



Article

Comprehensive Analysis of Rodent-Specific *Probasin* Gene Reveals Its Evolutionary Origin in Pseudoautosomal Region and Provides Novel Insights into Rodent Phylogeny

Stephan Maxeiner ^{1,2,*} , Lukas Walter ^{1,†}, Samuel Luca Zeitz ^{1,†} and Gabriela Krasteva-Christ ^{1,2} 

¹ Anatomy and Cell Biology, Saarland University, Kirrbergerstr. 100, Bldg. 61, 66424 Homburg, Germany; gabriela.krasteva-christ@uks.eu (G.K.-C.)

² Center for Gender-Specific Biology and Medicine (CGMB), Saarland University, 66424 Homburg, Germany

* Correspondence: stephan.maxeiner@uni-saarland.de

† These authors contributed equally to this work.

Simple Summary: The house mouse, *Mus musculus*, is the preeminent rodent employed in basic biomedical research to understand the systemic functions of individual genes in vivo. Genes involved in human pathologies regularly find their counterparts in the mouse genome; however, little is known about genes exclusive to primates or rodents. In this study, we aimed at finding the evolutionary origin of the *probasin* gene, which is only described in mice and rats and is absent in humans. We traced its origin back to when ancestors of mouse-like rodents such as hamsters, mice, and moles separated, consequently becoming absent outside the mammal order of rodents. *Probasin* initially emerged as a gene present on both sex chromosomes in a specialized region that behaves like autosomes, hence termed pseudoautosomal. During the evolution of mouse-like rodents, it recombined to the X-specific region on the X-chromosome. Our study highlights this transition, suggesting that *probasin* has remained, at present, largely conserved in mouse-like species. Close analysis of sequence variations in the *probasin* gene cause us to challenge the current understanding about the phylogenetic relationship among some tribes within the family of mice.



Academic Editor: Edward L. Braun

Received: 9 December 2024

Revised: 20 February 2025

Accepted: 22 February 2025

Published: 27 February 2025

Citation: Maxeiner, S.; Walter, L.; Zeitz, S.L.; Krasteva-Christ, G. Comprehensive Analysis of Rodent-Specific *Probasin* Gene Reveals Its Evolutionary Origin in Pseudoautosomal Region and Provides Novel Insights into Rodent Phylogeny. *Biology* **2025**, *14*, 239. <https://doi.org/10.3390/biology14030239>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: *Probasin* protein was originally identified as a basic protein present in rat prostate epithelium. So far, its physiological role, its origin, and its presence in other species including humans remain largely elusive. With the ever-growing number of genome assemblies, thus far, *probasin* genes (*Pbsn/PBSN*) have only been predicted in a subset of rodent genomes. In this study, we addressed the phylogeny of *probasin* genes and found them to be exclusively present in members of the superfamily Muroidea. It first emerged in the so-called pseudoautosomal region, a subtelomeric gene cluster of both mammalian sex chromosomes. During evolution of the Muroidea lineages, *probasin* recombined to the X-specific region of the X-chromosome in mice and hamster species. This event likely saved the gene from events that other pseudoautosomal genes suffered, namely displaying an increase in G and C nucleotide composition or accumulation of repetitive elements. We observed changes to its coding region, e.g., sequence insertions in exon 6, which challenge the current understanding of rodent phylogeny, in particular regarding the evolutionary history of tribe formation within the subfamily Murinae. Analyzing the evolution of *probasin* genes in Muroidea allows fostering understanding of phylogenetic relationships in one of the largest groups of mammalian species.

Keywords: pseudoautosomal region; rodents; Cricetidae; Muridae; Spalacidae; lipocalin; Muroidea; Eumuroidea

1. Introduction

Sex chromosomes, X- and Y-chromosomes, in Eutherian mammals display a subtelomeric region that behaves like autosomes, e.g., by allowing cross-over events during male meiosis, and is therefore named the ‘pseudautosomal region’ or PAR [1]. Recently, we demonstrated that in mouse-like rodents (superfamily Muroidea), the arrangement of genes within the PAR has undergone dramatic changes, whereas among most mammals, it is largely conserved [2]. The PAR has shrunk in size, accumulated repetitive sequences, and displayed an increase in its overall G and C nucleotide content. Some genes have even translocated to autosomes or been lost [2–4]. This loss of genes is intriguing since many of them have been demonstrated to be clinically significant to humans; among these are, e.g., *SHOX* (SHOX homeobox) and *ANOS1* (anosmin 1), responsible for the phenotypes of Turner syndrome or Kallmann syndrome, respectively [2].

Surprisingly, even though both the human and mouse genomes were sequenced and published over two decades ago, a comprehensive analysis of genes present in either the human or the mouse genome has not yet been conducted [5,6]. Current biomedical research highly depends on small rodents; in particular, the accessibility of the mouse genome for genetic manipulation is the avenue to take to understand the role of genes and their products in a systemic *in vivo* context. Following the fate of PAR genes during rodent evolution, we also observed that while there has been a loss of some genes, novel genes have also appeared. These genes require to be studied given that rodent species roughly comprise about 40% of all mammal species [7].

Recently, we came across *probasin* in a mouse (*Mus musculus*) in *Pbsn*, a gene in the vicinity of *Prkx* (protein kinase cAMP-dependent X-linked catalytic subunit) and *Tbl1x* (transducing beta-like 1 X-linked) localized on the X-chromosome. Both genes are frequently found in or close to the PAR in many mammals [8]. Presumptive ‘*probasin*’ or ‘*probasin-like*’ genes have only been identified in rodent genomes (cf. NCBI databases). Probasin was initially purified as an androgen-dependent basic protein expressed in rat prostate epithelium almost four decades ago and later identified in mice [8–10]. It displays a structural resemblance to lipocalins, a protein family present in all kingdoms of life, which include major urinary proteins (MUPs) and odorant-binding proteins (OBPs) in vertebrates; both families are responsible for shuttle lipophilic biomolecules [9,11,12]. Thus far, *Pbsn* knockout in mice has not been generated, leaving the physiological role of probasin protein largely unaddressed. However, its promoter region has been studied thoroughly in rats [9]. For example, androgen-responsive elements of the rat *probasin* promoter have been employed to generate transgenic mice which express oncogenes to study tumor formation in prostate cancer mouse models [13,14]. In a similar approach, transgenic mice expressing Cre-recombinase under the control of a rat *probasin* minimal promoter resulted in *lacZ* reporter gene expression postnatally and largely limited to the prostate epithelium [15].

In this study, we analyzed over sixty rodent genomes within the superfamily Muroidea in which *probasin* genes have been identified. We characterized its respective gene size, base content, and phylogeny. Our results confirm its close relationship with the family of lipocalins and MUPs/OBPs. Additionally, we determined differences in *Pbsn*/*PBSN* sequences, challenging the established relationship of distinct tribes within the subfamily Murinae.

2. Materials and Methods

2.1. Bioinformatical Analyses

Sequence analysis was essentially performed as has been reported recently [16]. Initially, the sequence information from the house mouse (Gene ID: 54192, NCBI, Bethesda,

MD, USA) was used to fetch homologous sequences from rodent genomes available from rodent assemblies curated by Genbank (NCBI) by submitting this sequence to the Blastn suite (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome (accessed on 23 November 2024)) with default settings and either selecting ‘blast’ or ‘megablast’ search options. All sequences were compared using the MultAlin tool (<http://multalin.toulouse.inra.fr/multalin/> (accessed on 23 November 2024)) [16]. Once the nucleotide sequence was retrieved for a given species, the exons were identified by their framing splice acceptor and donor sites; the coding region up to the presumptive poly-adenylation signal was isolated and submitted to the ‘Translate tool’ (<https://web.expasy.org/translate/> (accessed on 23 November 2024)) to generate the corresponding protein sequence. The newly assembled sequences were used alternatively for further blastn queries. The respective gene location of deduced sequences representing the presumptive start (ATG) of each coding sequence up to the poly-adenylation motif (AATAAA) and the protein sequences can be found in the Supplementary Data File. The genomic region of 61 *PBSN* genes could be fully analyzed (Table 1) regarding exon/intron structure and coding sequences (see Supplementary Data Sheet, section ‘gene structure & comments’), consequently 61 coding sequences could be assessed for base composition (cf. section ‘GC & GC3 content’) and translated into amino acid sequences. Only 42 of the 61 analyzed genes displayed no sequence disruption, hence allowing us to determine the full gene size. The content of the bases C and G either in the full-length coding sequence or in the third position (GC3) was assessed using the GC Content Calculator (https://jamiemcgowan.ie/bioinf/gc_content.html (accessed on 23 November 2024)) and Codon Usage Calculator (https://jamiemcgowan.ie/bioinf/codon_usage.html (accessed on 23 November 2024)) by Jamie McGowan (<https://jamiemcgowan.ie/> (accessed on 23 November 2024)). The results can be found in the Supplementary Data Sheet.

Table 1. Overview of analyzed genomes available from NCBI databases.

Family	Species
Cricetidae	<i>Alexandromys oeconomus</i> ¹ , <i>Arvicola amphibius</i> ² , <i>Cricetulus griseus</i> ² , <i>Dicrostonyx torquatus</i> ¹ , <i>Mesocricetus auratus</i> ¹ , <i>Microtus agrestis</i> ¹ , <i>M. arvalis</i> ¹ , <i>M. fortis</i> ² , <i>M. ochrogaster</i> ¹ , <i>M. oregoni</i> ² , <i>M. richardsoni</i> ¹ , <i>Myodes glareolus</i> ² , <i>Ondatra zibethicus</i> ¹ , <i>Onychomys torridus</i> ² , <i>Peromyscus attwateri</i> ¹ , <i>P. aztecus</i> ¹ , <i>P. californicus insignis</i> ² , <i>P. emericus</i> ² , <i>P. leucopus</i> ² , <i>P. maniculatus sonoriensis</i> ² , <i>P. nudipes</i> ¹ , <i>P. polionotus subgriseus</i> ¹ , <i>Phodopus roborovskii</i> ² , <i>P. sungorus</i> ¹ , <i>Phyllotis vaccarum</i> ¹ , <i>Sigmodon hispidus</i> ¹
Muridae	<i>Acomys cahirinus</i> ¹ , <i>A. dimidiatus</i> ¹ , <i>A. percivali</i> ¹ , <i>A. russatus</i> ² , <i>Apodemus agrarius</i> ¹ , <i>A. speciosus</i> ¹ , <i>A. sylvaticus</i> ² , <i>Arvicanthis niloticus</i> ² , <i>Grammomys surdaster</i> ² , <i>Hylomyscus alleni</i> ¹ , <i>Lophiomyys imhaus</i> ¹ , <i>Mastomys coucha</i> ² , <i>M. natalensis</i> ¹ , <i>Meriones psammophilus</i> ¹ , <i>M. unguiculatus</i> ² , <i>Mus minutoides</i> ¹ , <i>M. musculus</i> ^{2,3} , <i>M. musculus castaneus</i> ¹ , <i>M. pahari</i> ² , <i>M. spicilegus</i> ¹ , <i>M. spretus</i> ¹ , <i>Pachyuromys duprasi</i> ¹ , <i>Praomys delectorum</i> ¹ , <i>Psammomys obesus</i> ² , <i>Rattus norvegicus</i> ² , <i>R. rattus</i> ² , <i>Rhabdomys dilectus</i> ¹ , <i>R. pumilio</i> ¹ , <i>Rhombomys opimus</i> ¹ , <i>Rhynchomys soricooides</i> ¹ , <i>Uromys caudimaculus</i> ¹
Nesomyidae	<i>Cricetomys ansorgei</i> ¹ , <i>C. gambianus</i> ¹
Spalacidae	<i>Nannospalax galili</i> ² , <i>Rhizomys pruinosus</i> ¹

¹ Genomes without *PBSN* gene annotation. ² Genomes with *PBSN* or *PBSN*-like gene annotation. ³ Tissue processing for cDNA analysis. All species are presented in alphabetical order in each family section.

Once full nucleotide sequences of the coding regions were identified, they were aligned using Clustal Omega (<https://www.ebi.ac.uk/jdispatcher/msa/clustalo> (accessed on 23 November 2024)) in the Pearson/FASTA output format. The alignment in the FASTA format was saved and processed using Mega X [17,18]. Sequences were aligned, they were saved in the alignment format, and a phylogenetic tree was produced with default settings

using the Construct/Test Neighbor-Joining Tree option. To display the relationship and evolutionary branching points of individual rodent species, a text file listing all analyzed species was uploaded using the Time Tree option of Mega X (Figure S1). To assess genomic changes during the evolution of *probasin*, information available on annotated genomes was retrieved from the NCBI (<https://www.ncbi.nlm.nih.gov/gene/?term=probasin> (accessed on 23 November 2024)). Information regarding the structure of mouse probasin protein was derived from its entry in the AlphaFold protein structure database (<https://alphafold.ebi.ac.uk/entry/O08976> (accessed on 23 November 2024)). In an alignment of selected probasin protein sequences from different Muroidea lineages, the presence of β -sheets and α -helices were highlighted using their respective positions in the mouse. Furthermore, presumptive signal peptides of protein sequences were determined using SignalP-5.0 (<https://services.healthtech.dtu.dk/services/SignalP-5.0/> (accessed on 23 November 2024)) [19].

2.2. RT-PCR Analysis

Prostate tissue was salvaged from one male mouse which had been euthanized by an overdose of isoflurane inhalation (Piramal Critical Care Deutschland GmbH, Hallbergmoos, Germany) followed by cervical dislocation under a license regarding animal procedures for organ collection (see Institutional Review Board Statement) to meet the 3R principles. Prostate tissue was collected and homogenized in TRI Reagent[®] (Zymo Research Europe GmbH, Freiburg, Germany). Subsequently, the tissue was homogenized by employing a FastPrep-24 5G bead-beating lysis system (MP Biomedicals Germany GmbH, Eschwege, Germany). RNA was isolated using the Direct-zol[™] RNA MiniPrep Plus Kit according to the manufacturer's protocol (Zymo Research Europe GmbH), and RNA concentration was determined using a NanoDrop One spectrophotometer (Thermo Scientific, Waltham, MA, USA). We diluted 1 μ g of total RNA to a total volume of 8 μ L with RNase/DNase-free water, which was further subject to cDNA synthesis essentially as has been reported recently [20]. Briefly, the RNA was treated with DNase to remove traces of DNA potentially co-purified during the RNA extraction. The DNase-treated RNA was further transcribed into cDNA using the SuperScript II reverse transcription kit (Thermo Scientific). The cDNA was diluted in a 20 μ L of RNase/DNase-free water and PCR-reactions were set up using the Q5 High-Fidelity 2 \times Master Mix according to the manufacturer's protocol (New England Biolabs GmbH, Frankfurt, Germany). Oligonucleotide binding sites were identified by employing IDT's Realtime PCR Tool purchased from IDT (Coralville, IA, USA): KC2358 ACGAGTGGCTGGAGTTTTG; KC2359 GAAAATTCCTGATGTCATCTG; SMrt43 GAATAAGGAGGAGATGACGGAG; and SMrt44 GTTGAGTCACTAAGGTTTGATCTTG. PCRs were loaded on a T100 Thermal Cycler (Bio-Rad Laboratories, München, Germany). The PCR temperature settings were as follows: initial denaturation at 98 °C (30 s), 35 cycles at 98 °C (10 s), 60 °C (20 s), 72 °C (30 s), and final extension at 72 °C (1 min). The PCR products were separated on 2% agarose gel supplemented with MIDORI Green (1:10,000; Nippon Genetics Europe GmbH, Düren, Germany) and documented using the ChemiDoc XRS+ System (Bio-Rad Laboratories). Animal research followed the ARRIVE guidelines (where applicable), and animal handling followed the German guidelines for the care and use of laboratory animals approved by local authorities.

2.3. Statistical Analyses

Statistical significance was only assessed for the families Muridae and Cricetidae since their sample sizes exceeded $n > 3$. A normal distribution of data was confirmed by applying the Kolmogorov–Smirnov analysis before a paired Student's *t*-test was applied. *p*-values below 0.05 were considered significant. Statistical analysis was performed using GraphPad Prism 9 software (GraphPad Software, Boston, MA, USA).

3. Results

3.1. PBSN Coding Sequences Vary Among Muroidea Families

Thus far, probasin proteins and their underlying genes have only been studied in rats and mice due to their absence in other mammals including humans. In our phylogenetic analysis, we used the mouse *Pbsn* gene and its coding sequence as a reference to investigate its potential origin within the rodent lineage. Therefore, we employed a strategy that we most recently described in detail [16]. At the beginning of our study, the search for *probasin* or *probasin*-like genes retrieved only 36 hits in the NCBI/gene database. We extended this search by submitting the mouse or rat probasin coding sequences to the NCBI/blastn, which resulted in the retrieval of 61 full-length probasin coding sequences. The applied strategy is depicted as a flowchart in Figure 1. The mouse *Pbsn* gene (Gene ID: 54192), or *PBSN* in other rodent species, covered approximately 15.5 kb (distance covering the start codon to the poly-adenylation motif) and consists of seven exons (Figure 2A) [9]. The coding region started on exon 1 and terminated at a stop codon on exon 6, consequently leaving only the 3' untranslated region, 3'UTR, on exon 7. Similarity searches in the Muroidea family Platacanthomyidae as well as in other rodent families were negative [21]. When we compared the exon/intron structure of all the different rodent species, we noticed that the coding region in rats (family: Muridae), *Peromyscus* spec. (Cricetidae), and *Cricetomys* spec. (Nesomyidae) terminated at the beginning of exon 7 and that their exon 6 was consistently 24 nucleotides shorter than that in *M. musculus* (Figure 2B).

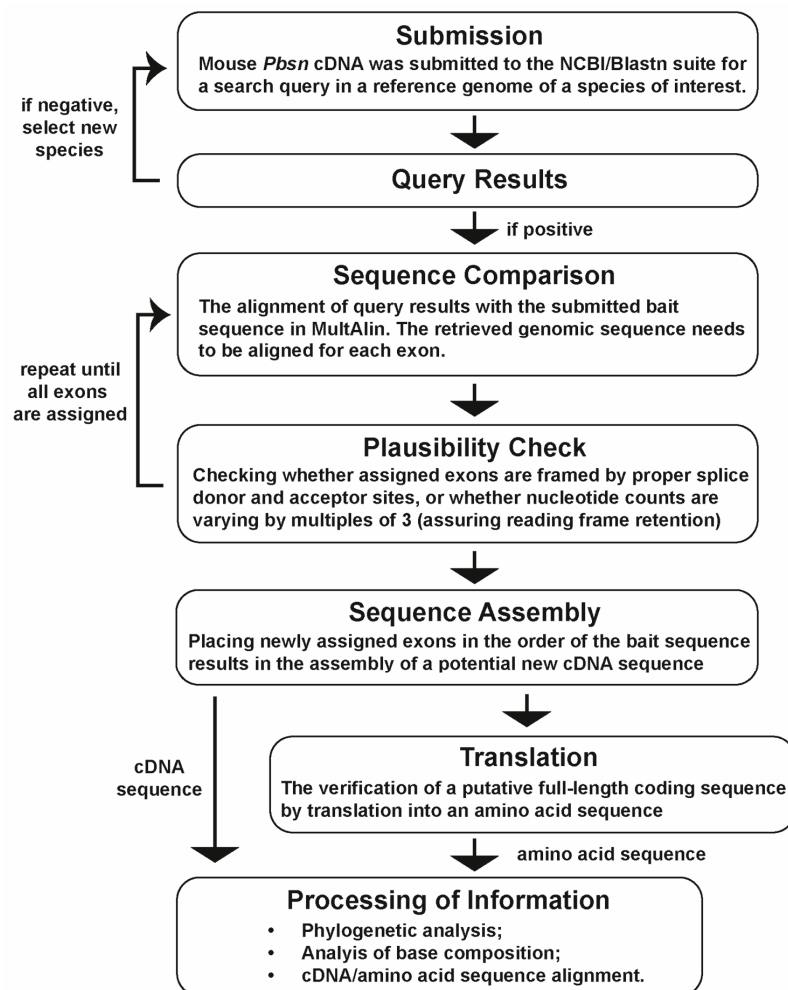


Figure 1. A flowchart depicting each step of the process of identifying novel, yet undescribed, probasin coding sequences.

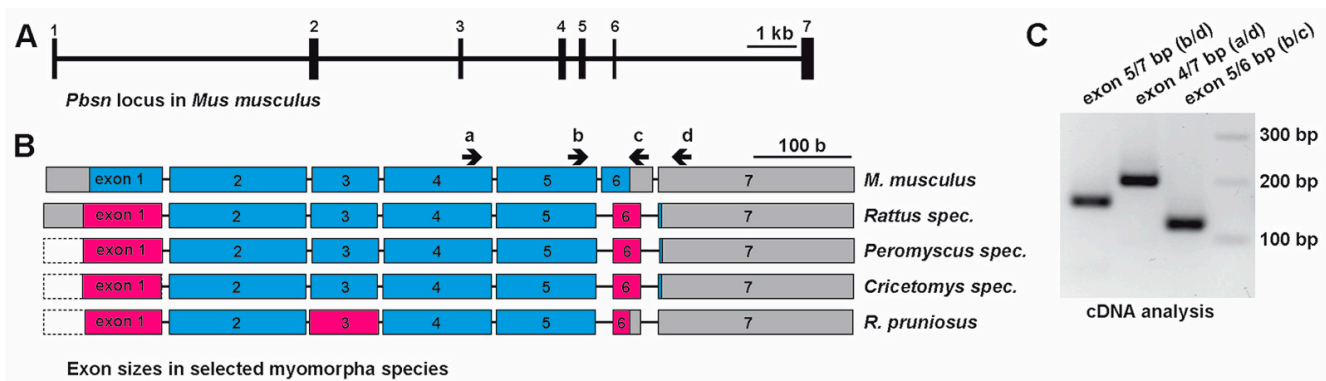


Figure 2. Characterizations of probasin gene features in the rodent superfamily Muroidea. **(A)** A general overview of the *probasin* gene locus in the house mouse, *Mus musculus*. Exons are numbered and their respective positions are drawn to scale. **(B)** A comparison of relative exon sizes (numbered and drawn to scale), coding region differences relative to the house mouse (blue, shared sizes; magenta, deviating sizes), and untranslated regions (gray). Dashed-lined open boxes indicate that the size of the untranslated region is not determined. The house mouse and rats, *Rattus spec.*, represent the family Muridae, *Peromyscus spec.* represent Cricetidae (hamsters), *Cricetomys spec.* represent the Nesomyidae family, and the *R. pruniosus* species gene information is derived from a Spalacidae representative. Black arrows with respective lower case letters a to d indicate relative positions of different primers used in the RT-PCR experiment. **(C)** cDNA analysis derived from mouse prostate tissue employing primer combinations localizing to different exons confirms the presence of exon 7 in *Pbsn* transcripts in the house mouse. Abbreviations: b, base; kb, kilobase; base pair.

Similarly to mice, the only two available Spalacidae sequences (*N. galili*, *R. pruniosus*) had a preliminary stop codon on exon 6. Generally, exon 2 to exon 5 displayed identical sizes in all analyzed species, i.e., 140, 68, 111, and 102 bases, respectively. The only exceptions were found in the family Spalacidae (cf. Data Set S1). A PCR analysis of mouse prostate cDNA using primer sets annealing to different exons revealed that exon 7 was indeed incorporated into the mRNA despite that it did not account for the coding region (Figure 2C).

A more detailed analysis based on alignments of sequences from representative Muroidea species revealed that in *M. musculus*, *Apodemus spec.*, and *Praomys delectorum* (Figures 3A and 4), an insertion of 24 bases occurred into exon 6, introducing a preliminary stop codon. However, the splice donor site ('GT' nucleotides) following exon 6 and the splice acceptor site ('AG' nucleotides) of exon 7 carrying the 'designated' stop codon remained largely conserved (Figure 3A,B).

However, there are exceptions in the species *M. musculus*, *M. caroli*, *M. spicilegus*, and *M. spretus*, which all belong to the subgenus 'Mus' within the genus 'Mus'. A mutation from an A to G nucleotide generated a novel splice acceptor site eight bases upstream of exon 7 ('AA' to 'AG'), which was confirmed by comparing the differences in the exon 6/7 junctions with deposited sequence information of the house mouse (NCBI's nucleotide repository; AF005204.1, Johnson and Greenberg, see: <https://www.ncbi.nlm.nih.gov/nuccore/AF005204.1> (accessed on 23 November 2024)) and rat (M27156.1) [23] (Figure 3C).

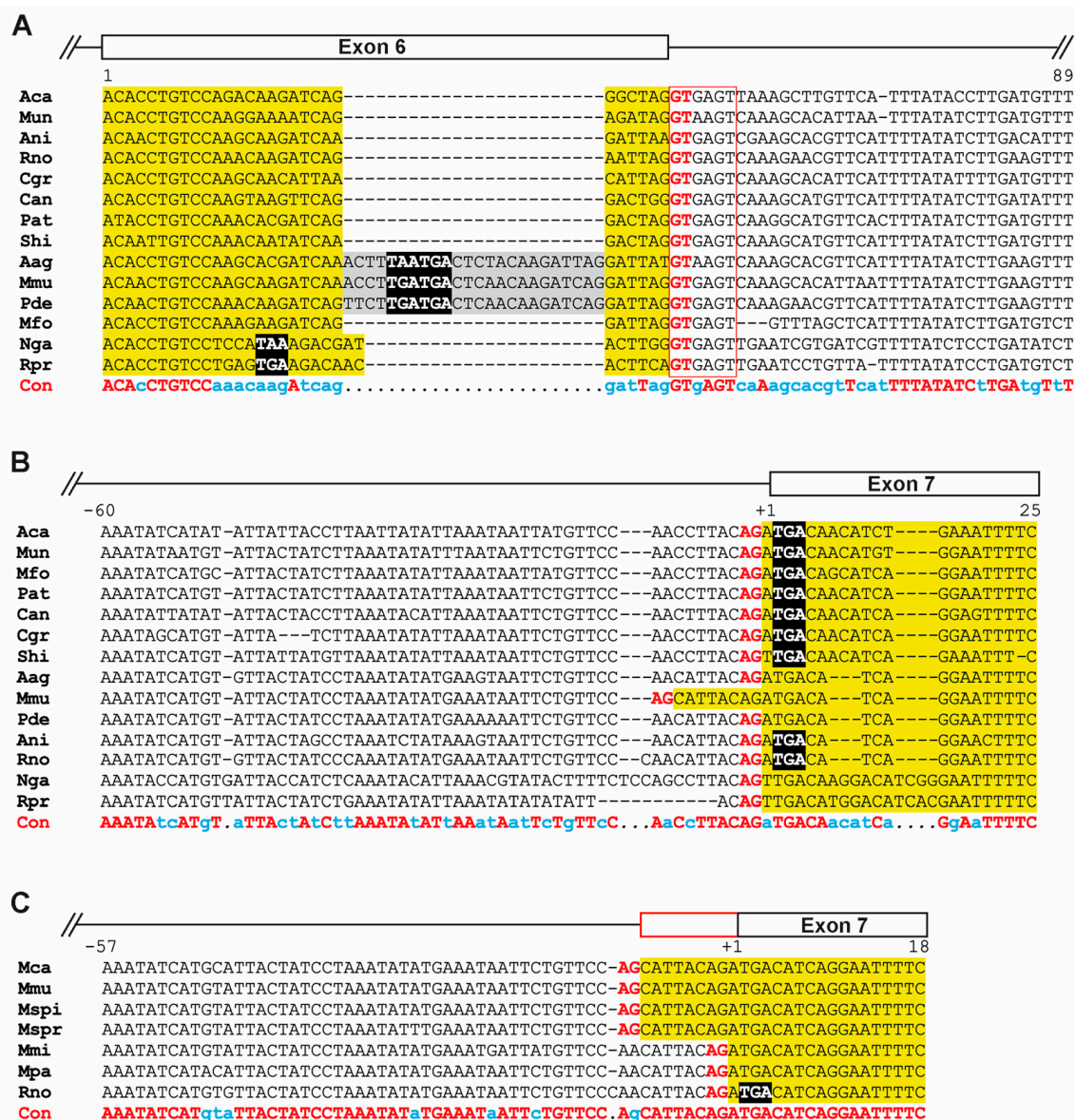


Figure 3. Differences within probasin-coding nucleotide sequences. (A) The alignment of selected nucleotide sequences comprising full-length exon 6 and parts of the subsequent intron. (B) The alignment of selected nucleotide sequences comprising the beginning of exon 7 as well as parts of the upstream intron sequence. (C) The alignment of selected nucleotide sequences from the genus *Mus* and from rats comprising the beginning of exon 7 as well as parts of the upstream intron sequence. Alignment was performed using MultAlin [22] with default settings using the Symbol comparison table ‘DNA-5-0’; in the case of (B), Gap penalties at the ‘end’ were chosen. In (A–C), nucleotide sequences highlighted in yellow represent exon sequences, and gray highlights insertions exclusive to the Murinae species Aag, Mmu, and Pde. The list of abbreviated species can be found in Appendix A. All stop codons are depicted in black with a white font. The consensus (Con) shows conserved nucleotides using a red font and upper case; a blue font with lower case indicates frequently used nucleotides in the alignment. (A) The open box in red indicates a splice donor site following the coding region of exon 6. All splice donor (GT) and splice acceptor sites (AG) are depicted in an upper-case red font.

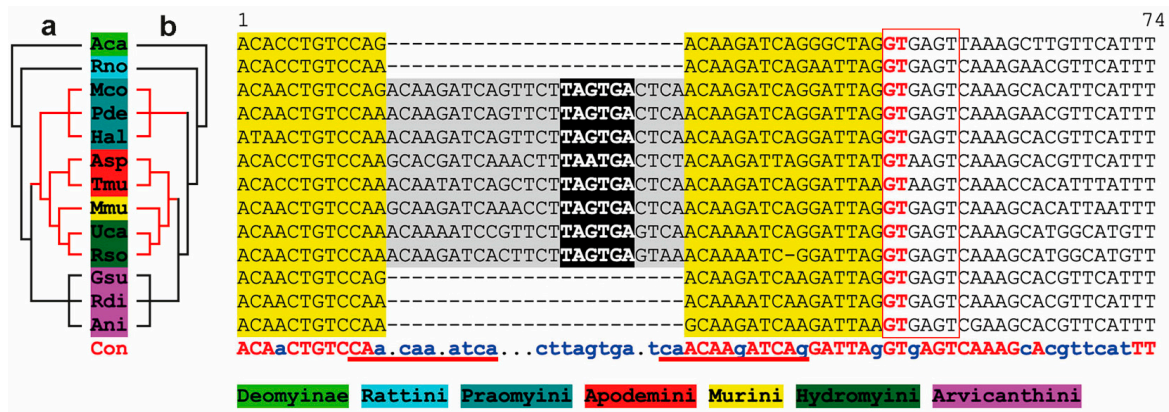


Figure 4. A comparison of sequence insertions into *probasin* exon 6 within the subfamilies Murinae and Deomyiinae. Selected sequences of 13 different representative Murinae and Deomyiinae species are aligned, comprising full-length exon 6 (yellow) and the downstream intron sequence. Stop codons emerging in the reading frame are highlighted in black with a white font. Insertions of additional sequences in a subset of species are depicted in gray. Alignment was performed using MultAlin with default settings using the Symbol comparison table ‘DNA-5-0’ [22]. A consensus sequence with identical bases (upper case, red) and frequent bases (lower case, blue) is displayed underneath the species’ alignment. Potential sequence duplications are underlined (red). The species are color-coded referring to the respective subfamily or tribe. (a) The revised relative phylogenetic relationship based on our findings. (b) The relative phylogenetic relationship presented by Aghová and co-workers [2]. Species abbreviations can be found in Appendix A. The open red box indicates splice donor site following the coding region of exon 6. The splice donor (GT) is depicted in an upper-case red font.

3.2. PBSN Gene Sizes and Nucleotide Compositions Are Comparable Among Muroidea Species

Given that the Spalacidae sequences show a few distinguishing features compared to all other Muroidea families, we assessed differences in gene size (Figure 5A), the content of G or C nucleotides in the coding sequence (GC content, see Figure 5B), and the GC content in the third codon position (GC3 content, Figure 5C). Where statistical evaluation was applicable, we found a modest but significant difference in the GC content by comparing Cricetidae (41.43%) to Muridae species (40.81%) (Figure 5B). Any other comparisons showed no significant difference. The GC-to-GC3 relationship in Cricetidae, Muridae, and Nesomyidae species followed a linear relationship. Whereas the Spalacidae species *N. galili* spearheaded this relationship, *R. pruinosus* emerged as an outlier (Figure 5D).

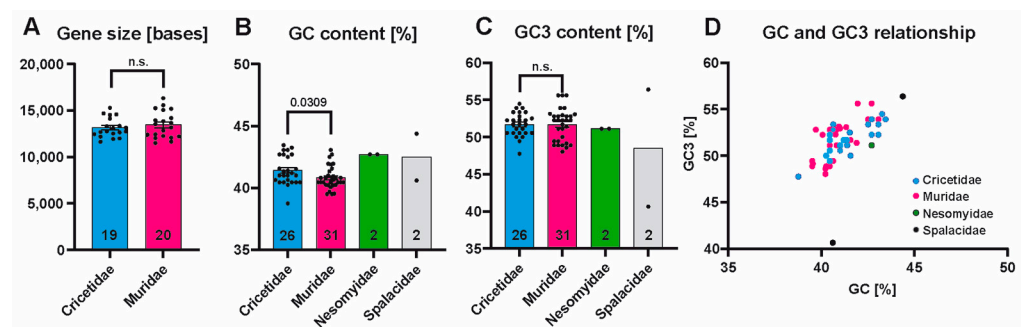


Figure 5. (A) A comparison of gene sizes between Cricetidae and Muridae species. (B,C) Comparisons of the average GC content in the coding sequence (B) or in the third position of each codon (C). (D) An XY scatter blot depicting the GC-and-GC3 relationship in different species. In (A–C), all columns are labeled with the number (n) of species included, and the sizes represent the mean size/content of analyzed species, and, where applicable, error bars are inserted representing the S.E.M. (standard error of the mean). Abbreviations: GC, guanosine/cytidine nucleotides; GC3, guanosine/cytidine nucleotides in the third position of a codon; n.s., not significant.

3.3. Phylogenetic Analysis of Probasin-Coding Sequences Outgroups the Spalacida Family

The phylogenetic relationship was inferred based on the probasin-coding nucleotide sequence of 61 species (Figure 6). The phylogenetic distance of both Spalacidae sequences from all other sequences was consistent with the Time Tree results (Figure S1). The Cricetidae and Muridae families further separated from each other, while the Nesomyidae species were incorporated in the Muridae sub-branch. The above-mentioned shifts in the exon 7 splice acceptor sites in species of the subgenus *Mus* clustered together, despite that exon 7 sequences were not incorporated in this analysis.

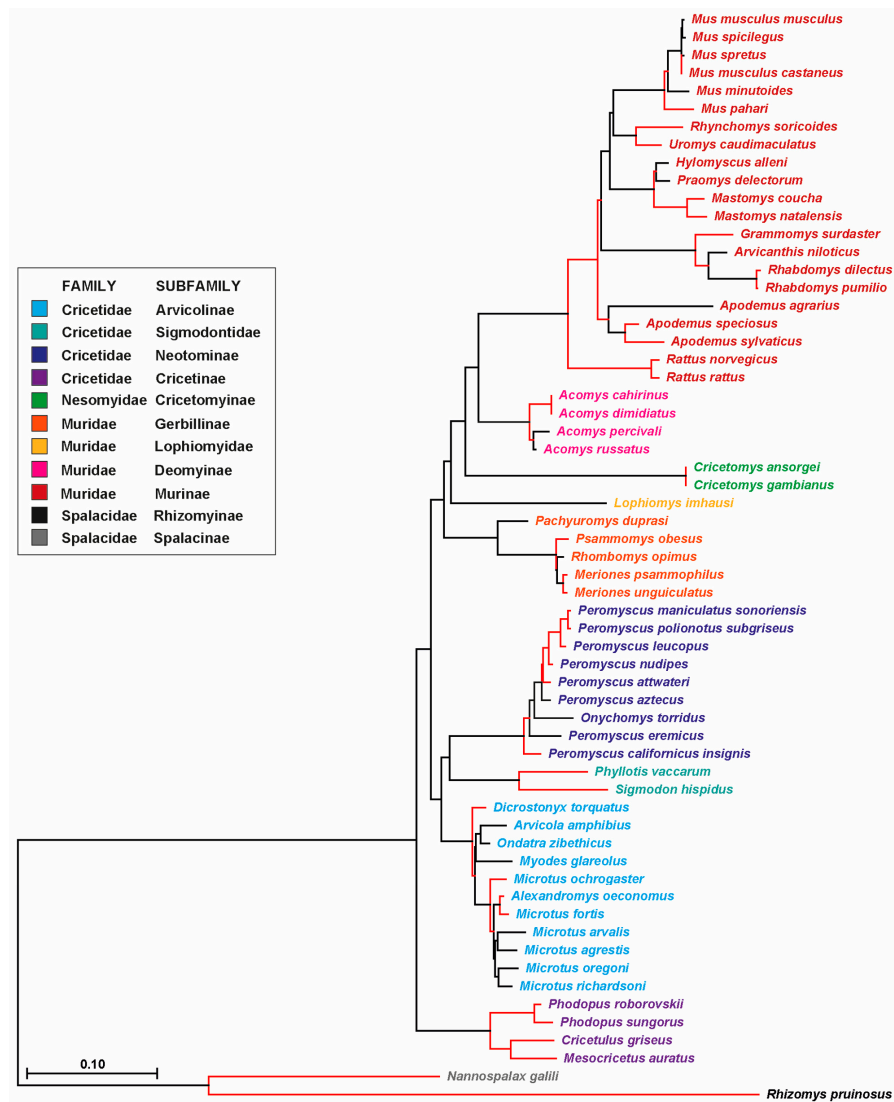


Figure 6. A phylogenetic tree of probasin-coding nucleotide sequences. The evolutionary history was inferred using the Neighbor-Joining method [24]. The optimal tree with the sum of a branch length = 3.11734955 is shown. The percentage (>70) of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) [25] are indicated by red branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to produce the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method [26] and are in the units of the number of base substitutions per site. The rate variation among sites was modeled with a gamma distribution (shape parameter = 1). The differences in the composition bias among sequences were considered in evolutionary comparisons [27]. This analysis involved 61 nucleotide sequences. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There was a total of 540 positions in the final data set. Evolutionary analyses were conducted in Mega X [17,18].

3.4. Genomic Recombination Events Affect *PBSN* Gene Localization on the X-Chromosome

Currently, *probasin* sequences are only partially annotated or assigned to a chromosomal locus. To clarify the origin and evolution of *PBSN* beyond what has been published in mice and rats, we chose to compare the genomic vicinities of *PBSN* to determine whether this has been subject to change during rodent evolution (Figure 7). *PBSN* occurred first in Muroidea families. In *N. galili*, it localizes upstream of *PRKX* and *MXRA5* (matrix remodeling associated 5), two genes that are present in the X-specific part of the X-chromosome immediately proximal to the human PAR. In other mammals, both genes mostly reside within the PAR [1]. In Cricetidae (e.g., *P. leucopus*) and different Muridae species (rat, mouse, etc.), recombination events lead to a translocation of *PRKX* and *PBSN* downstream of *TBL1X* (Figure 7). Additionally, a gene cluster comprising *PRRG1* (proline-rich and Gla domain 1), *LANCL3* (LanC-like family member 3), and *XK* (X-linked Kx-blood-group antigen, Kell, and VSP 13A binding protein) split up during the separation of Muridae and Cricetidae lineages from Spalacidae, eventually ‘homing’ the recombined *TBL1X/PRKX/PBSN* gene cluster.

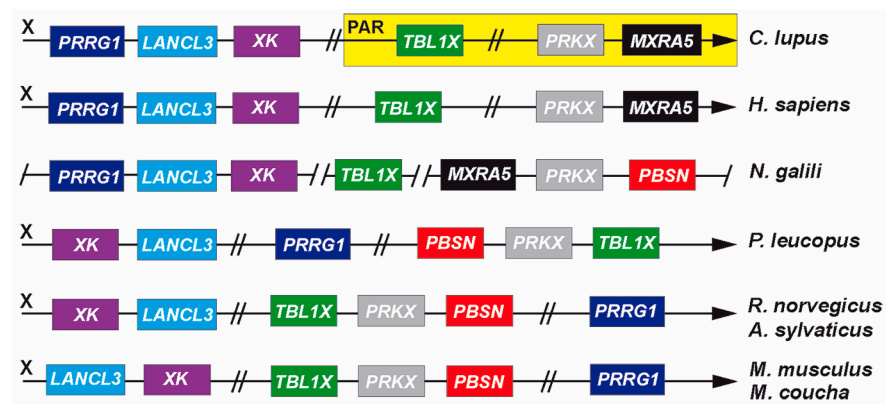


Figure 7. Recombination events affecting *PBSN* gene localization on the X-chromosome. A comparison of the *probasin* gene location relative to its neighboring genes in selected annotated X-chromosomes of the domestic dog (*C. lupus familiaris*), humans (*H. sapiens*), the Cricetidae species *Peromyscus leucopus* (‘white-footed mouse’), and several Muridae species: *Rattus norvegicus*, *Apodemus sylvaticus*, *Mastomys coucha*, and the house mouse. The dog genome is depicted to indicate that the genes *TBL1X*, *PRKX*, and *MXRA5* are located in the pseudoautosomal region (PAR) of the X- and Y-chromosomes (yellow-shaded region), whereas in humans, this genomic vicinity is already incorporated in the X-specific region of the X-chromosome [1]. In the genome of the Spalacidae species *N. galili*, the vicinity of several genes was identified but lacks formal assignment to the X-chromosome. Arrowheads point to the telomer. Double-hyphenated bars indicate unspecified distances between depicted genes. Genes are presented in the order of appearance. The occurrence of recombination events in the evolution of the X-chromosome in Cricetidae and Muridae indicate that *PRKX* and *PBSN* recombined next to *TBL1X*. Additionally, the gene cluster, including *PRRG1*, *LANCL3*, and *XK*, fell apart, now homing among other genes the cluster comprising *TBL1X*, *PRKX*, and *PBSN*.

Further changes happened in the orientation of *LANCL3* and *XK* on the X-chromosome along the separation into different Muridae lineages. Since the sequence identities and the genomic vicinities of two ‘*probasin*-like’ genes in the Dipodidae species *J. jaculus* are fundamentally different, they were not included in our analysis.

3.5. *Probasin* Proteins Retain Structural Features of Lipocalin Protein Family

Finally, we compared the protein sequences of species representing different Muroidea branches and inferred the structural information available from mouse *probasin* protein (Figure 8).

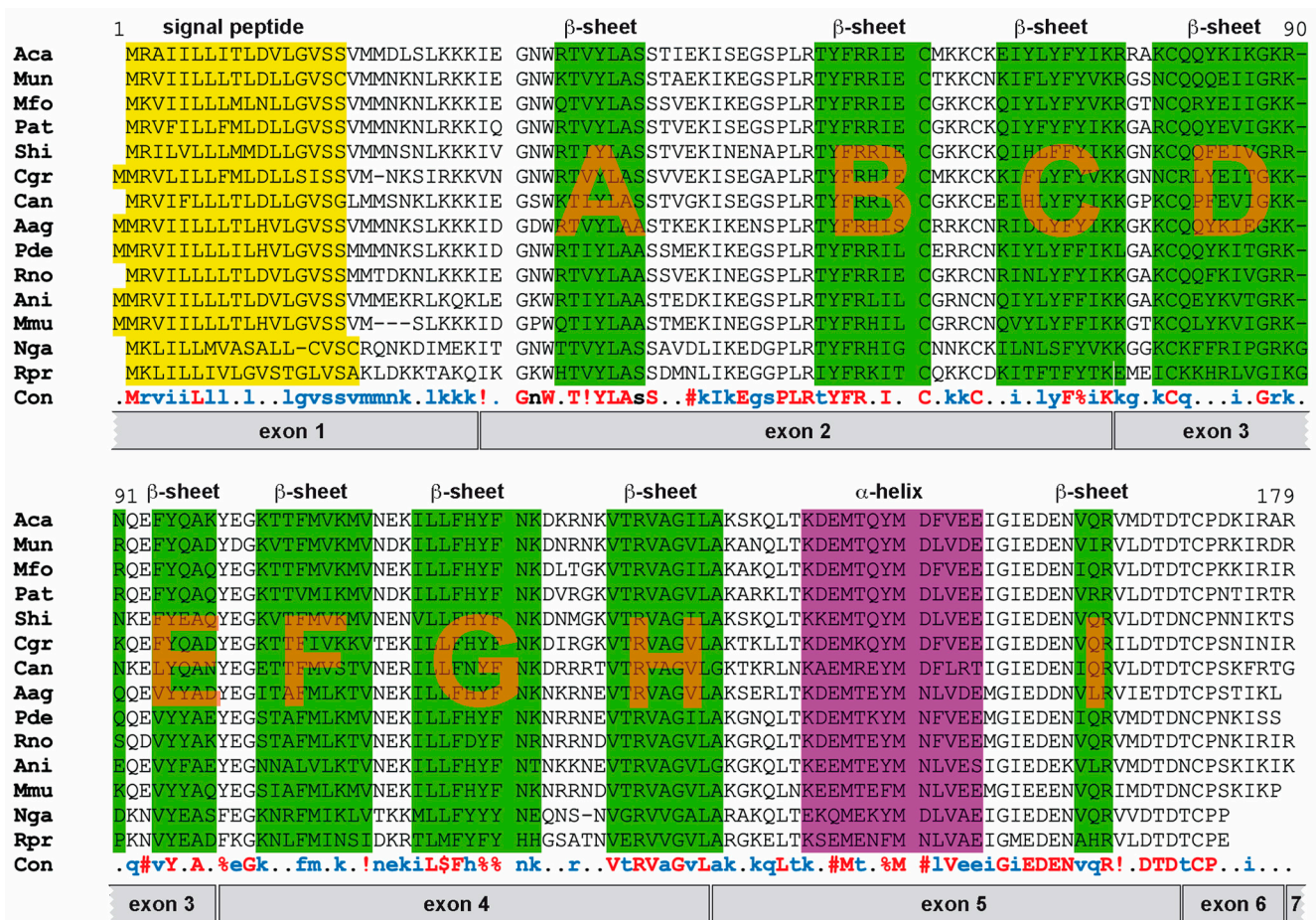


Figure 8. A comparison of probasin protein sequences representing different Muroidea families and subfamilies. Protein sequences were aligned using MultAlin [22], and different protein features are highlighted, imposing the information available for the presumptive structure of mouse probasin selected from the AlphaFold2.0 database [28,29]. Green highlighted regions represent β -sheets, and the lilac region depicts an α -helix. β -sheets are listed alphabetically in their order of appearance. Signal peptides are highlighted in yellow. Exon coverage is depicted in the gray-shaded boxes underneath the translated consensus protein sequence. Amino acids are represented by one-letter code. The color-codes and symbols of the amino acid consensus (Con) are the following: red, identical amino acids; blue, frequent consensus of a specific amino acid; !, either I or V; \$, either L or M; %, F or Y; #, N, D, Q, E, B, or Z. Species abbreviations can be found in Appendix A.

All sequences retain a signaling peptide, one α -helix, and nine β -sheets (A to I). These features are conserved, suggesting their similarity to members of the lipocalin and MUP family [9,12]. The first eight β -sheets form a barrel-like structure necessary for the process of shuttling lipophilic biomolecules. Several other features are conserved as well, e.g., a GxW motif (glycine and tryptophane separated by one amino acid) prior to β -sheet A, a CxxxC motif (two cystines separated by three amino acids) following β -sheet B, and a cysteine at the beginning of β -sheet D forming a disulfide bond with a cysteine close to the C-terminus (Figure 8) [9].

4. Discussion

The emergence of the lipocalin-like protein probasin remains elusive given that its presence has yet only been described in rats and mice [8,9], and current genomic annotations are only available for *probasin* or ‘*probasin-like*’ genes within rodents (NCBI). Using the mouse and rat coding sequences as references, we were able to identify *probasin* genes by similarity searches in over sixty rodent genomes, which all belong to the Muroidea

superfamily. This dates to a first emergence roughly 44–48 million years ago, before the Spalacidae family separated from the Cricetidae, Muridae, and Nesomyidae families, which were also subsumed by the umbrella term Eumuroidea (Figure S1) [30]. The genomic vicinity of the *probasin* gene in the Spalacidae and Eumuroidea species suggests that it localizes to the X-chromosome, and that at the branching point of both lineages, further recombination events happened. Whereas *PBSN* is located in *N. galili* in a highly conserved gene cluster that is frequently found in the mammalian PAR, in Eumuroidea, *PBSN* as well as its neighboring gene *PRKX* both recombined with *TBL1X* to an X-specific region, an event recently described for other genes within this clade [2,3]. Within the evolution of the Eumuroidea species, additional recombination events occurred, but this happened without further disrupting the newly joined *TBL1X/PRKX/PBSN* gene cluster.

Previously, we reported that the Eumuroidea PAR is subject to erosion, i.e., genes translocated to autosomes as well as to X-specific regions on the X-chromosome [2,3]. Illustrated by the evolution of the *neuroligin-4* gene, we demonstrated that gene sizes collapsed, alternatively spliced exons were lost, and the GC/GC3 content in synonymous codons increased substantially [3,4]. In contrast to that, generally, *PBSN* genes were unaffected by the aforementioned changes. The GC and GC3 contents averaged between 40 and 45% and 45 and 55%, respectively, and when plotted together, they were distributed in a linear fashion. This was, for example, not the case for *neuroligin-4*, which had GC/GC3 levels starting at 70% and up to over 90%, eventually exiting the linear GC/GC3 relationship, forming a ‘Supermyomorpha island’ [4]. Our observation suggests that *probasin*’s recombination with its neighboring genes to an X-specific localization on the X-chromosome might have happened evolutionarily earlier before its branching from Spalacidae and Eumuroidea, saving *PBSN* from the onset of PAR erosion.

Our analysis aimed at placing mouse *Pbsn* in the broader context of other rodent *PBSN* sequences. Interestingly, based on our results, we conclude that mouse *Pbsn* is an outlier compared to the common *PBSN* gene pattern of other rodents in two specific aspects: (a) an insertion into exon 6 introducing a preliminary stop codon and (b) a shift in the splice acceptor site of exon 7. Both features help to assess the proximity of a phylogenetic relationship among certain rodents. The exon 7 splice acceptor shift affects representative species of the subgenus *Mus* within the genus *Mus*, whereas other *Mus* subgenera, i.e., *Coelomys* and *Nannomys*, are unaffected. Since the entire genus *Mus* harbors its *PBSN* stop codon on exon 6, this shift leaves the gene and its translational product unaffected. In our phylogenetic analysis, the sequence insertion in exon 6 appears—whenever present in Muridae species—of common origin. This conflicts with the ‘Time Tree’ depiction (Figure S1), suggesting that the underlying data regarding different Muridae species rely on different source material and need to be re-evaluated carefully. We found insertions in the Muridae tribes Praomyini (e.g., *P. delectorum*), Murini (*M. musculus*), Apodemini (*A. agarius*), and Hydromyini (*U. caudimaculatus*). However, Arvicanthini species (e.g., *A. niloticus*) were unaffected (Figure 5). The insertion into exon 6 challenges the current separation into different tribe branches within the Murinae subfamily. Analysis of nuclear gene sequences and fossil data sets indicate that Arvicanthini separated evolutionarily later. This would imply that at least two independent insertions at the very same position occurred [31,32]. Due to this being unlikely, our results put forward a different scenario in which the Praomyini, Murini, Apodemini, and Hydromyini tribes separated evolutionarily later from Arvicanthini. Discordances between current species trees and gene trees might be due to several reasons, such as horizontal gene transfer, gene duplication, and lineage sorting [33]. Our analysis is based on a single gene which emerged first in the common ancestor of Spalacidae and Eumuroidea despite its generally similarity to other genes of the large lipocalin family. Our estimate of Muroidea phylogeny based on *probasin* is

similar to recent rodent phylogenies at the level of relationships among families and subfamilies [31,32,34]. The underlying genes as well as *probasin* in this report have the potential to exhibit differences from the species tree [33]. Moreover, all genes in recombining regions including the X-specific part of the mammalian X-chromosome may represent a mosaic of regions with slightly different gene trees [35]. Despite these potential complexities, we found that the evolutionary history of the *probasin* gene, inferred using Neighbor-Joining, exhibited limited topological discordance with the rodent species tree. Recently, Steppan and Schenk acknowledged, regarding the Muridae phylogeny, that despite testing several genes, not all basal core murine nodes were strongly supported; hence, further geographical and morphological features should be taken into account [34].

Probasin proteins appear largely conserved across all analyzed species, displaying the highest similarity in important structural elements and motifs suggesting that functional aspects are retained. Thus far, the structural relationship to lipocalins indicates a potential role in shuttling lipophilic biomolecules. For instance, in co-purification experiments, heparin-binding growth factor-1 has been identified as interacting with probasin [36]. Whether the role of these biomolecules is exclusive to the prostate remains elusive and needs to be addressed in future experiments. Probasin has a close structural relationship to other lipocalins such as major urinary proteins and odorant-binding proteins, which are both involved in the shuttling of pheromones [12]. Combinations of these proteins are supposed to potentially have a signature role for different species [12,37]. The addition of probasin to this 'cocktail', which is produced by the prostate and, hence, added to seminal fluid, might contribute to a signature blend of odorants shuttled by MUPs and OBPs. We need to await the generation of probasin-deficient mice to assess its potential beneficial evolutionary role.

5. Conclusions

Probasin is a unique lipocalin-like gene found in a subset of mouse-like rodents, Muroidea. It likely emerged in common ancestors within the pseudoautosomal region of both sex chromosomes. However, recombination events during evolution resulted in its current placement in the X-specific region of the X-chromosome in, e.g., the house mouse. Analysis of the coding region suggests that the translocation from the pseudoautosomal region elsewhere occurred prior to changes that affected its coding region, i.e., no increase in the GC or the GC3 content in its coding region is observed. Subtle changes to its coding region such as preliminary stop codons in exon 6 or changes to splice acceptor sites in exon 7 help to clarify evolutionary relationships among closely related species on the infra-family level of Muridae.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/biology14030239/s1>. Figure S1: Evolutionary relationship of different species depicted in Time Tree; Data Set S1: Gene analyses.

Author Contributions: Conceptualization, S.M.; formal analysis, S.M.; investigation, S.M., L.W. and S.L.Z.; resources, G.K.-C.; data curation, S.M.; writing—original draft preparation, S.M.; writing—review and editing, L.W., S.L.Z. and G.K.-C.; visualization, S.M., L.W. and S.L.Z.; supervision, S.M.; project administration, S.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: All animal procedures (animal care procedures, anesthesia, and sacrificing mice for organ collection) were reviewed and approved by the animal welfare and ethics committee of Saarland University (Tierschutzbeauftragte der Universität des Saarlandes, Homburg) (code VM2023-21; effective 01/2024).

Informed Consent Statement: Not applicable.

Data Availability Statement: The relevant data supporting the findings of this study are available in this article and its Supplementary Information Files.

Acknowledgments: We would like to thank Monika Hollenhorst for the critical reading of this manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

In Figures 3, 4 and 8, the following species abbreviations are used: Aag, *Apodemus agrarius*; Aca, *Acomys cahirinus*; Ani, *Arvicanthis niloticus*; Asp, *Apodemus speciosus*; Can, *Cricetomys ansorgei*; Cgr, *Cricetulus griseus*; Gsu, *Grammomys surdaster*; Hal, *Hylomyscus alleni*; Mca, *Mus caroli*; Mco, *Mastomys coucha*; Mfo, *Microtus fortis*; Mmi, *Mus minutoides*; Mmu, *Mus musculus*; Mpa, *Mus pahari*; Mspi, *Mus spicilegus*; Mspr, *Mus spretus*; Mun, *Meriones unguiculatus*; Nga, *Nannospalax galili*; Pat, *Peromyscus attwateri*; Pde, *Praomys delectorum*; Rdi, *Rhabdomys dilectus*; Rno, *Rattus norvegicus*; Rpr, *Rhizomys prunioides*; Rso, *Rhynchomys soricoides*; Shi, *Sigmodon hispidus*; Tmu, *Tokudai muenninki*; Uca, *Uromys caudimaculatus*.

The following abbreviations are used for nucleotide bases: A, adenine; C, cytosine, G, guanine, and T, thymine.

References

- Raudsepp, T.; Chowdhary, B.P. The Eutherian Pseudoautosomal Region. *Cytogenet. Genome Res.* **2015**, *147*, 81–94. [[CrossRef](#)]
- Maxeiner, S.; Gebhardt, S.; Schweizer, F.; Venghaus, A.E.; Krasteva-Christ, G. Of mice and men-and guinea pigs? *Ann. Anat.* **2021**, *238*, 151765. [[CrossRef](#)]
- Maxeiner, S.; Benseler, F.; Krasteva-Christ, G.; Brose, N.; Südhof, T.C. Evolution of the Autism-Associated Neurologin-4 Gene Reveals Broad Erosion of Pseudoautosomal Regions in Rodents. *Mol. Biol. Evol.* **2020**, *37*, 1243–1258. [[CrossRef](#)] [[PubMed](#)]
- Maxeiner, S.; Benseler, F.; Brose, N.; Krasteva-Christ, G. Of Humans and Gerbils- Independent Diversification of Neurologin-4 Into X- and Y-Specific Genes in Primates and Rodents. *Front. Mol. Neurosci.* **2022**, *15*, 838262. [[CrossRef](#)]
- International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860–921. [[CrossRef](#)] [[PubMed](#)]
- Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature* **2002**, *420*, 520–562. [[CrossRef](#)]
- Burgin, C.J.; Colella, J.P.; Kahn, P.L.; Upham, N.S. How many species of mammals are there? *J. Mammal.* **2018**, *99*, 1–14. [[CrossRef](#)]
- Matuo, Y.; Nishi, N.; Muguruma, Y.; Yoshitake, Y.; Kurata, N.; Wada, F. Localization of prostatic basic protein (“probasin”) in the rat prostates by use of monoclonal antibody. *Biochem. Biophys. Res. Commun.* **1985**, *130*, 293–300. [[CrossRef](#)] [[PubMed](#)]
- Kasper, S.; Matusik, R.J. Rat probasin: Structure and function of an outlier lipocalin. *Biochim. Biophys. Acta* **2000**, *1482*, 249–258. [[CrossRef](#)] [[PubMed](#)]
- Johnson, M.A.; Hernandez, I.; Wie, Y.; Greenberg, N. Isolation and characterization of mouse probasin: An androgen-regulated protein specifically expressed in the differentiated prostate. *Prostate* **2000**, *43*, 255–262. [[CrossRef](#)]
- Charkoftaki, G.; Wang, Y.; McAndrews, M.; Bruford, E.A.; Thompson, D.C.; Vasiliou, V.; Nebert, D.W. Update on the human and mouse lipocalin (LCN) gene family; including evidence the mouse Mup cluster is result of an “evolutionary bloom”. *Hum. Genom.* **2019**, *13*, 11. [[CrossRef](#)]
- Pelosi, P.; Knoll, W. Odorant-binding proteins of mammals. *Biol. Rev. Camb. Philos. Soc.* **2022**, *97*, 20–44. [[CrossRef](#)] [[PubMed](#)]
- Greenberg, N.M.; DeMayo, F.; Finegold, M.J.; Medina, D.; Tilley, W.D.; Aspinall, J.O.; Cunha, G.R.; Donjacour, A.A.; Matusik, R.J.; Rosen, J.M. Prostate cancer in a transgenic mouse. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 3439–3443. [[CrossRef](#)] [[PubMed](#)]
- Li, Z.; Szabolcs, M.; Terwilliger, J.D.; Efstratiadis, A. Prostatic intraepithelial neoplasia and adenocarcinoma in mice expressing a probasin-Neu oncogenic transgene. *Carcinogenesis* **2006**, *27*, 1054–1067. [[CrossRef](#)]
- Wu, X.; Wu, J.; Huang, J.; Powell, W.C.; Zhang, J.; Matusik, R.J.; Sangiorgi, F.O.; Maxson, R.E.; Sucov, H.M.; Roy-Burman, P. Generation of a prostate epithelial cell-specific Cre transgenic mouse model for tissue-specific gene ablation. *Mech. Dev.* **2001**, *101*, 61–69. [[CrossRef](#)]
- Maxeiner, S.; Krasteva-Christ, G.; Althaus, M. Pitfalls of using sequence databases for heterologous expression studies—a technical review. *J. Physiol.* **2023**, *601*, 1611–1623. [[CrossRef](#)] [[PubMed](#)]

17. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)]
18. Stecher, G.; Tamura, K.; Kumar, S. Molecular Evolutionary Genetics Analysis (MEGA) for macOS. *Mol. Biol. Evol.* **2020**, *37*, 1237–1239. [[CrossRef](#)]
19. Almagro Armenteros, J.J.; Tsirigos, K.D.; Sønderby, C.K.; Petersen, T.N.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **2019**, *37*, 420–423. [[CrossRef](#)]
20. Hollenhorst, M.I.; Husnik, T.; Zylka, M.; Duda, N.; Flockerzi, V.; Tschernig, T.; Maxeiner, S.; Krasteva-Christ, G. Human airway tuft cells influence the mucociliary clearance through cholinergic signalling. *Respir. Res.* **2023**, *24*, 267. [[CrossRef](#)] [[PubMed](#)]
21. D'Elía, G.; Fabre, P.H.; Lessa, E.P. Rodent systematics in an age of discovery: Recent advances and prospects. *J. Mammal.* **2019**, *100*, 852–871. [[CrossRef](#)]
22. Corpet, F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **1988**, *16*, 10881–10890. [[CrossRef](#)] [[PubMed](#)]
23. Spence, A.M.; Sheppard, P.C.; Davie, J.R.; Matuo, Y.; Nishi, N.; McKeehan, W.L.; Dodd, J.G.; Matusik, R.J. Regulation of a bifunctional mRNA results in synthesis of secreted and nuclear probasin. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 7843–7847. [[CrossRef](#)]
24. Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425.
25. Felsenstein, J. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **1985**, *39*, 783–791. [[CrossRef](#)]
26. Tamura, K.; Nei, M.; Kumar, S. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 11030–11035. [[CrossRef](#)] [[PubMed](#)]
27. Tamura, K.; Kumar, S. Evolutionary distance estimation under heterogeneous substitution pattern among lineages. *Mol. Biol. Evol.* **2002**, *19*, 1727–1736. [[CrossRef](#)] [[PubMed](#)]
28. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Židek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)] [[PubMed](#)]
29. Varadi, M.; Bertoni, D.; Magana, P.; Paramval, U.; Pidruchna, I.; Radhakrishnan, M.; Tsenkov, M.; Nair, S.; Mirdita, M.; Yeo, J.; et al. AlphaFold Protein Structure Database in 2024: Providing structure coverage for over 214 million protein sequences. *Nucleic Acids Res.* **2024**, *52*, D368–D375. [[CrossRef](#)]
30. Steppan, S.; Adkins, R.; Anderson, J. Phylogeny and divergence-date estimates of rapid radiations in muroid rodents based on multiple nuclear genes. *Syst. Biol.* **2004**, *53*, 533–553. [[CrossRef](#)]
31. Fabre, P.H.; Herrel, A.; Fitriana, Y.; Meslin, L.; Hautier, L. Masticatory muscle architecture in a water-rat from Australasia (Murinae: Hydromys) and its implication for the evolution of carnivory in rodents. *J. Anat.* **2017**, *231*, 380–397. [[CrossRef](#)]
32. Aghová, T.; Kimura, Y.; Bryja, J.; Dobigny, G.; Granjon, L.; Kergoat, G.J. Fossils know it best: Using a new set of fossil calibrations to improve the temporal phylogenetic framework of murid rodents (Rodentia: Muridae). *Mol. Phylogenet. Evol.* **2018**, *128*, 98–111. [[CrossRef](#)] [[PubMed](#)]
33. Maddison, W.P. Gene trees in species trees. *Syst. Biol.* **1997**, *46*, 523–536. [[CrossRef](#)]
34. Steppan, S.J.; Schenk, J.J. Muroid rodent phylogenetics: 900-species tree reveals increasing diversification rates. *PLoS ONE* **2017**, *12*, e0183070. [[CrossRef](#)]
35. Hobolth, A.; Christensen, O.F.; Mailund, T.; Schierup, M.H. Genomic Relationships and Speciation Times of Human, Chimpanzee, and Gorilla Inferred from a Coalescent Hidden Markov Model. *PLoS Genet.* **2007**, *3*, e7. [[CrossRef](#)] [[PubMed](#)]
36. Matuo, Y.; Adams, P.S.; Nishi, N.; Yasumitsu, H.; Crabb, J.W.; Matusik, R.J.; McKeehan, W.L. The androgen-dependent rat prostate protein, probasin, is a heparin-binding protein that co-purifies with heparin-binding growth factor-1. *In Vitro Cell. Dev. Biol.* **1989**, *25*, 581–584.
37. Roberts, S.A.; Prescott, M.C.; Davidson, A.J.; McLean, L.; Beynon, R.J.; Hurst, J.L. Individual odour signatures that mice learn are shaped by involatile major urinary proteins (MUPs). *BMC Biol.* **2018**, *16*, 48. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.