.

**Methodology:** Ana Águeda-Pinto, Ana Lemos de Matos, Ana Pinheiro, Fabiana Neves, Patrícia de Sousa-Pereira.

**Project administration:** Pedro J. Esteves.

**Resources:** Pedro J. Esteves.

**Supervision:** Ana Lemos de Matos, Pedro J. Esteves.

**Validation:** Ana Lemos de Matos, Ana Pinheiro, Fabiana Neves, Pedro J. Esteves.

**Visualization:** Ana Águeda-Pinto.

**Writing – original draft:** Ana Águeda-Pinto.

**Writing – review & editing:** Ana Águeda-Pinto, Ana Lemos de Matos, Ana Pinheiro, Fabiana Neves, Patrícia de Sousa-Pereira, Pedro J. Esteves.

## References

1. Duggal NK, Emerman M. Evolutionary conflicts between viruses and restriction factors shape immunity. *Nature Reviews Immunology*. 2012; 12:687. <https://doi.org/10.1038/nri3295> PMID: 22976433
2. Jern P, Coffin JM. Effects of Retroviruses on Host Genome Function. *Annual Review of Genetics*. 2008; 42(1):709–32. <https://doi.org/10.1146/annurev.genet.42.110807.091501> PMID: 18694346.
3. Chan YK, Gack MU. Viral evasion of intracellular DNA and RNA sensing. *Nature Reviews Microbiology*. 2016; 14:360. <https://doi.org/10.1038/nrmicro.2016.45> PMID: 27174148
4. Hall JC, Rosen A. Type I interferons: crucial participants in disease amplification in autoimmunity. *Nature Reviews Rheumatology*. 2010; 6:40. <https://doi.org/10.1038/nrrheum.2009.237> PMID: 20046205
5. Malim MH, Bieniasz PD. HIV restriction factors and mechanisms of evasion. *Cold Spring Harbor perspectives in medicine*. 2012; 2(5):a006940–a. <https://doi.org/10.1101/cshperspect.a006940> PMID: 22553496.
6. Hatzioannou T, Bieniasz PD. Antiretroviral restriction factors. *Current opinion in virology*. 2011; 1(6):526–32. <https://doi.org/10.1016/j.coviro.2011.10.007> PMID: 22278313.
7. Boso G, Buckler-White A, Kozak CA. Ancient Evolutionary Origin and Positive Selection of the Retroviral Restriction Factor *Fv1* in Muroid Rodents. *J Virol*. 2018; 92(18). <https://doi.org/10.1128/jvi.00850-18> PMID: 29976659
8. Pertel T, Hausmann S, Morger D, Züger S, Guerra J, Lascano J, et al. TRIM5 is an innate immune sensor for the retrovirus capsid lattice. *Nature*. 2011; 472(7343):361–5. <https://doi.org/10.1038/nature09976> PMID: 21512573.

9. Sawyer SL, Wu LI, Emerman M, Malik HS. Positive selection of primate *TRIM5α* identifies a critical species-specific retroviral restriction domain. *Proc Natl Acad Sci U S A*. 2005; 102(8):2832–7. <https://doi.org/10.1073/pnas.0409853102> PMID: 15689398
10. Han K, Lou DI, Sawyer SL. Identification of a genomic reservoir for new *TRIM* genes in primate genomes. *PLOS Genetics*. 2011; 7(12):e1002388. <https://doi.org/10.1371/journal.pgen.1002388> PMID: 22144910
11. Goldschmidt V, Ciuffi A, Ortiz M, Brawand D, Muñoz M, Kaessmann H, et al. Antiretroviral Activity of Ancestral *TRIM5α*. *J Virol*. 2008; 82(5):2089–96. <https://doi.org/10.1128/JVI.01828-07> PMID: 18077724
12. Sawyer SL, Emerman M, Malik HS. Discordant evolution of the adjacent antiretroviral genes *TRIM22* and *TRIM5* in Mammals. *PLOS Pathogens*. 2007; 3(12):e197. <https://doi.org/10.1371/journal.ppat.0030197> PMID: 18159944
13. Schaller T, Hué S, Towers GJ. An active *TRIM5* protein in rabbits indicates a common antiviral ancestor for mammalian *TRIM5* proteins. *J Virol*. 2007; 81(21):11713–21. <https://doi.org/10.1128/JVI.01468-07> PMID: 17728224
14. Si Z, Vandegraaff N, O'hUigin C, Song B, Yuan W, Xu C, et al. Evolution of a cytoplasmic tripartite motif (*TRIM*) protein in cows that restricts retroviral infection. *Proceedings of the National Academy of Sciences*. 2006; 103(19):7454–9. <https://doi.org/10.1073/pnas.0600771103> PMID: 16648259
15. Hron T, Farkašová H, Padhi A, Pačes J, Elleder D. Life history of the oldest Lentivirus: characterization of ELVg<sub>v</sub> integrations in the dermopteran genome. *Molecular Biology and Evolution*. 2016; 33(10):2659–69. <https://doi.org/10.1093/molbev/msw149> PMID: 27507840
16. de Matos AL, van der Loo W, Areal H, Lanning DK, Esteves PJ. Study of *Sylvilagus* rabbit *TRIM5α* species-specific domain: how ancient endoviruses could have shaped the antiviral repertoire in Lagomorpha. *BMC Evolutionary Biology*. 2011; 11(1):294. <https://doi.org/10.1186/1471-2148-11-294> PMID: 21982459
17. Melo-Ferreira J, de Matos AL, 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. <https://doi.org/10.1016/j.ympev.2015.01.004> WOS:000352173500021. PMID: 25637497
18. 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. <https://doi.org/10.1371/journal.pone.0059668> PMID: 23573205
19. Yap MW, Stoye JP. Apparent effect of rabbit endogenous lentivirus type K acquisition on retrovirus restriction by lagomorph *Trim5α*s. *Philosophical transactions of the Royal Society of London Series B, Biological sciences*. 368(1626):20120498–. <https://doi.org/10.1098/rstb.2012.0498> PMID: 23938750.
20. Keckesova Z, Ylänen LMJ, Towers GJ, Gifford RJ, Katzourakis A. Identification of a RELIK orthologue in the European hare (*Lepus europaeus*) reveals a minimum age of 12 million years for the lagomorph lentiviruses. *Virology*. 2009; 384(1):7–11. <https://doi.org/10.1016/j.virol.2008.10.045> PMID: 19070882
21. Fletcher AJ, Hue S, Schaller T, Pillay D, Towers GJ. Hare *TRIM5α* restricts divergent retroviruses and exhibits significant sequence variation from closely related Lagomorpha *TRIM5* genes. *J Virol*. 2010; 84(23):12463–8. <https://doi.org/10.1128/JVI.01514-10> WOS:000283799500037. PMID: 20861252
22. Yap MW, Stoye JP. Apparent effect of rabbit endogenous lentivirus type K acquisition on retrovirus restriction by lagomorph *Trim5α* s. *Philos Trans R Soc B-Biol Sci*. 2013; 368(1626):11. <https://doi.org/10.1098/rstb.2012.0498> WOS:000331222100004. PMID: 23938750
23. Song B, Javanbakht H, Perron M, Park DH, Stremlau M, Sodroski J. Retrovirus restriction by *TRIM5α* variants from Old World and New World Primates. *J Virol*. 2005; 79(7):3930–7. <https://doi.org/10.1128/JVI.79.7.3930-3937.2005> PMID: 15767395
24. Song B, Gold B, O'hUigin C, Javanbakht H, Li X, Stremlau M, et al. The B30.2 (SPRY) domain of the retroviral restriction factor *TRIM5α* exhibits lineage-specific length and sequence variation in primates. *J Virol*. 2005; 79(10):6111–21. <https://doi.org/10.1128/JVI.79.10.6111-6121.2005> PMID: 15857996
25. Soares EA, Menezes AN, Schrago CG, Moreira MAM, Bonvicino CR, Soares MA, et al. Evolution of *TRIM5α* B30.2 (SPRY) domain in New World primates. *Infection, Genetics and Evolution*. 2010; 10(2):246–53. <https://doi.org/10.1016/j.meegid.2009.11.012> PMID: 19931648
26. Yang H, Ji X, Zhao G, Ning J, Zhao Q, Aiken C, et al. Structural insight into HIV-1 capsid recognition by rhesus *TRIM5*2012.
27. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst Biol*. 2010; 59(3):307–21. <https://doi.org/10.1093/sysbio/syq010> WOS:000276528300006. PMID: 20525638

28. Wang H-Y, Tang H, Shen CKJ, Wu C-I. Rapidly evolving genes in human. I. The glycoporphins and their possible role in evading malaria parasites. *Molecular Biology and Evolution*. 2003; 20(11):1795–804. <https://doi.org/10.1093/molbev/msg185> PMID: 12949139
29. Kosiol C, Vinař T, da Fonseca RR, Hubisz MJ, Bustamante CD, Nielsen R, et al. Patterns of positive selection in six mammalian genomes. *PLOS Genetics*. 2008; 4(8):e1000144. <https://doi.org/10.1371/journal.pgen.1000144> PMID: 18670650
30. Aguileta G, Refrégier G, Yockteng R, Fournier E, Giraud T. Rapidly evolving genes in pathogens: Methods for detecting positive selection and examples among fungi, bacteria, viruses and protists. *Infection, Genetics and Evolution*. 2009; 9(4):656–70. <https://doi.org/10.1016/j.meegid.2009.03.010> PMID: 19442589
31. James LC, Keeble AH, Khan Z, Rhodes DA, Trowsdale J. Structural basis for PRYSPRY-mediated tripartite motif (TRIM) protein function. *Proc Natl Acad Sci U S A*. 2007; 104(15):6200–5. <https://doi.org/10.1073/pnas.0609174104> WOS:000245737500021. PMID: 17400754
32. Yap MW, Nisole S, Stoye JP. A single amino acid change in the SPRY domain of human TRIM5 alpha leads to HIV-1 restriction. *Curr Biol*. 2005; 15(1):73–8. <https://doi.org/10.1016/j.cub.2004.12.042> WOS:000226715000029. PMID: 15649369
33. Kumar S, Stecher G, Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*. 2016; 33(7):1870–4. <https://doi.org/10.1093/molbev/msw054> PMID: 27004904
34. Perelman P, Johnson WE, Roos C, Seuánez HN, Horvath JE, Moreira MAM, et al. A molecular phylogeny of living primates. *PLOS Genetics*. 2011; 7(3):e1001342. <https://doi.org/10.1371/journal.pgen.1001342> PMID: 21436896
35. Diehl WE, Johnson WE, Hunter E. Elevated rate of fixation of endogenous retroviral elements in Haplorhini TRIM5 and TRIM22 genomic sequences: impact on transcriptional regulation. *PLoS one*. 2013; 8(3):e58532–e. <https://doi.org/10.1371/journal.pone.0058532> PMID: 23516500
36. Stremmler M, Perron M, Lee M, Li Y, Song B, Javanbakht H, et al. Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5 $\alpha$  restriction factor. *Proceedings of the National Academy of Sciences*. 2006; 103(14):5514–9. <https://doi.org/10.1073/pnas.0509996103> PMID: 16540544
37. McCarthy KR, Kirmaier A, Autissier P, Johnson WE. Evolutionary and functional analysis of old world primate TRIM5 reveals the ancient emergence of primate lentiviruses and convergent evolution targeting a conserved capsid interface. *PLoS pathogens*. 2015; 11(8):e1005085–e. <https://doi.org/10.1371/journal.ppat.1005085> PMID: 26291613
38. Kratovac Z, Virgen CA, Bibollet-Ruche F, Hahn BH, Bieniasz PD, Hatziioannou T. Primate lentivirus capsid sensitivity to TRIM5 proteins. *J Virol*. 2008; 82(13):6772–7. <https://doi.org/10.1128/JVI.00410-08> PMID: 18417575
39. Johnson WE, Sawyer SL. Molecular evolution of the antiretroviral *TRIM5* gene. *Immunogenetics*. 2009; 61(3):163–76. <https://doi.org/10.1007/s00251-009-0358-y> PMID: 19238338
40. Li Y-L, Chandrasekaran V, Carter SD, Woodward CL, Christensen DE, Dryden KA, et al. Primate TRIM5 proteins form hexagonal nets on HIV-1 capsids. *eLife*. 2016; 5:e16269. <https://doi.org/10.7554/eLife.16269> PMID: 27253068
41. Javanbakht H, An P, Gold B, Petersen DC, O'Huigin C, Nelson GW, et al. Effects of human TRIM5 $\alpha$  polymorphisms on antiretroviral function and susceptibility to human immunodeficiency virus infection. *Virology*. 2006; 354(1):15–27. <https://doi.org/10.1016/j.virol.2006.06.031> PMID: 16887163
42. Kovalskyy DB, Ivanov DN. Recognition of the hiv capsid by the TRIM5 alpha restriction factor is mediated by a subset of pre-existing conformations of the TRIM5 alpha SPRY domain. *Biochemistry*. 2014; 53(9):1466–76. <https://doi.org/10.1021/bi4014962> WOS:000332913600009. PMID: 24506064
43. Katzourakis A, Tristem M, Pybus OG, Gifford RJ. Discovery and analysis of the first endogenous lentivirus. *Proceedings of the National Academy of Sciences*. 2007; 104(15):6261–5. <https://doi.org/10.1073/pnas.0700471104> PMID: 17384150
44. Chuong EB, Elde NC, Feschotte C. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science*. 2016; 351(6277):1083–7. <https://doi.org/10.1126/science.aad5497> PMID: 26941318
45. Patel MR, Emerman M, Malik HS. Paleovirology—ghosts and gifts of viruses past. *Current Opinion in Virology*. 2011; 1(4):304–9. <https://doi.org/10.1016/j.coviro.2011.06.007> PMID: 22003379
46. Rivas-Carrillo SD, Pettersson ME, Rubin C-J, Jern P. Whole-genome comparison of endogenous retrovirus segregation across wild and domestic host species populations. *Proceedings of the National Academy of Sciences*. 2018; 115(43):11012–7. <https://doi.org/10.1073/pnas.1815056115> PMID: 30297425

47. Hayward A, Cornwallis CK, Jern P. Pan-vertebrate comparative genomics unmasks retrovirus macro-evolution. *Proceedings of the National Academy of Sciences*. 2015; 112(2):464–9. <https://doi.org/10.1073/pnas.1414980112> PMID: 25535393
48. de Matos AL, de Sousa-Pereira P, Lissovsky AA, van der Loo W, Melo-Ferreira J, Cui J, et al. Endogenization of mouse mammary tumor virus (MMTV)-like elements in genomes of pikas (*Ochotona* sp.). *Virus Res*. 2015; 210:22–6. <https://doi.org/10.1016/j.virusres.2015.06.021> WOS:000365455500004. PMID: 26151606
49. Neves F, Abrantes J, Esteves PJ. Evolution of CCL11: genetic characterization in lagomorphs and evidence of positive and purifying selection in mammals. *Innate Immunity*. 2016; 22(5):336–43. <https://doi.org/10.1177/1753425916647471> PMID: 27189425.
50. de Sousa-Pereira P, Abrantes J, Baldauf H-M, Keppler OT, Esteves PJ. Evolutionary study of leporid CD4 reveals a hotspot of genetic variability within the D2 domain. *Immunogenetics*. 2016; 68(6):477–82. <https://doi.org/10.1007/s00251-016-0909-y> PMID: 26979977
51. Pinheiro A, Woof JM, Almeida T, Abrantes J, Alves PC, Gortázar C, et al. Leporid immunoglobulin G shows evidence of strong selective pressure on the hinge and CH3 domains. *Open biology*. 2014; 4(9):140088. <https://doi.org/10.1098/rsob.140088> PMID: 25185680
52. de Matos AL, McFadden G, Esteves PJ. Evolution of viral sensing RIG-I-like receptor genes in Leporidae genera *Oryctolagus*, *Sylvilagus*, and *Lepus*. *Immunogenetics*. 2014; 66(1):43–52. <https://doi.org/10.1007/s00251-013-0740-7> WOS:000329096900006. PMID: 24220721
53. Hall TA, editor BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic acids symposium series*; 1999: [London]: Information Retrieval Ltd., c1979–c2000.
54. 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. <https://doi.org/10.1093/nar/22.24.4673> WOS:A1994PU19900018. PMID: 7984417
55. Posada D. jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution*. 2008; 25(7):1253–6. <https://doi.org/10.1093/molbev/msn083> PMID: 18397919
56. Posada D, Crandall KA. The effect of recombination on the accuracy of phylogeny estimation. *Journal of Molecular Evolution*. 2002; 54(3):396–402. <https://doi.org/10.1007/s00239-001-0034-9> PMID: 11847565
57. Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefevre P. RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics*. 2010; 26(19):2462–+. <https://doi.org/10.1093/bioinformatics/btq467> WOS:000282170000017. PMID: 20798170
58. Yang Z. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and Evolution*. 2007; 24(8):1586–91. <https://doi.org/10.1093/molbev/msm088> PMID: 17483113
59. Yang Z, Wong WSW, Nielsen R. Bayes Empirical Bayes inference of amino acid sites under positive selection. *Molecular Biology and Evolution*. 2005; 22(4):1107–18. <https://doi.org/10.1093/molbev/msi097> PMID: 15689528
60. Pond SLK, Frost SDW. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics*. 2005; 21(10):2531–3. <https://doi.org/10.1093/bioinformatics/bti320> PMID: 15713735
61. Delpont W, Poon AFY, Frost SDW, Kosakovsky Pond SL. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*. 2010; 26(19):2455–7. <https://doi.org/10.1093/bioinformatics/btq429> PMID: 20671151
62. Kosakovsky Pond SL, Frost SDW. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution*. 2005; 22(5):1208–22. <https://doi.org/10.1093/molbev/msi105> PMID: 15703242
63. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. Detecting individual sites subject to episodic diversifying selection. *PLOS Genetics*. 2012; 8(7):e1002764. <https://doi.org/10.1371/journal.pgen.1002764> PMID: 22807683
64. Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, et al. FUBAR: a fast, unconstrained Bayesian Approximation for inferring selection. *Molecular Biology and Evolution*. 2013; 30(5):1196–205. <https://doi.org/10.1093/molbev/mst030> PMID: 23420840
65. de Matos AL, Liu J, McFadden G, Esteves PJ. Evolution and divergence of the mammalian *SAMD9/SAMD9L* gene family. *Bmc Evolutionary Biology*. 2013; 13:16. <https://doi.org/10.1186/1471-2148-13-16> WOS:000320538300001.
66. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. 2009; 25(11):1451–2. <https://doi.org/10.1093/bioinformatics/btp187> PMID: 19346325