Tracing the evolution of prestin's area-motor activity through ancestral sequence reconstruction and structural modeling

Nicolas Fuentes-Ugarte*

Departamento de Biologia, Facultad de Ciencias, Universidad de Chile, Santiago,Chile.

Tiaren Ruiz-Rojas*

Facultad de Ingenieria, Universidad San Sebastian, Santiago, Chile.

Alvaro Ruiz-Fernandez

Computational Biology Lab, Centro Cientifico y Tecnologico de Excelencia, Fundacion Ciencia & Vida, Santiago, Chile, Santiago, Chile.

Jose Antonio Garate

Facultad de Ingenieria, Universidad San Sebastian, Santiago, Chile

Victor Castro-Fernandez†

Departamento de Biologia, Facultad de Ciencias, Universidad de Chile, Santiago, Chile.

Raul Araya-Secchi‡

Facultad de Ingenieria, Universidad San Sebastian, Santiago, Chile and Centro Cientifico y Tecnologico de Excelencia, Fundacion Ciencia & Vida, Santiago, Chile

Abstract

Prestin, a member of the SLC26A family, is essential for the electromotility of mammalian outer hair cells, converting voltage changes into mechanical work. In contrast, non-mammalian orthologues function as anion transporters, lacking electromechanical activity. To investigate the molecular and structural basis of this functional divergence, we performed ancestral sequence reconstruction (ASR) of prestin across vertebrates, followed by comparative structural modeling using AlphaFold2-multimer. We identified more than 200 amino acid substitutions along the lineage that lead to placental mammals, with early substitutions concentrated in the transmembrane domain (TMD) and late substitutions clustering in the STAS domain, particularly in the intervening sequence (IVS). Structural modeling revealed that early substitutions modulate protein–lipid interactions and interhelical contacts. In placental mammals, the IVS-loop adopts a distinct conformation that places a negatively charged patch near the chloride access pathway, potentially affecting the ion dynamics and voltage responsiveness. These structural transitions occurred without major rearrangements of the global fold of prestin, supporting a model in which the novel function evolved through distributed substitutions within a conserved scaffold. Our findings illustrate how exaptation and incremental structural remodeling enabled the repurposing of an ancestral anion transporter into a voltage-sensitive area-motor, providing a framework for understanding the molecular evolution of complex biophysical traits.

INTRODUCTION

The auditory system of mammals exhibits remarkable sensitivity and frequency selectivity. These features rely on the cochlear amplifier, a mechanism that enhances sound-induced vibrations within the cochlea to improve sensitivity and frequency resolution. At the heart of this amplification process are outer hair cells (OHCs), specialized sensory cells that exhibit electromotility (EM), the ability to rapidly change body length in response to electrical stimulation[1, 2, 3, 4, 5]. EM is driven by the area-motor protein prestin, densely packed in the lateral membrane of OHCs[3, 6, 7].

Prestin is a member of the SLC26A family, which predominantly comprises anion trans-

* These authors contributed equally to this work

† Corresponding author: vcasfe@uchile.cl

‡ Corresponding author: raul.araya@uss.cl

porters[8]. As such, non-mammalian prestin orthologs function as non-electrogenic anion exchangers[9, 10, 11], while mammalian prestin has evolved into a direct voltage-to-displacement converter, operating on millisecond or faster timescales, which are essential for high-frequency sound amplification[3, 5, 7, 12, 13].

Phylogenetic analyzes have revealed that mammalian prestin evolved from an anion transporter present in a pre-mammalian ancestor[14, 15, 16]. Mammalian prestin exhibits minimal transport capacity and pronounced nonlinear capacitance (NLC), a hallmark of voltage sensitivity[11, 17, 18]. Furthermore, an inverse relationship has been established between anion transport and motor function, suggesting that transport capacity was lost as voltage-dependent NLC was gained [10, 19]. Thus, the unique function of mammalian prestin is the result of evolutionary adaptations that have repurposed an ancestral anion transporter into a highly specialized motor protein. This functional shift from ion transport to electromotility remains incompletely understood, particularly with regard to how key substitutions have fine-tuned this mechanism in placental mammals.

Considerable efforts have been devoted to identify the regions and residues where the differences between mammalian and non-mammalian prestin are most pronounced and critical for the generation of NLC and EM. Early work identified specific segments within the transmembrane domain (TMD), notably TM1–2 and TM9–11, as essential to NLC/EM and determined that replacing these regions in the zefrafish prestin by those of gerbil, a synthetic prestin capable of transport and NLC could be constructed[20]. Others showed that the extracellular loop between TM3 and TM4 (EC-loop) is also essential for NLC/EM[21, 22, 23]. Furthermore, increased flexibility of TM6 was proposed to contribute to NLC/EM[24], as well as certain conserved and non-conserved cysteines[25]. In addition, membrane thickness has been shown to affect mammalian and non-mammalian prestin differently[26].

Recently, the structures of human(HsPres)[27], dolphin(TtPres)[24], and gerbil (MuPres)[28] prestin have provided invaluable information on its functional organization and its interactions with the lipid membrane. From these structures, molecular dynamics (MD) simulation studies have shown the existence of elevator-like movements as part of its conformational change cycle[29] and how prestin structure affects the organization of the membrane in its surroundings[30]. However, fundamental questions remain regarding the identity of the voltage sensor (extrinsic or intrinsic)[9, 31, 32], how voltage triggers conformational

change[24, 33, 34], and how disordered regions, absent in experimental structures such as the IVS-loop, contribute to the specialized function of mammalian prestin[35].

Here, we address these gaps employing ancestral sequence reconstruction (ASR) combined with structural modeling and conformational sampling to identify and contextualize key substitutions in prestin evolution. We map these changes to prestin domains and regions and then use conformational sampling with AlphaFold2-multimer (AF2-multimer)[36, 37] to evaluate their structural effects. This approach not only corroborates the relevance of known functional sites, but unveils previously unexplored features, notably in the IVS-loop of the STAS domain. By modeling non-mammalian prestin structures for the first time, we offer an evolutionary perspective on how incremental substitutions collectively reshaped an ancestral anionic transporter into the high-speed area-motor protein essential for mammalian hearing.

Our results indicate that early substitutions alter contacts between residues on the TMD and between the protein and the surrounding lipids, which point toward a gradual reduction in transport capacity along with an enhanced voltage sensitivity. Despite these functional changes, the TMD architecture remains largely unchanged until more punctuated substitutions appear after the divergence between mammals and reptiles. These later substitutions cluster in the EC-loop and, more strikingly, in the IVS of the STAS domain, giving rise to pronounced structural differences between non-mammalian and mammalian prestin. In particular, alterations in the IVS-loop result in the emergence, in placental mammals, of a patch of negatively charged residues (negative patch or NP) that could influence the access/exit of chloride ions from the binding site, suggesting a pivotal role for this region in the unique electromechanical behavior of mammalian prestin.

RESULTS AND DISCUSSION

Ancestral Sequence Reconstruction Identifies Critical Evolutionary Substitutions in Prestin Domains.

We constructed a phylogenetic tree for the chordate SLC26A5 subfamily as shown in Figure 1 A (for a full version of the tree, see Supplementary File 1), identifying five monophyletic groups (fish, amphibians, reptiles, birds and mammals) with a consistent topology compared to previous studies [38, 39]. Using ancestral sequence reconstruction (ASR)[40], we inferred

12 ancestral sequences, listed in Table I, with high confidence (mean posterior probability > 0.95; see also Supplementary Figures S1, S2 and S3).

Our analysis focused on the evolutionary pathway from fish to placental mammalians (ancPis $\rightarrow$ ancRhi $\rightarrow$ ancTre $\rightarrow$ ancAmn $\rightarrow$ ancMam $\rightarrow$ ancThe $\rightarrow$ ancPla; following the nomenclature introduced in I) using as initial reference (outgroup) the closest homologous transporter, human SLC26A6 (hsA6), and human SLC26A5 (hsPres) as a model of current placental mammalian prestin. Sequence identity was high within groups (90% in non-mammals, 80% in mammals), and dropped to 60% between them (Supplementary Table S1).

We identified 207 substitutions along this evolutionary pathway: 42 in the N-terminal domain, 83 in the transmembrane domain (TMD) and 82 in the sulfate transporter and anti-sigma (STAS) domain. Substitutions were classified as early (ancAmn $\rightarrow$ ancMam; n=108), intermediate (ancMam $\rightarrow$ ancThe; n=19), or late (ancThe $\rightarrow$ ancPla; n=21), plus four earlier substitutions (ancTre $\rightarrow$ ancAmn). We also found substitutions showing two (n = 38) and three steps (n=17) in the transition from ancAmn to ancPla (or hsPres) that are indicative of regions with greater variability (the complete list of substitutions is presented in the Supplementary Tables S2, S3, S4, S5 and in the Supplementary File 2). In terms of proportions, *i.e.* the percentage of substituted residues, the N-terminal domain presents a 56% of substitutions, the TMD 19% and the STAS domain 35%, suggesting that the TMD is more conserved and less prone to substitutions.

Regarding the nature of the substitutions, we found that 47% (n = 97) of the substitutions are non-conservative, i.e., they significantly change the chemical properties of the residue (indicated in **bold*** throughout the text). They correspond to 31% of the substitutions in the N-terminal domain, 45% in the TMD, and 52% in the STAS domain. According to our findings, these substitutions tend to correspond to late, 2-step, or 3-step substitutions. Interestingly, the most prevalent types of non-conservative substitutions were those that result in the addition of glycines or prolines and those that result in the replacement of an apolar for an aromatic sidechain.

In summary, our reconstruction reveals that prestin evolution involved a substantial number of amino acid substitutions, with a strong bias toward late, non-conservative changes in the STAS domain and to a lesser extent in the TMD. The prevalence of substitutions that add glycine or proline, and apolar-to-aromatic substitutions highlights potentially function-

Table I. Robustness of the Ancestral Sequence Reconstructions.

Ancestor	Description	Node Support	ASR Mean Posterior probability ± SD
ancPis	Last common ancestor from Pisces	99.98%	0.974 ± 0.092
ancRhi	Last common ancestor from Rhipidistia	96.68%	0.972 ± 0.095
ancTre	Last common ancestor from Tetrapoda	99.98%	0.952 ± 0.125
ancAmp	Last common ancestor from Amphibia	77.94%	0.973 ± 0.090
ancAmn	Last common ancestor from Amniota	99.98%	0.974 ± 0.088
ancRep	Last common ancestor from Reptilia	93.41%	0.977 ± 0.083
ancArch	Last common ancestor from Archosauria	99.98%	0.978 ± 0.081
ancAvi	Last common ancestor from Avis	99.98%	0.976 ± 0.084
ancMam	Last common ancestor from Mammalia	99.98%	0.959 ± 0.108
ancThe	Last common ancestor from Theria	99.98%	0.962 ± 0.113
ancMar	Last common ancestor from Marsupialia	99.98%	0.996 ± 0.033
ancPla	Last common ancestor from Placentaria	99.98%	0.999 ± 0.001

altering events that occurred during the transition from ancestral transporters to the specialized motor protein observed in placental mammals.

Conformational Sampling Reveals that the IVS-loop of the STAS domain differentiates mammalian Prestin.

Placing the substitutions in the context of prestin structure is crucial to understanding their functional relevance. Thus, we generated a full-length model of human prestin (hsPres) using AlphaFold2-multimer (AF2-multimer) (V3)[36, 37] to account for the regions missing in the available mammalian prestin structures, particularly the IVS-loop. In addition, we generated structural models for ancestral and extant prestin representatives (Figure 2 and Supplementary Figure S4). The models closely resembled the domain-swapped dimeric structure of members of the SLC26A family (Figure 1 B) and showed high overall confidence, indicated by the high values of the predicted local distance difference test (pLDDT) observed for most of the well-folded part of the TMD and STAS domain (Figure 1 B, Figure S4 and Supplementary Figure S4). The flexible and disordered parts (N- and C-terminal, EC-loop,

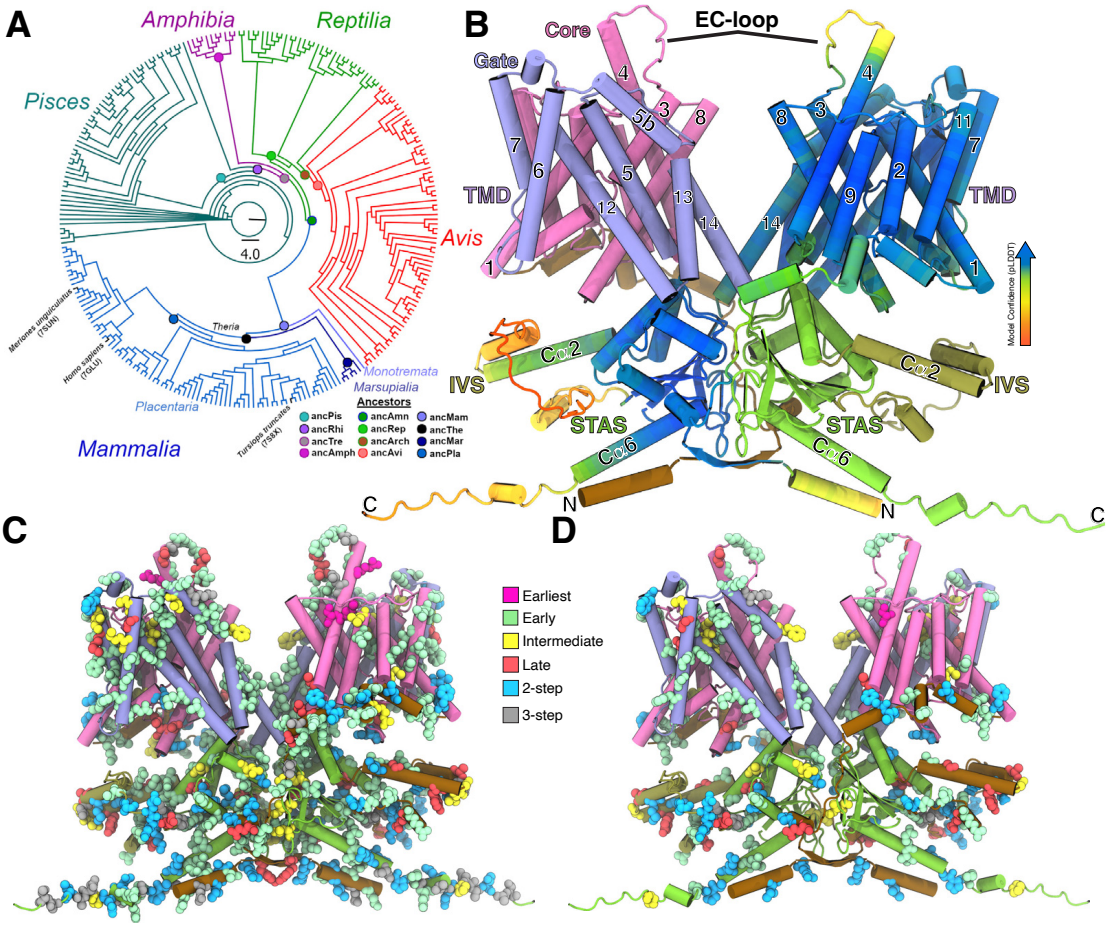

Figure 1. **Ancestral sequence reconstruction of SLC26A5 members**(A) Maximum likelihood phylogenetic tree of prestin (SLC26A5) from Chordata. Sequences colored as follow: Dark cyan Pisces, dark magenta Amphibian, green Reptilian, red Birds, light blue monotremata mammalians, navy blue marsupialia mammalians and bright navy blue placentaria mammalians. The extant sequences of *Homo sapiens (hsPres)*, *Meriones unguiculatus (muPres)* and *Tursiops truncates (ttPres)* are highlighted with their respective PDB accession codes. Circles indicate nodes where ancestral sequences were inferred in this study. (B) Structural model of full-lenght hsPres. Protein shown in cartoon representation with the monomer on the right labeled by AF2 confidence score (pLDDT) and the other by domain (N-terminal domain, ochre; core domain, pink; gate domain, violet: STAS domain, light-green and IVS, tan. (C-D) Mapping of the substitutions onto the hsPres model showing (C) all and (D) non-conservative substitutions. Domains colored as in B and substitutions colored as: earliest, magenta; early, lime; intermediate, light-yellow; late, red; 2-step light-blue; and 3-step, gray.

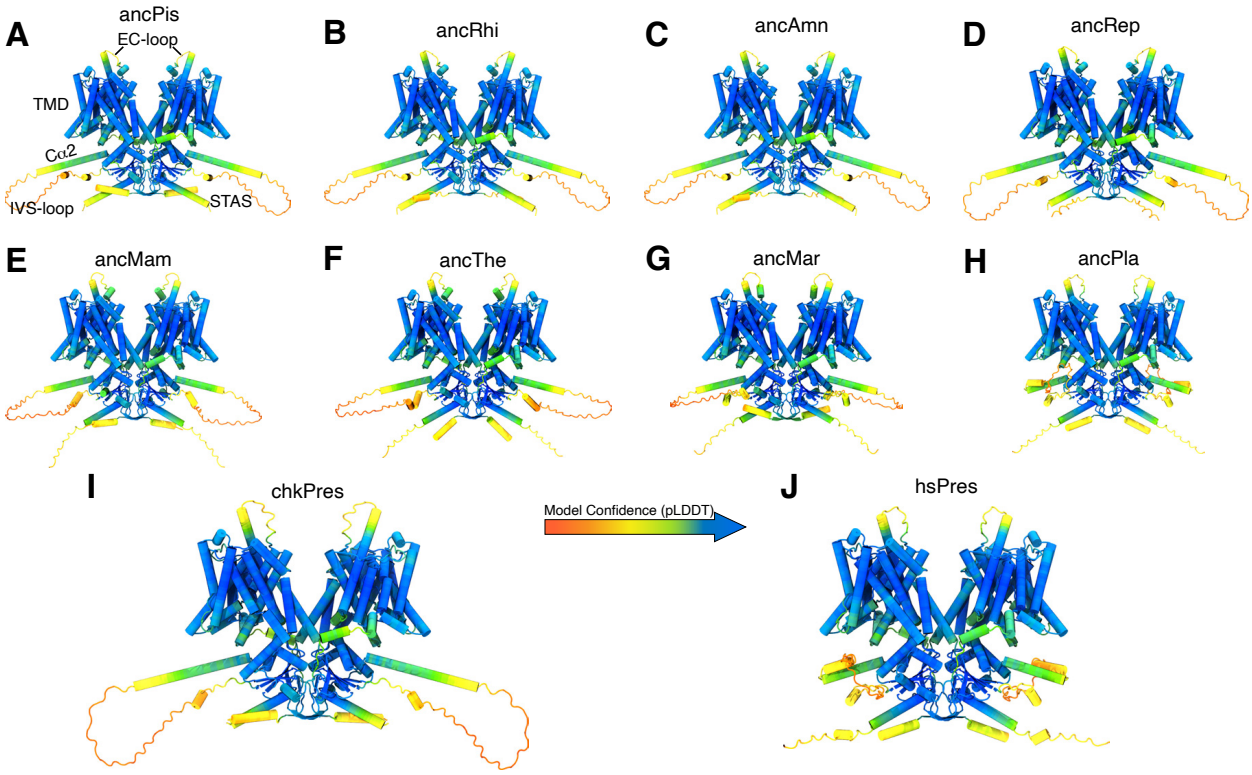

Figure 2. **Reconstruction of ancestral prestin structures using AF2-multimer(V3)**. Structural predictions of ancestors from Pisces (A), Rhipidistia (B), Amniota (C), Reptilia (D), Mammalia (E), Theria (F), Marsupialia (G), Placentaria (H), chicken prestin (I) and human prestin (J) shown in cartoon representation colored by confidence (pLDDT).

IVS-loop) show the lowest pLDTT values, indicating low confidence.

Since structural predictions of flexible regions such as the EC-loop and the IVS-loop show low confidence, the use of structural ensembles provides a broader scope of their functional relevance. Thus, we generated structural predictions of hsPres, prestin ancestors (ancAmn, ancRep, ancThe, ancMam, ancPla) and chicken prestin (chkPres) increasing their conformational variability by using multiple seeds and the *dropout* option (see Methods) of AF2-multimer(V3) as reported in[41]. This resulted in the generation of 1280 predictions for each sequence that explored the flexibility of the EC- and IVS-loop while showing minor variation in the TMD and folded parts of the STAS domain. However, it is important to note that, while this approach enhances conformational sampling, the resulting ensembles have to be interpreted as a qualitative exploration of possible conformations rather than a quantitative representation of a thermodynamic ensemble of populations.

Across all reconstructed prestin models, the transmembrane domain (TMD) exhibited high structural similarity (RMSD-C$\alpha \leq 3.0$ Å), with variation in the TM3-M4 loop (extracelular or EC-loop) and the TM6-TM7 loops. Similarly, for the structured part of the STAS domain. Strikingly, the only major structural divergence observed between non-mammalian and mammalian models (ancestral and extant) was in the conformation and orientation of the intervening sequence (IVS) loop.

In non-mammalians ancestors and extant representatives, the IVS consistently formed an extended Cα2 helix (Figure 2A-D and Supplementary Figure S4 A-I). In contrast, mammalian ancestors and extant representatives exhibited a distinct rearrangement of this region. Early mammalians and marsupials (ancMam, ancThe, platypus, opposum, wombat) showed a shorter Cα2 helix followed by a disordered IVS-loop (Figure 2 EF and Supplementary Figure S4 JL) while placental mammals (ancPla, human, gerbil, dolphin, and bat) showed a broken helix (Cα2 + Cα2') repositioning the IVS loop toward the intracellular entrance of the adjacent monomer (Figure 2 GH and Supplementary Figure S4 MH).

This consistent difference in conformation and orientation of the IVS-loop is correlated with the differences observed between sequences of non-mammalian and mammalian ancestors and between early mammalians, marsupials and placentals (Supplementary Figure S3 and Supplementary File 2). Furthermore, they are in line with the progression of prestin function from ion transport in non-mammals followed by intermediate area-motor adaptations in monotremes to robust, finely tuned electromotility in therian mammals[16, 42, 43]. Thus, the IVS-loop represents the most prominent structural difference distinguishing mammalian prestin from their ancestral and non-mammalian counterparts. We discuss the relationship between these differences and the substitutions that we found in that region in a later section of this work.

Mapping of substitutions into the structural models of hsPres and the reconstructed ancestors.

Mapping the substitutions on the hsPres model (Figures 1 CD and 4), we observed that the early substitutions are distributed throughout the TMD and STAS domains with non-conservative early substitutions accumulating in the TM5b, TM6, EC-loop and STAS do-

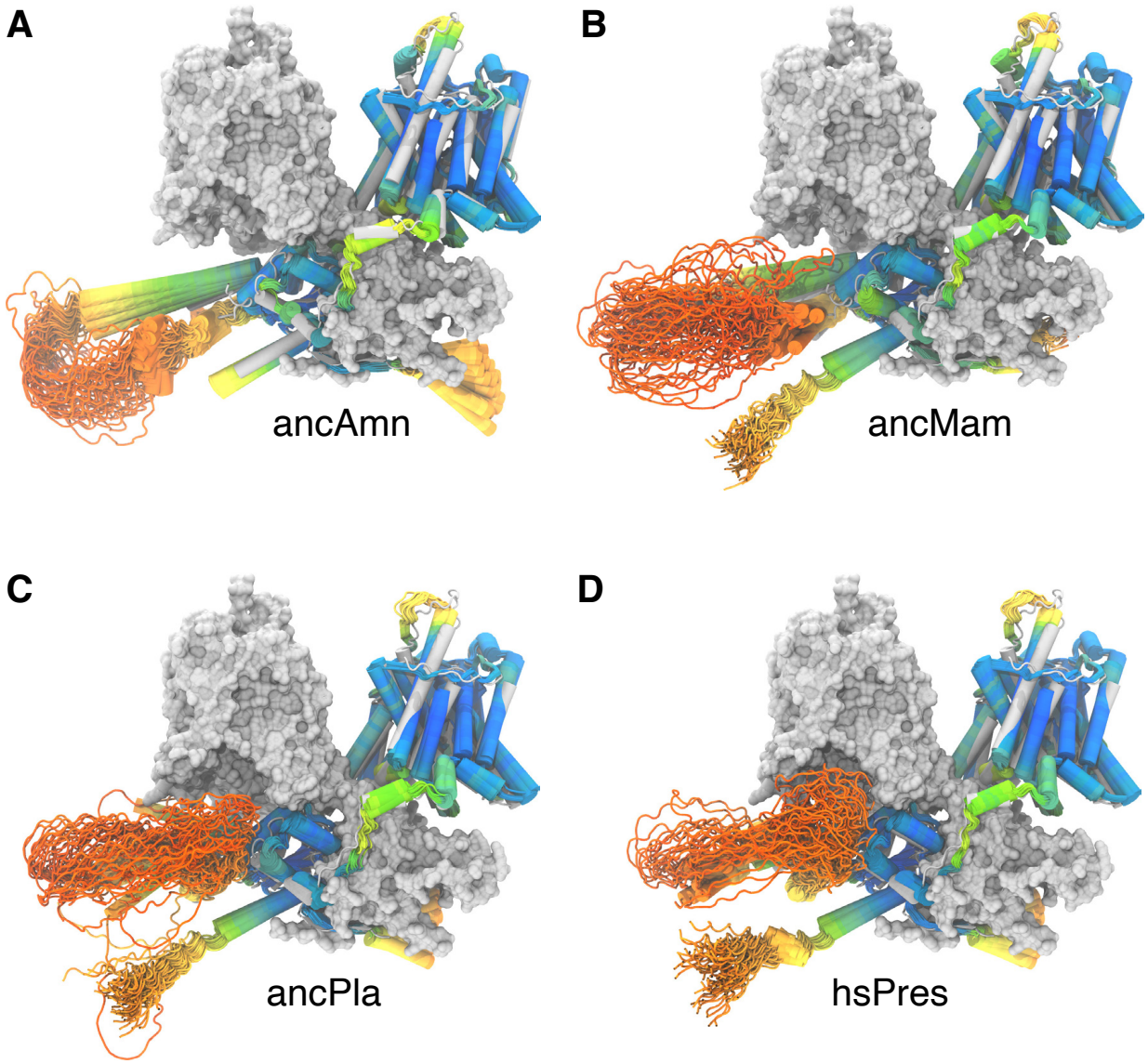

Figure 3. **Conformational sampling of prestin models**(A-D) Representative conformations obtained for human and ancestral prestin models. Only one monomer of the models is shown in cartoon representation colored as in Figure 2. The other monomer corresponds to hsPres structure 7LGU[27] shown as a gray surface.

mains (Figure 1 D, 4). Intermediate substitutions are found mainly in TM5b, TM6 and in the STAS domain. Late substitutions are found in the EC-loop, the TM5b-TM6 loop, and in the STAS domain. Finally, two- and three-step substitutions are found in the EC-loop and STAS domains, indicating positions with the least restriction to substitutions. In the following sections, we detail our analysis for each domain separately.

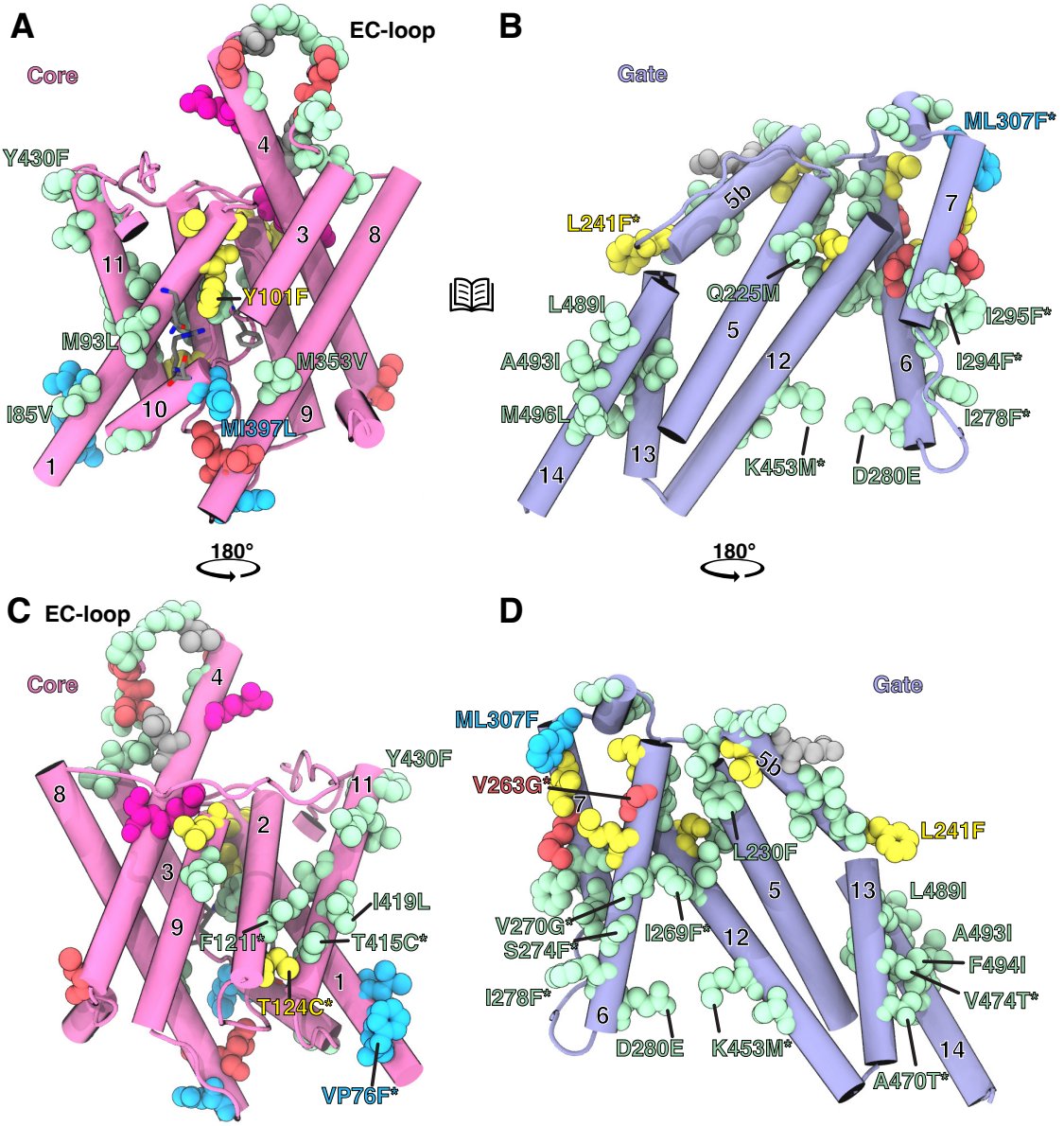

Figure 4. **Substitutions on the TMD**(A-B) Open-book view of the core (A) and gate (B) domain interface. (C-D) 180° rotation along the z-axis of the core domain (C) and gate domain (D) with respect to A and B respectively. In all panels the protein is shown in cartoon representation and substituted residues in VdW representation. The same color scheme as in Figure 1 BD was used. Two-step substitutions are represented as, for example, VP76F, with residue numbering based on human prestin (HsPres).

Substitutions to the TMD show clustering in the Gate Domain.

Almost 20% of the residues in the TMD (n=83) are substituted during the ancAmn →→ ancPla transition, with 9% corresponding to non-conservative substitutions. In our analysis, we found that the early and intermediate substitutions are located near and far from the chloride-binding site, whereas the late substitutions are positioned farther away from it. (Figure 1 CD and Figure 4). This suggests that early and intermediate changes, which are widespread across the transmembrane domain (TMD) and do not necessarily affect chloride binding, may have contributed to the loss of ion transport and the emergence of voltage sensitivity. However, late substitutions, mapped to a few clustered locations, may be responsible for the fine-tuning of prestin to provide the high sensitivity and selectivity observed in placental mammals[16].

Table II. Summary of TM Helices Substitutions

TM Helix	Substitutions	NC Substitutions	Evolutionary Step
TM1	8/29	5	2-step, early and intermediate
TM2	2/17	2	1 early, 1 intermediate
TM3	3/14	2	All early
TM4	5/27	4	3 earliest, 1 late, 1 3-step
TM5	2/23	2	Early
TM5b	7/11	4	4 Early, 2 intermediate, 1 3-step
TM6	11/23	10	7 Early, 2 intermediate, 2 late
TM7	9/15	6	6 early, 1 intermediate, 1 late, 1 2-step
TM8	4/35	4	1 early, 2 late, 1 2-step
TM9	2/17	1	1 early, 1 intermediate
TM10	2/9	2	1 early, 1 2-step
TM11	7/21	3	all early
TM12	5/27	5	4 early, 1 intermediate
TM13	4/15	4	all early
TM14	4/20	2	all early

Not all TM helices are equally substituted, and we found more substitutions in the gate domain (scaffold) than in the core domain (transport) with the highest number of substitutions found in TM5b, TM6 and TM12 (Figure4, II and Supplementary FigureS3). Furthermore, in both domains the majority of substitutions are located in the external face of the helix, with the exception of TM6 and TM7. In these helices, their lipid-facing region, which also

forms the interface between the two, concentrates all the substitutions (Figure 4 BD).

It has been shown that the function of prestin and other members of SLC26A involves changes in the orientation between the core and gate domains [24, 27], thus, we expected to find an accumulation of substitutions in their interface. However, as can be seen in Figure 4 AB, very few substitutions are found at the core-gate interface. From the core domain, we found M93L (TM1), Y101F (TM1), M353V (TM8) and MI397L (TM10) and from the Gate domain, we found only Q225M* (TM5). Some of these substitutions are located near the chloride binding site, which highlights their relevance for NLC/EM. For example, in TM1 substitutions M93L and Y101F (early) and **T136P*** (intermediate) when mutated individually from their mammalian to non-mammalian version (that is, L93M, F101Y, P136T) resulted in changes in the parameters of NLC and when mutated together completely abolished NLC/EM [20]. Furthermore, mutated V353 and L139 were related to changes in solvent accessibility to the chloride binding site[29].

Despite the lack of substitution on the core-gate interface, we found a network of substitutions located near the linkers between both domains between TM7-TM8 and TM11-TM12. These include a combination of early, intermediate, and late substitutions located on the extracellular side of TM5, TM5b, TM6, TM7, and TM12 in the gate domain and the extracellular side of TM11 in the core domain (Figure 4) and may have changed the coupling between domains and therefore the functional conformational change.

One of the helices in the TMD with the largest amount of substitutions is TM6, with the majority being non-conservative. This helix has been suggested to play a crucial role in NLC/EM due to the accumulation of glycines, in a unique pattern in mammalian prestin[24]. Our analysis showed that of the glycines found in placental prestin, two correspond to early substitutions and one to a late substitution (Figure 5 D). Furthermore, we found several substitutions in TM6 that could also be relevant for its flexibility since they participate in interactions with TM7 and TM12 (Figure 4 BD and Figure 5) and could restrict rotation or displacement of TM6. Furthermore, a comparison with the non-mammalian ancestor ancAmn shows that early substitutions D280E in TM6, and **K543M*** in TM12 perturb an electrostatic contact network between these helices (Figure 5), which would reduce the strength of the interaction between them and allow TM6 to move more freely. Other charged residues that are part of this network, in mammalian prestin, have been shown to affect

NLC/EM when mutated to glutamine[32], highlighting the role of the interaction of TM6 - TM12 in prestin function.

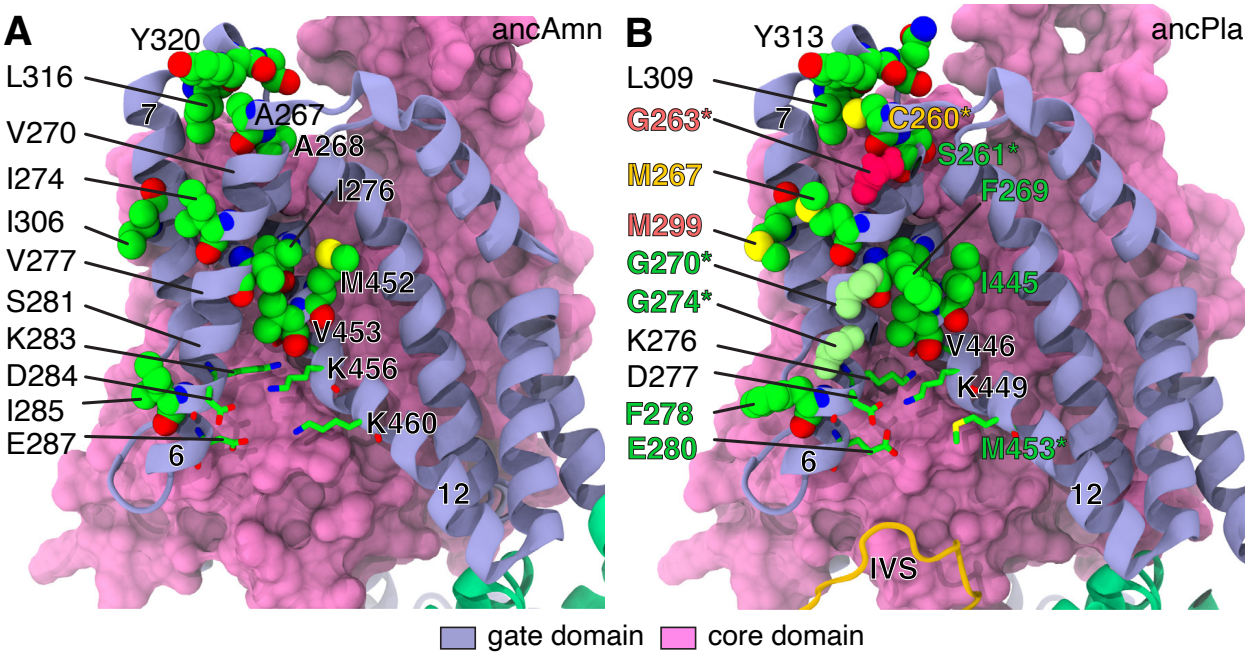

Figure 5. **Substitutions in the TMD**. (A-B) Substitutions in TM6 in the ancAmn (A) and ancPla (B) models. The gate domain is shown as ribbons and the core domain as surface colored as in 1. Substitution labels are colored as in 1C and shown in VdW or licorice representation colored by element (carbon, green; oxygen, red; nitrogen, blue and sulphur, yellow). Only the TMD of one monomer is shown for clarity. Non-conservative substitutions marked with a *.

Several Substitutions in the TMD affect protein-lipid interfase.

Mammalian prestin is known to be sensitive to changes in membrane tension[44, 45, 46], thickness [26, 47, 48], and cholesterol concentration[49, 50, 51] thus it called our attention that most of the substitutions found in the TMD are located on the lipid-facing side of the helices, particularly in the gate domain (Figure 4). This suggests that the evolutionary transition from transporter to area-motor may have involved significant changes in how prestin interacts with the membrane. Interestingly, previous studies have identified differences in the response to changes in membrane thickness between mammalian and non-mammalian prestin [26, 48], providing some support for our observations.

We identified two sets of non-conservative substitutions that may be the most relevant: The

first set comprises eight substitutions in which an apolar residue is replaced by phenylalanine (and one where a tyrosine is replaced by a phenylalanine). They are mainly located near the extracellular or intracellular end of an helix and concentrate in the gate domain (TM5b, TM6 and TM7) with only two found in the core domain (TM1 and TM11) (4). The second set of substitutions comprised four early non-conservative substitutions in TM13 that result in a unique pattern found only in mammalian prestin (**V470T***, **V474T***, **A475T*** and **A478S***). In our mammalian models, residues V474, S478, and T479 (residue numbering according to hsPres) are located on the external face of the helix, while T475 faces inward toward TM5. This results in three polar residues placed inside the lipid bilayer in the intermonomer region (Figure6). These residues could use their side chain hydroxyl groups to form intrahelix (TM13) and interhelix (TM13-TM14) hydrogen bonds, a characteristic that has previously been observed when these types of residue are located in transmembrane helices [52]. Furthermore, the presence of polar residues such as Ser or Thr in the lipid bilayer has been shown to affect the thickness of the membrane and the organization of lipids around the protein [52]. More importantly, these residues increase the polarity in the intermonomer region, facilitating the interaction with the poplar head-groups of lipids or cholesterol.

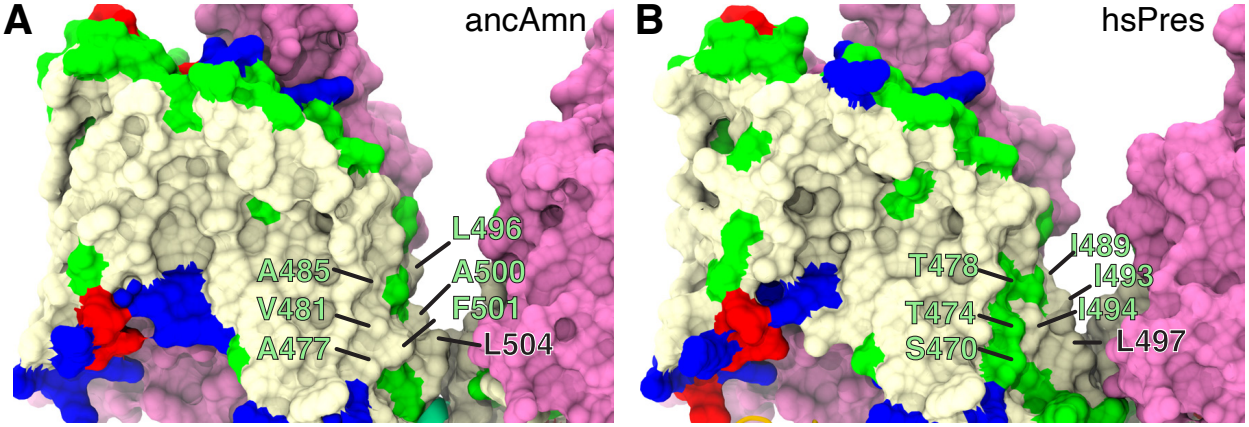

Figure 6. **Substitutions on the TM13-14**(A-B) (A-B) Surface representation with the core domain colored as in Figure 1B and the gate domain colored by residue type (hydrophobic = white, polar = green, acidic = red, basic = blue). (A) ancAmn (B) ancPla. The same color scheme as in Figure 1 (C) was used to label the substitutions.

In summary, we found several substitutions (mostly early) that suggest that one of the initial changes in mammalian prestin is related to its interaction with surrounding lipids. It is well known that substitutions in transmembrane helices that replace an apolar residue

with a phenylalanine or a threonine can affect the disposition of the TM helix or result in remodeling of protein-lipid interactions, membrane embedding and orientation, and helix-helix interactions and dynamics[53, 54, 55]. Furthermore, these changes not only alter the protein itself but also cause the surrounding lipid and protein environment to adapt, affecting protein stability, conformational equilibrium, and interactions. Thus, these substitutions could be associated with the membrane deformation around human prestin and its lipid-mediated organization[27, 30] and the differences in response between mammalian and non-mamalian prestin to changes in membrane thickness[26].

EC-loop configurations correlate with functional differences between non-mammalian and mammalian prestin.

In the TMD we also find the extracellular loop (EC-loop), which joins TM3 and TM4 (residues 151 to 170 in hsPres). It has been shown to provide pendrin (SLC26A4) with NLC / EM when its EC-loop is replaced by that of gerbil prestin[21], while other studies have proposed that it plays a key role in EM[22, 23]. We map five early, three late, and one 3-step substitutions in this region, a remarkable accumulation for a 20 residue loop (Figure 7A). As expected, substitutions resulting in a PGG motif produce different conformation and orientation of the loop between non-mammalian and mammalian ancestors (Figure **??**). In non-mammalian ancestors, the loop faces the extracellular surface of its monomer, facilitating the formation of interactions with residues from TM4, the TM5-TM5b loop, and TM5b. In mammalian ancestors, the loop points outwards, decoupling TM3 and the loop from other regions, a situation that is more marked for the placental ancestor ancPla and hsPres.

Conformational sampling further highlights these differences, as shown by the distinct distance distributions between the center of mass (COM) of the EC-loop and the COM of TM5b obtained from the models sampled in Figure 7 B): non-mammalians (ancAmn, ancRep, average distance $\sim$ 21 Å), early mamalianians (ancMam, ancThe, average distance $\sim$ 22 Å) and placentals (ancPla, hsPres, $\sim$ 24 Å).

Taken together, our results suggest that substitutions in the EC-loop alter its conformational flexibility and its capacity to interact with other regions of the TMD, particularly by

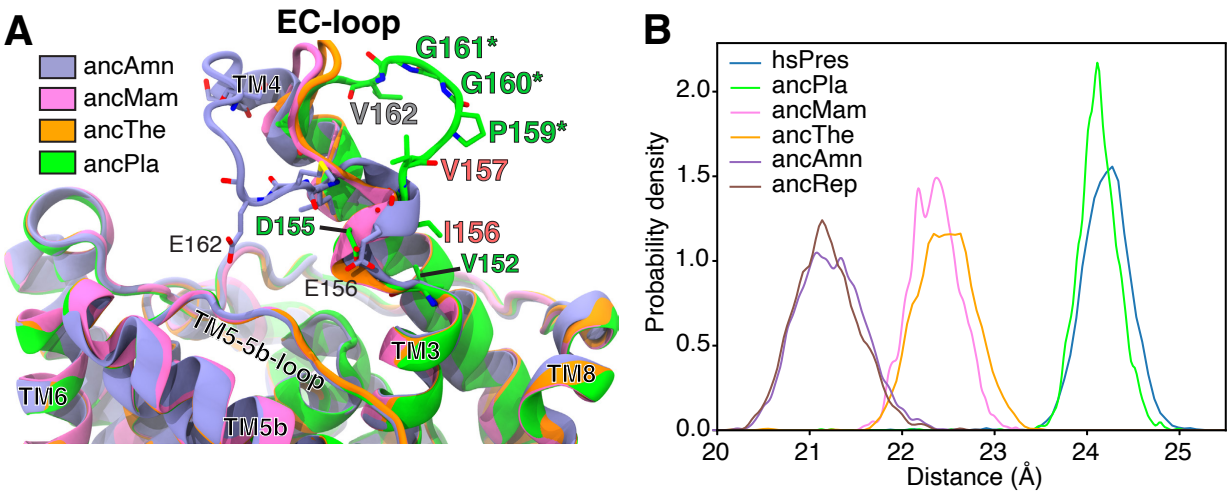

Figure 7. **Structural effect of substitutions in the EC-loop**. (A) Substitutions in the EC-loop. Proteins shown as ribbons. Substitution labels colored as in Figure 1B. Non-conservative substitutions marked with a *. (B) Distribution of the distance between the Center of Mass (COM) of the EC-loop and the COM TM5b obtained from the models AF2-conformational sampling models

decoupling it from the TM5-5b loop. However, our results do not allow for a mechanistic explanation of the role of the EC-loop on the generation of NLC/EM. Interestingly, TM5b and the TM5-TM5b loop present several early substitutions and its structural/functional role remains unexplored, so it is an interesting area for future research.

A Unique IVS-loop in Placental Mammalian Prestin results from accumulation of substitutions

In the STAS domain, the greatest accumulation of substitutions is found in the intervening sequence (IVS) that connects Cα1 to sheet Cβ4 (residues 561 to 614 according to hsPres) and contains helix Cα1 (8AB). This segment is highly variable across the SLC26A family, both in length and in amino acid composition[8].

In the IVS there is a flexible loop (IVS-loop) that connects helix Cα2 with sheet Cβ4 (residues 580 to 614 according to hsPres) that is unresolved in all currently available prestin structures and in most of the members of the SLC26 family. The only exception is a set of SLC26A3 cryo-EM strucutres (PDB-ID: 7XUH, 7XUJ, 7XUL, and 8IET) in which the IVS loop is resolved and positioned near the intracellular entry of the adjacent monomer

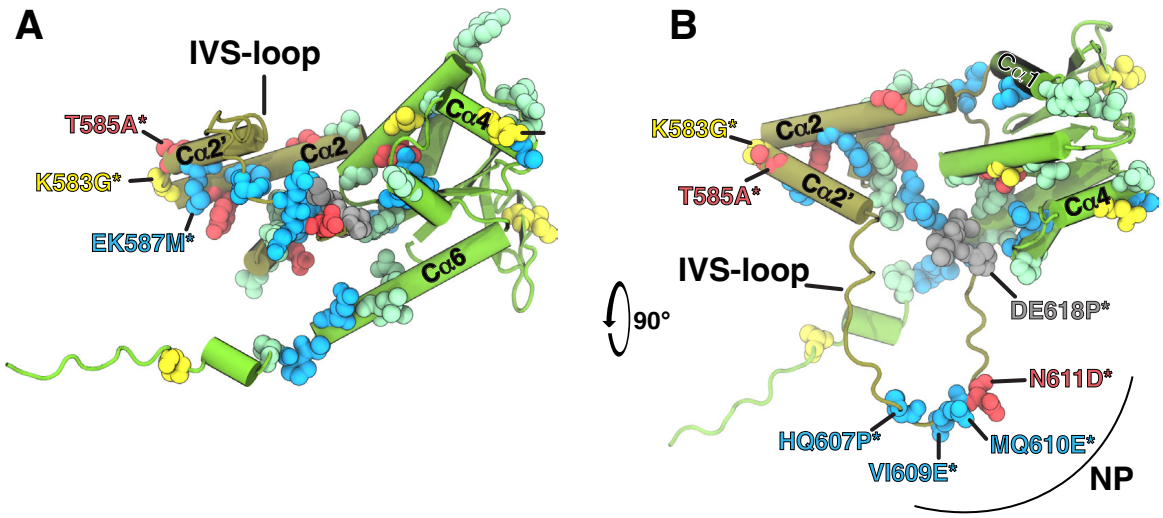

Figure 8. **Substitutions in the IVS** (A-B) Detail of the STAS domain showing only the non-conservative substitutions, highlighting substitutions found in the IVS-loop and the negative patch (NP). (A) Side view and (B) Top view (90° rotation of A along the x-axis). In all panels the protein is shown in cartoon representation and substituted residues in VdW representation. The same color scheme as in Figure 1 (BD) was used.

(FigureS5). Unfortunately, there is no publication associated with these structures.

Despite its variability, it has received limited attention in studies of the function and regulation of prestin and other members of the family. Some reports suggest potential roles in protein-protein interactions with regulators such as calmodulin[56, 57] or CTFR [58, 59]. One study found that mutations in positively and negatively charged residues within the mammalian prestin IVS loop caused modest changes in voltage operating parameters[35]. More recently, the IVS-loop in SLC26A9 has been implicated in modulating intracellular chloride dynamics[60].

In this region, we found 30 substitutions that result in clear sequence differences between non-mammalian, early mammalian and placental ancestors (Figure 9A and Supplementary Figure S3), which include the elimination of a positively charged patch (PP in Figure 9 A) and the appearance of a negatively charged patch in placental mammals (NP in Figure 9A). These differences also translate into structural differences in our models. Mammalian ancestors (ancThe onward) present the non-conservative intermediate substitution **K583G*** which result in a broken Cα2 in two helical segments: Cα2 and Cα2' as observed in the

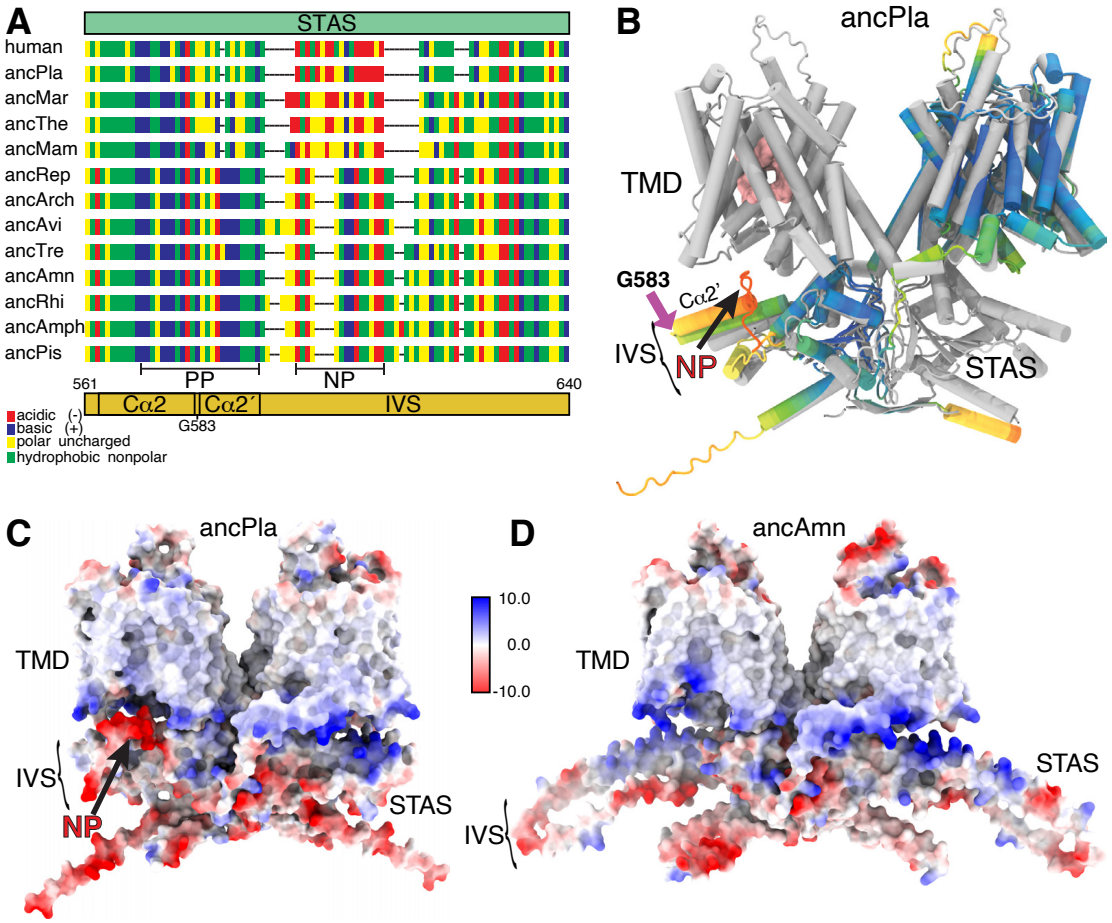

Figure 9. **Detail of the unique features of the placental mammal IVS**. (A) Fingerprint sequence alignment of the reconstructed ancestors and hsPres. Residues colored by hydropathy (acidic, red; basic, blue, polar uncharged, yellow and hydrophobic nonpolar, green). NP and PP refer to negative and positive patch, respectively. (B) Superposition of a monomer of the ancPla model (colored by pLDDT as in Figure 2) onto the structure of human prestin 7LGU[27] (grey) shown in cartoon representation. (C-D) Electrostatic potential maps of ancPla (B) and ancAmn (C).

ancPla and hsPres models (Figure 2 (GH and J), 16 BE, Figure 9 B) (and other mammalian representatives; Supplementary Figure S4 MM). This characteristic is more pronounced in the ancPla and hsPres models (see Figure 9 B), perhaps due to substitutions at positions 584 (3-step), 585 (late) and 587 (2-step), resulting in the NANMANA motif that has a higher helix propensity. After Cα2', we find Pro607 (two-step) and, more notably, three two-step substitutions that replace hydrophobic or polar residues with negatively charged amino acids (Glu609, Glu610, and Asp611) (16 BE) increasing the number of negatively charged residues resulting in a negative patch (NP) formed by residues Glu609, Glu610, and Asp611 (Figure 9 A). This combination of substitutions results in the placement of the NP near the entrance of the intracellular (IC) cavity of the adjacent monomer (Figure 9B-D). The NP is adjacent to the PIVKSTF motif, which is unique to placentals and originated from a series of two- and three-step substitutions. This motif forms contacts with residues from helix Cα3 and with a group of substitutions in helix Cα5 (Pro696 (early non-conservative, S to P), Ala697 (two-step, T to S to A), Leu698 (two-step, T to V to L)) that increase the hydrophobicity of this helix (16BE). Thus, changes that lead to the formation of Cα2' and NP are accompanied by changes in other regions of the STAS domain that could stabilize this unique conformation.

Conformational sampling showed that non-mammalian ancestors present a continuous and extended Cα2 positioning of the IVS-loop far from the intracellular (IC) entrance of the adjacent monomer. In contrast, mammalian models, especially those representing placental mammals (ancPla and hsPres), show a distinct break in Cα 2, allowing the IVS-loop to adopt conformations closer to or farther from the entry of the IC (Figure 10 CD). To quantify this, we measured the distance between the center of mass (COM) of the IVS-loop and the COM of the chloride binding site of the adjacent monomer (Figure 10 F) and the helical content in the IVS of the models (Figure 10 G). Our results identified two distinct structural states in placental mammals (bimodal distribution), contrasting with earlier-mammalian and non-mammalian models that predominantly retain a single "far" conformation characterized by a continuous Cα2 helix. Our findings suggest that intermediate and late substitutions significantly altered the IVS-loop in mammalian ancestors, introducing a structural hinge near residue G583 (Figure 10 G) and a unique negatively charged patch (NP) close to the chloride entry/exit point of the adjacent monomer. Although previous studies have

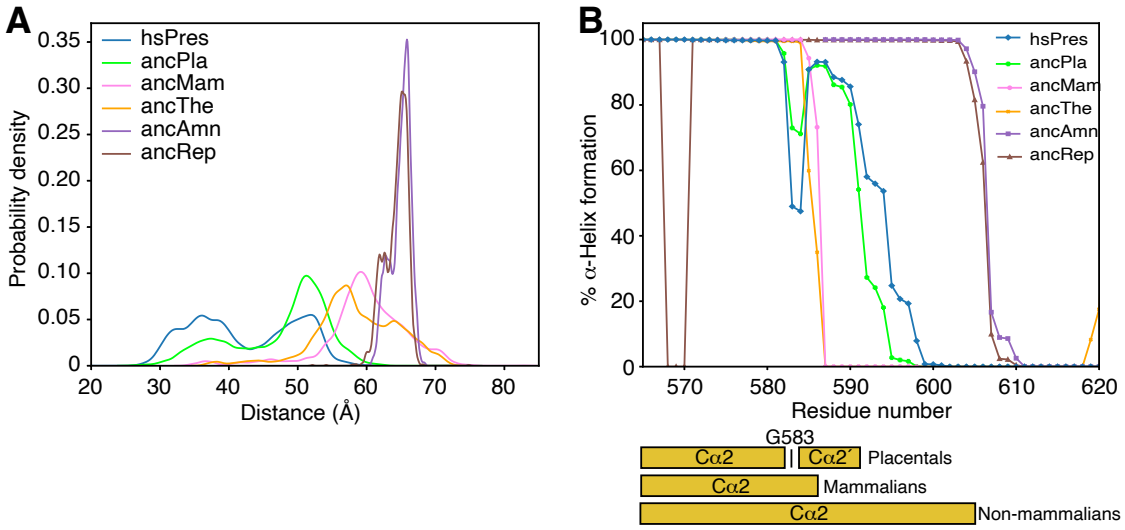

Figure 10. **Conformational and secondary structure analysis of the IVS-loop**(A) Distribution of the distance between the COM of the IVS-loop of one monomer and the chloride-binding site of the adjacent monomer obtained from the models. (B) Percentage of alpha-helix conformation for each residue in the IVS of models. The plot show the percentage of models in which each residue is assigned an alpha-helix conformation by DSSP [61]. The dark yellow blocks below the panel highlight the length of the helix for each prestin group (Placental, Mammalians and Non-mammalians) and the location of residue G583. The residue numbers correspond to hsPres.

indicated that these charged residues slightly influence nonlinear capacitance (NLC) [35], their precise functional role remains unclear. However, the distinct conformation of the placental IVS-loop raises the possibility that it could modulate chloride ion dynamics by restricting ion exit or trapping ions near the binding site, potentially influencing voltage sensing or allosteric coupling - an idea supported by recent findings in SLC26A9 that point in a similar direction[60].

In summary, the IVS-loop of the STAS domain stands out as a highly diverse and lineage-specific region that underwent substantial remodeling during prestin evolution. Structural and sequence changes in placental mammals, including the formation of a hinge at G583, the emergence of a negatively charged patch, and associated substitutions in neighboring helices, appear to support a unique conformation near the intracellular entrance of the adjacent protomer. Although the precise functional consequences remain to be determined, our results suggest that these changes could contribute to fine-tuning of the voltage sensitivity and

chloride accessibility, highlighting the IVS-loop as a potential regulatory element in the electromechanical behavior of prestin.

A structural and evolutionary model for the emergence of area-motor function

Our findings recapitulate the progression of prestin function from ion transport in non-mammals through intermediate area-motor adaptations in monotremes to robust, finely tuned electromotility and corresponding frequency selectivity in therian mammals by a distributed remodeling of its architecture as it transitioned from a passive anion transporter to an active area-motor protein.

Although the overall scaffold of the TMD and STAS domains remained conserved, key substitutions introduced new physicochemical properties that collectively support the voltage-driven mechanical output.

In the TMD, substitutions altered membrane-facing surfaces and interhelical contacts. Aromatic and polar residues became more prevalent at the protein-lipid interface, potentially enhancing membrane anchoring and deformation.

The most striking innovations occurred in the IVS-loop. This flexible element underwent structural rearrangements in placental mammals, adopting a conformation that places a negatively charged patch near the ion access site, a feature that may directly modulate chloride access or voltage sensitivity.

Together, these changes support an evolutionary scenario in which the area-motor function emerged through fine-tuning of existing elements within a conserved scaffold. This mosaic of adaptations may be the basis for the ability of prestin to convert voltage into mechanical work with extraordinary speed and precision.

CONCLUSIONS

Our findings provide a structural and evolutionary framework for understanding how prestin acquired its unique area-motor function in mammals. By integrating ancestral sequence reconstruction with structural modeling and comparative analysis, we reveal that the transformation of prestin from a canonical anion transporter into a voltage-sensitive motor protein

involved a distributed set of molecular innovations.

Rather than arising from a single defining mutation or domain acquisition, the functional shift of prestin was driven by the accumulation of substitutions that modulate membrane embedding, helix packing, and interdomain communication. These include changes to lipid-facing residues in the transmembrane domain and reorientation of flexible loop regions.

This stepwise repurposing of conserved structural elements offers a model for how novel functional traits can evolve through subtle refinement of existing architectures, illustrating how exaptation operates at the molecular level[62]. Our work highlights the broader principles of functional innovation in membrane proteins.

METHODS

Multiple sequence alignment, phylogenetic inference, and ancestral sequence reconstruction

For the multiple sequence alignment (MSA) of chordate prestin, sequences were retrieved from the non-redundant protein database (nr) using BLASTp, with hsPres (Uniprot accession: P58743-1) as reference. Sequences exhibiting an identity greater than 98% were filtered through CD-HIT[63], while those with lengths greater than 870 and less than 570 residues were excluded. The resulting set of sequences (n = 248, see the sequence list in Supplementary Tables S6to S9) was aligned with a structural profile derived from the Cryo-EM structures of hsPres ([27]; PDB-ID: 7GLU), ttPres ([24]; PDB-ID: 7S8X) and muPres ([28];PDB-ID: 7SUN) using ClustalW[64] and STAMP[65]. Evolutive models were selected using SMS2.0[66], followed by inference of maximum likelihood (ML) phylogenetic trees using PhyML[67]. Ancestral sequences were inferred using the empirical Bayes method with PAML 4.4[68], employing the Lazarus tool as described by Handson-Smith[69]. Probability calculations for each amino acid were calculated at each alignment position for each target node, selecting the amino acids with the highest probabilities. Gap positions were inferred via machine learning (ML), using the discrete character model in Mesquite[70] as previously reported[71].

Generation of structural models

Full-length dimeric structural predictions of ancestral prestin sequences together with human prestin (hsPres), chicken (chkPres) and other representatives of mammalian and non-mammalian prestin (see Supplementary Figure S5) were generated using AlphaFold2-multimer(V3) [36, 37] within the colabfold-batch implementation 1.5.5[72]. We used five recycles and five ensembles to produce five models for each sequence. The model with the best ranking after relaxation was selected for further analysis. Templates were not used in this process.

Conformational sampling with AF2

We enhanced the conformational sampling of AlphaFold2-multimer(V3) using multiple seeds and the *dropout* option that introduces epistemic uncertainty and leads to a wider exploration of potential structures by the model[41]. We used 256 seeds with 1 recycle and 1 ensemble to generate 5 models per seed for a total of 1280 models per sequence. All models were considered for analysis.

Data analysis and visualization

The analysis of structural models was performed using VMD[73] plugins and custom Tcl and Python scripts. Distance measurements for intramonomer distances as shown as a combination of the two monomers. The intermonomer distances are shown as combinations of distances A-B and B-A.

The assignment of a secondary structure to follow the conformations of Cα2 was done with DSSP [61]. Molecular renderings were performed with VMD.

DATA AVAILABILITY

All the outpus from the ASR and AF2 structure predictions can be found at DOI:10.5281/zenodo.15197370

ACKNOWLEDGMENTS

This work is supported by the following grants: ANID FONDECYT Regular 1231164 and ANID PAI 77200112 (RAS),ANID FONDECYT Regular 1221260 (JAG), ANID FONDECYT De Iniciación 11221268 (ARF), ANID FONDECYT Regular 1221667 (VCF), ANID Beca Doctorado Nacional 21221449 (NFU) and Financiamiento Basal para Centros Cientificos y Tecnologicos de Excelencia de ANID" Centro Ciencia& Vida, FB210008 (to Fundación Ciencia & Vida). The authors thank Felipe Engelberger for his helpful advice and discussion.

REFERENCES

---

[1] William E. Brownell, Charles R. Bader, Daniel Bertrand, and Yves De Ribaupierre. Evoked mechanical responses of isolated cochlear outer hair cells. *Science*, 227:194–196, 1 1985. ISSN 00368075. doi:10.1126/science.3966153. URL `http://www.ncbi.nlm.nih.gov/pubmed/3966153`.

[2] J. F. Ashmore. A fast motile response in guinea-pig outer hair cells: the cellular basis of the cochlear amplifier. *The Journal of Physiology*, 388:323–347, 7 1987. ISSN 14697793. doi:10.1113/jphysiol.1987.sp016617. URL `https://pubmed.ncbi.nlm.nih.gov/3656195/`.

[3] Peter Dallos and Bernd Fakler. Prestin, a new type of motor protein. *Nature Reviews Molecular Cell Biology*, 2002. doi:10.1038/nrm730.

[4] P. Dallos, X. Wu, M. A. Cheatham, J. Gao, J. Zheng, C. T. Anderson, S. Jia, X. Wang, W. H. Cheng, S. Sengupta, D. Z. He, and J. Zuo. Prestin-based outer hair cell motility is necessary for mammalian cochlear amplification. *Neuron*, 58(3):333–339, 2008. doi:10.1016/j.neuron.2008.02.028.

[5] Jonathan Ashmore. Cochlear outer hair cell motility. *Physiological Reviews*, 88:173–210, 1 2008. ISSN 0031-9333. doi:10.1152/physrev.00044.2006. URL `http://physrev.physiology.org/cgi/doi/10.1152/physrev.00044.2006`.

[6] Jing Zheng, Weixing Shen, David Z.Z. He, Kevin B. Long, Laird D. Madison, and Peter Dallos. Prestin is the motor protein of cochlear outer hair cells. *Nature*, 405:149–155, 5 2000. ISSN

00280836. doi:10.1038/35012009. URL http://www.ncbi.nlm.nih.gov/pubmed/10821263.

[7] Peter Dallos, Jing Zheng, and Mary Ann Cheatham. Prestin and the cochlear amplifier, 10 2006. ISSN 00223751.

[8] David B. Mount and Michael F. Romero. The slc26 gene family of multifunctional anion exchangers. *Pflugers Archiv European Journal of Physiology*, 447:710–721, 2 2004. ISSN 00316768. doi:10.1007/s00424-003-1090-3. URL http://www.ncbi.nlm.nih.gov/pubmed/12759755.

[9] Dominik Oliver, Dominik Oliver, David Z.Z. He, David Z.Z. He, Nikolaj Klöcker, Nikolaj Klöcker, Nikolaj Klöcker, Jost Ludwig, Jost Ludwig, Uwe Schulte, Uwe Schulte, Siegfried Waldegger, Siegfried Waldegger, J. P. Ruppersberg, J. P. Ruppersberg, Peter Dallos, Peter Dallos, Bernd Fakler, Bernd Fakler, and Bernd Fakler. Intracellular anions as the voltage sensor of prestin, the outer hair cell motor protein. *Science*, 292(5525):2340–2343, June 2001. doi:10.1126/science.1060939.

[10] Xiaodong Tan, Jason L. Pecka, Jie Tang, Oseremen E. Okoruwa, Qian Zhang, Kirk W. Beisel, and David Z.Z. He. From zebrafish to mammal: Functional evolution of prestin, the motor protein of cochlear outer hair cells. *Journal of Neurophysiology*, 2011. doi:10.1152/jn.00234.2010.

[11] Thorsten J Schaechinger and Dominik Oliver. Nonmammalian orthologs of prestin (slc26a5) are electrogenic divalent/chloride anion exchangers. *Proceedings of the National Academy of Sciences of the United States of America*, 104:7693–8, 5 2007. ISSN 0027-8424. doi:10.1073/pnas.0608583104. URL http://www.ncbi.nlm.nih.gov/pubmed/17442754http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1863495.

[12] Jost Ludwig, Dominik Oliver, Gerhard Frank, Nikolaj Klöcker, Anthony W. Gummer, and Bernd Fakler. Reciprocal electromechanical properties of rat prestin: The motor molecule from rat outer hair cells. *Proceedings of the National Academy of Sciences*, 2001. doi:10.1073/pnas.071613498.

[13] Stephen J. Rossiter, Shuyi Zhang, and Yang Liu. Prestin and high frequency hearing in mammals. *Communicative Integrative Biology*, 2011. doi:10.4161/cib.4.2.14647.

[14] Lucía F. Franchini and A. Belén Elgoyhen. Adaptive evolution in mammalian proteins involved in cochlear outer hair cell electromotility. *Molecular Phylogenetics and Evolution*, 41:622–635, 12 2006. ISSN 10557903. doi:10.1016/j.ympev.2006.05.042. URL https://pubmed.ncbi.

nlm.nih.gov/16854604/.

[15] Oseremen E. Okoruwa, Michael D. Weston, Divvya C. Sanjeevi, Amanda R. Millemon, Bernd Fritzsch, Richard Hallworth, and Kirk W. Beisel. Evolutionary insights into the unique electromotility motor of mammalian outer hair cells. *Evolution Development*, 2008. doi:10.1111/j.1525-142x.2008.00239.x.

[16] Liu Z, Li GH, Huang JF, Murphy RW, and Shi P. Hearing aid for vertebrates via multiple episodic adaptive events on prestin genes. *Mol Biol Evol*, 29(9):2187–98, Sep 2012. doi: 10.1093/molbev/mss087.

[17] M. Schänzler and C. Fahlke. Anion transport by the cochlear motor protein prestin. *The Journal of Physiology*, 590:259–272, 2011. doi:10.1113/jphysiol.2011.209577.

[18] P. Mistrík, N. Daudet, K. Morandell, and J. F. Ashmore. Mammalian prestin is a weak cl-/hco3 - electrogenic antiporter. *Journal of Physiology*, 590:5597–5610, 11 2012. ISSN 00223751. doi:10.1113/jphysiol.2012.241448. URL http://www.ncbi.nlm.nih.gov/pubmed/22890707http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3528980.

[19] Oseremen E. Okoruwa, Michael D. Weston, Divvya C. Sanjeevi, Amanda R. Millemon, Bernd Fritzsch, Richard Hallworth, and Kirk W. Beisel. Evolutionary insights into the unique electromotility motor of mammalian outer hair cells. *Evolution & Development*, 10:300–315, 5 2008. ISSN 1520541X. doi:10.1111/j.1525-142X.2008.00239.x. URL http://doi.wiley.com/10.1111/j.1525-142X.2008.00239.x.

[20] T. J. Schaechinger, D. Gorbunov, C. R. Halaszovich, T. Moser, S. Kügler, B. Fakler, and D. Oliver. A synthetic prestin reveals protein domains and molecular operation of outer hair cell piezoelectricity. *EMBO Journal*, 30(14):2793–2804, 2011. doi:10.1038/emboj.2011.202.

[21] Jie Tang, Jason L. Pecka, Xiaodong Tan, Kirk W. Beisel, and David Z.Z. He. Engineered pendrin protein, an anion transporter and molecular motor. *Journal of Biological Chemistry*, 286:31014–31021, 9 2011. ISSN 00219258. doi:10.1074/jbc.M111.259564. URL http://www.ncbi.nlm.nih.gov/pubmed/21757707http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3162460.

[22] Xiaodong Tan, Jason L. Pecka, Jie Tang, Sándor Lovas, Kirk W. Beisel, and David Z.Z. He. A motif of eleven amino acids is a structural adaptation that facilitates motor capability of eutherian prestin. *Journal of Cell Science*, 2012. doi:10.1242/jcs.097337.

[23] Makoto Kuwabara, Koichiro Wasano, Satoe Takahashi, Justin Bodner, Tomotaka Komori, Tomotaka Komori, Tomotaka Komori, Sotaro Uemura, Jing Zheng, Tomohiro Shima, and Kazuaki Homma. The extracellular loop of pendrin and prestin modulates their voltage-sensing property. *Journal of Biological Chemistry*, 293(26):9970–9980, June 2018. doi: 10.1074/jbc.ra118.001831.

[24] Navid Bavi, Michael Clark, Gloria Contreras, Rong Shen, Bharat Reddy, Wieslawa Milewski, and Eduardo Perozo. The conformational cycle of prestin underlies outer-hair cell electromotility. *Nature*, 2021. doi:10.1038/s41586-021-04152-4.

[25] B Currall, D Rossino, H Jensen-Smith, and R Hallworth. The roles of conserved and non-conserved cysteinyl residues in the oligomerization and function of mammalian prestin. *J Neurophysiol*, 106(5):2358–67, Nov 2011. doi:10.1152/jn.00496.2011.

[26] C Izumi, JE Bird, and KH Iwasa. Membrane thickness sensitivity of prestin orthologs: the evolution of a piezoelectric protein. *Biophys J*, 100(11):2614–22, Jun 2011. doi: 10.1016/j.bpj.2011.04.032.

[27] J Ge, J Elferich, S Dehghani-Ghahnaviyeh, Z Zhao, M Meadows, H von Gersdorff, E Tajkhorshid, and E Gouaux. Molecular mechanism of prestin electromotive signal amplification. *Cell*, 184(18):4669–4679.e13, 2021. ISSN 0092-8674. doi:10.1016/j.cell.2021.07.034. URL `https://www.sciencedirect.com/science/article/pii/S0092867421008824`. Epub 2021 Aug 13.

[28] Carmen Butan, Qiang Song, Jun-Ping Bai, Winston J T Tan, Dhasakumar Navaratnam, and Joseph Santos-Sacchi. Single particle cryo-EM structure of the outer hair cell motor protein prestin. *Nat. Commun.*, 13(1):290, January 2022.

[29] Makoto F. Kuwabara. Elevator-like movements of prestin mediate outer hair cell electromotility. *Nature Communications*, 2023. doi:10.1038/s41467-023-42489-8.

[30] Sepehr Dehghani-Ghahnaviyeh, Zhiyu Zhao, and Emad Tajkhorshid. Lipid-mediated prestin organization in outer hair cell membranes and its implications in sound amplification. *Nature Communications*, 13(1):6877, November 2022. ISSN 2041-1723. doi:10.1038/s41467-022-34596-9. URL `https://doi.org/10.1038/s41467-022-34596-9`.

[31] Volodymyr Rybalchenko and Joseph Santos-Sacchi. Anion control of voltage sensing by the motor protein prestin in outer hair cells. *Biophysical Journal*, 2008. doi: 10.1529/biophysj.108.134197.

[32] Jun Ping Bai, Alexei Surguchev, Simone Montoya, Peter S. Aronson, Joseph Santos-Sacchi, and Dhasakumar Navaratnam. Prestin's anion transport and voltage-sensing capabilities are independent. *Biophysical Journal*, 96:3179–3186, 4 2009. ISSN 15420086. doi: 10.1016/j.bpj.2008.12.3948. URL `http://www.ncbi.nlm.nih.gov/pubmed/19383462http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2718310`.

[33] Lei Song, Lei Song, Joseph Santos-Sacchi, and Joseph Santos-Sacchi. Disparities in voltage-sensor charge and electromotility imply slow chloride-driven state transitions in the solute carrier SLC26a5. *Proceedings of the National Academy of Sciences of the United States of America*, 110(10):3883–3888, March 2013. doi:10.1073/pnas.1218341110.

[34] Dmitry Gorbunov, Mattia Sturlese, Florian Nies, Murielle Kluge, Massimo Bellanda, Roberto Battistutta, and Dominik Oliver. Molecular architecture and the structural basis for anion interaction in prestin and SLC26 transporters. *Nature Communications*, 5(1):3622–3622, April 2014. doi:10.1038/ncomms4622.

[35] J. P. Bai, D. Navaratnam, H. Samaranayake, and J. Santos-Sacchi. En block c-terminal charge cluster reversals in prestin (slc26a5): effects on voltage-dependent electromechanical activity. *Neuroscience Letters*, 404(3):270–275, Sep 2006. doi:10.1016/j.neulet.2006.05.062.

[36] J. Jumper et al. Highly accurate protein structure prediction with alphafold. *Nature*, 596: 583–589, Aug 2021.

[37] R. Evans et al. Protein complex prediction with alphafold-multimer. *bioRxiv*, 2021. doi: 10.1101/2021.10.04.463034v1. Preprint.

[38] S.E. Evans. Evolution and phylogeny of amniotes. In M.D. Binder, N. Hirokawa, and U. Windhorst, editors, *Encyclopedia of Neuroscience*. Springer, Berlin, Heidelberg, 2009. doi: 10.1007/978-3-540-29678-2_3113.

[39] A. Meyer and R. Zardoya. Recent advances in the (molecular) phylogeny of vertebrates. *Annual Review of Ecology, Evolution, and Systematics*, 34(1):311–338, 2003.

[40] GKA Hochberg and JW Thornton. Reconstructing ancient proteins to understand the causes of structure and function. *Annual Review of Biophysics*, 46:247–269, 2017. doi: 10.1146/annurev-biophys-070816-033631.

[41] Björn Wallner. AFsample: improving multimer prediction with AlphaFold using massive sampling. *Bioinformatics*, 39(9):btad573, 09 2023. ISSN 1367-4811. doi:10.1093/bioinformatics/btad573. URL `https://doi.org/10.1093/bioinformatics/`

btad573.

[42] Xiaodong Tan, Jason L. Pecka, Jie Tang, Oseremen E. Okoruwa, Qian Zhang, Kirk W. Beisel, and David Z.Z. He. From Zebrafish to Mammal: Functional Evolution of Prestin, the Motor Protein of Cochlear Outer Hair Cells. *Journal of Neurophysiology*, 105(1):36–44, January 2011. doi:10.1152/jn.00234.2010.

[43] David Z.Z. He, Sándor Lovas, Yu Ai, Yi Li, and Kirk W. Beisel. Prestin at year 14: Progress and prospect. *Hearing Research*, 311:25–35, May 2014. doi:10.1016/j.heares.2013.12.002.

[44] K.H. Iwasa. Effect of stress on the membrane capacitance of the auditory outer hair cell. *Biophysical Journal*, 65(1):492–498, 1993. ISSN 0006-3495. doi: https://doi.org/10.1016/S0006-3495(93)81053-1. URL https://www.sciencedirect.com/science/article/pii/S0006349593810531.

[45] Joseph Santos-Sacchi, Weixing Shen, Jing Zheng, and Peter Dallos. Effects of membrane potential and tension on prestin, the outer hair cell lateral membrane motor protein. *The Journal of Physiology*, 531(3):661–666, 2001. doi:https://doi.org/10.1111/j.1469-7793.2001.0661h.x. URL https://physoc.onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-7793.2001.0661h.x.

[46] Joseph Santos-Sacchi and Winston Tan. Complex nonlinear capacitance in outer hair cell macro-patches: effects of membrane tension. *Scientific Reports*, 10(1), 2020. doi: 10.1038/s41598-020-63201-6. URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-85083235474&doi=10.1038%2fs41598-020-63201-6&partnerID=40&md5=ea8b7e9d127c64a13724d45c1b6ed492.

[47] Jing Fang, Chisako Izumi, and Kuni H. Iwasa. Sensitivity of prestin-based membrane motor to membrane thickness. *Biophysical Journal*, 98(12):2831–2838, 2010. doi: 10.1016/j.bpj.2010.03.034.

[48] Chisako Izumi, Jonathan E. Bird, and Kuni H. Iwasa. Membrane thickness sensitivity of prestin orthologs: The evolution of a piezoelectric protein. *Biophysical Journal*, 2011. doi: 10.1016/j.bpj.2011.04.032.

[49] R. I. Kamar, L. E. Organ-Darling, and R. M. Raphael. Membrane cholesterol strongly influences confined diffusion of prestin. *Biophysical Journal*, 103:1627–1636, 10 2012. ISSN 00063495. doi:10.1016/j.bpj.2012.07.052. URL http://dx.doi.org/10.1016/j.bpj.2012.07.052https://pubmed.ncbi.nlm.nih.gov/23083705/.

[50] Lavanya Rajagopalan, Jennifer N. Greeson, Anping Xia, Haiying Liu, Angela Sturm, Robert M. Raphael, Amy L. Davidson, John S. Oghalai, Fred A. Pereira, and William E. Brownell. Tuning of the outer hair cell motor by membrane cholesterol. *Journal of Biological Chemistry*, 282:36659–36670, 12 2007. ISSN 00219258. doi:10.1074/jbc.M705078200. URL `http://www.jbc.org`.

[51] John Sfondouris, Lavanya Rajagopalan, Fred A. Pereira, and William E. Brownell. Membrane composition modulates prestin-associated charge movement*. *Journal of Biological Chemistry*, 283(33):22473–22481, 2008. ISSN 0021-9258. doi:https://doi.org/10.1074/jbc.M803722200. URL `https://www.sciencedirect.com/science/article/pii/S0021925820747794`.

[52] Esther S. Brielle and Isaiah T. Arkin. Quantitative analysis of multiplex h-bonds. *Journal of the American Chemical Society*, 2020. doi:10.1021/jacs.0c04357.

[53] Kelsey A. Sparks, Nicholas J. Gleason, Renetra Gist, Rebekah Langston, Denise V. Greathouse, and Roger E. II Koeppe. Comparisons of interfacial phe, tyr, and trp residues as determinants of orientation and dynamics for gwalp transmembrane peptides. *Biochemistry*, 53 (22):3637–3645, 2014. doi:10.1021/bi500439x. URL `https://doi.org/10.1021/bi500439x`. PMID: 24829070.

[54] Coral del Val, Stephen H. White, and Ana-Nicoleta Bondar. Ser/thr motifs in transmembrane proteins: Conservation patterns and effects on local protein structure and dynamics. *The Journal of Membrane Biology*, 245(11):717–730, 2012. doi:10.1007/s00232-012-9452-4. URL `https://doi.org/10.1007/s00232-012-9452-4`.

[55] Shyam S. Krishnakumar and Erwin London. The control of transmembrane helix transverse position in membranes by hydrophilic residues. *Journal of Molecular Biology*, 374 (5):1251–1269, 2007. ISSN 0022-2836. doi:https://doi.org/10.1016/j.jmb.2007.10.032. URL `https://www.sciencedirect.com/science/article/pii/S0022283607013708`.

[56] Jacob P. Keller, Kazuaki Homma, Chongwen Duan, Jing Zheng, Mary Ann Cheatham, and Peter Dallos. Functional regulation of the slc26-family protein prestin by calcium/calmodulin. *Journal of Neuroscience*, 2014. doi:10.1523/jneurosci.4020-13.2014.

[57] Satoe Takahashi and Kazuaki Homma. The molecular principles underlying diverse functions of the SLC26 family of proteins. *J. Biol. Chem.*, 300(5):107261, May 2024.

[58] Kazuaki Homma, Kazuaki Homma, Katharine Miller, Katharine K. Miller, Charles T. Anderson, Charles T. Anderson, Charles T. Anderson, Charles T. Anderson, Soma Sengupta,

Soma Sengupta, Guo Guang Du, Guo Guang Du, Salvador Aguiñaga, Salvador Aguinaga, Mary Ann Cheatham, Mary Ann Cheatham, Peter Dallos, Peter Dallos, Jing Zheng, and Jing Zheng. Interaction between cftr and prestin (slc26a5). *Biochimica et Biophysica Acta*, 2010. doi:10.1016/j.bbamem.2010.02.001.

[59] Carol A. Bertrand, Shalini Mitra, Sanjay Mishra, Xiaohui Wang, Yu Zhao, Joseph M. Pilewski, Dean R. Madden, and Raymond A. Frizzell. The cftr trafficking mutation f508del inhibits the constitutive activity of slc26a9. *Ajp Lung Cellular and Molecular Physiology*, 2017. doi: 10.1152/ajplung.00178.2016.

[60] An-Ping Chen, Heather Holmes, James W. Decker, Min-Hwang Chang, and Michael F. Romero. Disorder within order: identification of the disordered loop of stas domain as the inhibitory domain in slc26a9 chloride channel. *Journal of Biological Chemistry*, page 108145, 2024. ISSN 0021-9258. doi:https://doi.org/10.1016/j.jbc.2024.108145. URL `https://www.sciencedirect.com/science/article/pii/S0021925824026474`.

[61] Wolfgang Kabsch and Christian Sander. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22(12):2577–2637, 1983. doi:https://doi.org/10.1002/bip.360221211. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/bip.360221211`.

[62] Jürgen Brosius. Exaptation at the molecular genetic level. *Science China Life Sciences*, 62 (4):437–452, 2019. doi:10.1007/s11427-018-9447-8.

[63] Y. Huang, B. Niu, Y. Gao, L. Fu, and W. Li. Cd-hit suite: A web server for clustering and comparing biological sequences. *Bioinformatics*, 26(5):680–682, 2010. doi: 10.1093/bioinformatics/btq003.

[64] J. D. Thompson, D. G. Higgins, and T. J. Gibson. Clustal w: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22(22):4673–4680, 1994. doi: 10.1093/nar/22.22.4673.

[65] R. B. Russell and G. J. Barton. Multiple protein sequence alignment from tertiary structure comparison: Assignment of global and residue confidence levels. *Proteins: Structure, Function, and Bioinformatics*, 14(2):309–323, 1992. doi:10.1002/prot.340140216.

[66] V. Lefort, J. E. Longueville, and O. Gascuel. Sms: Smart model selection in phyml. *Molecular Biology and Evolution*, 34(9):2422–2424, 2017. doi:10.1093/MOLBEV/MSX149.

[67] S. Guindon, J. F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and O. Gascuel. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of phyml 3.0. *Systematic Biology*, 59(3):307–321, 2010. doi:10.1093/sysbio/syq010.

[68] Z. Yang. Paml 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, 24(8):1586–1591, 2007. doi:10.1093/molbev/msm088.

[69] V. Hanson-Smith, B. Kolaczkowski, and J. W. Thornton. Robustness of ancestral sequence reconstruction to phylogenetic uncertainty. *Molecular Biology and Evolution*, 27(9):1988–1999, 2010. doi:10.1093/molbev/msq081.

[70] W. P. Maddison and D. R. Maddison. Mesquite: a modular system for evolutionary analysis. Version 3.81, 2023. URL `http://www.mesquiteproject.org`.

[71] Victor Castro-Fernandez, Ariel Herrera-Morande, Rodrigo Zamora, Fernando Merino, Felipe Gonzalez-Ordenes, Fernando Padilla-Salinas, Herika M. Pereira, José Brandão-Neto, Richard C. Garratt, and Victor Guixe. Reconstructed ancestral enzymes reveal that negative selection drove the evolution of substrate specificity in adp-dependent kinases. *Journal of Biological Chemistry*, 292(38):15598–15610, Sep 2017. doi:10.1074/jbc.M117.790865. Erratum in: J Biol Chem. 2017 Dec 22;292(51):21218.

[72] M. Mirdita et al. Colabfold: making protein folding accessible to all. *Nature Methods*, 19: 679–682, May 2022.

[73] William Humphrey, Andrew Dalke, and Klaus Schulten. VMD – Visual Molecular Dynamics. *Journal of Molecular Graphics*, 14:33–38, 1996. doi:10.1016/0263-7855(96)00018-5.

[74] Justin D. Walter, Justin D. Walter, Marta Sawicka, Marta Sawicka, Raimund Dutzler, and Raimund Dutzler. Cryo-EM structures and functional characterization of murine Slc26a9 reveal mechanism of uncoupled chloride transport. *eLife*, 8, July 2019. doi:10.7554/elife.46986.

[75] Sudhir Kumar, Michael Suleski, Jack M Craig, Adrienne E Kasprowicz, Maxwell Sanderford, Michael Li, Glen Stecher, and S Blair Hedges. Timetree 5: An expanded resource for species divergence times. *Molecular Biology and Evolution*, 39(8):msac174, 08 2022. ISSN 1537-1719. doi:10.1093/molbev/msac174. URL `https://doi.org/10.1093/molbev/msac174`.

Supplementary Material

Reconstruction of prestin ancestors sheds light on the evolution of its area-motor activity

Nicolas Fuentes-Ugarte*

Departamento de Biologia, Facultad de Ciencias, Universidad de Chile, Santiago,Chile.

Tiaren Ruiz-Rojas*

Facultad de Ingenieria, Tecnologia y Diseño, Universidad San Sebastian, Santiago, Chile

Alvaro Ruiz-Fernandez

Computational Biology Lab, Centro Cientifico y Tecnologico de Excelencia, Fundacion Ciencia & Vida, Santiago, Chile, Santiago, Chile.

Jose Antonio Garate

Facultad de Ingenieria, Tecnologia y Diseño, Universidad San Sebastian, Santiago, Chile

Victor Castro-Fernandez[†]

Departamento de Biologia, Facultad de Ciencias, Universidad de Chile, Santiago,Chile.

Raul Araya-Secchi[‡]

Facultad de Ingenieria, Tecnologia y Diseño, Universidad San Sebastian, Santiago, Chile

Centro Cientifico y Tecnologico de Excelencia, Fundacion Ciencia & Vida, Santiago, Chile

*These authors contributed equally to this work

[†]Corresponding author: vcasfe@uchile.cl

[‡]Corresponding author: raul.araya@uss.cl

SUPPLEMENTARY DISCUSSION

Mapping the substitutions to the STAS domain.

The cytoplasmic STAS domain is a structural hallmark of the SLC26A family, and since the publication of the first structure[74] it has been considered an important factor for the movement of the TMD during functional conformational change. This domain is structurally conserved between members of the prestin and SLC26 family and constitute 15% of the total buried interface between prestin protomers through STAS–STAS interactions, while forming extensive contacts with the opposing N-terminal domain, reinforcing the importance of STAS in dimer stabilization. It has also been suggested to act as a platform for binding of regulatory proteins such as CFTR and calmodulin[56, 58, 59].

In the STAS domain, the substitutions are broadly distributed as shown in Supplementary Figure 16A. However, focusing only on the non-conservative substitutions reveals some clustering (Figure 16BE). Some substitutions can be assigned to regions of interaction between adjacent STAS domains, while others could affect the interaction between the STAS domain and its own or adjacent TMD.

However, the lack of experimental mutational studies or comparisons between the STAS domain of mammalian an non-mammalian prestin or among the SLC26A family difficult the interpretation of our findings.

SUPPLEMENTARY FIGURES

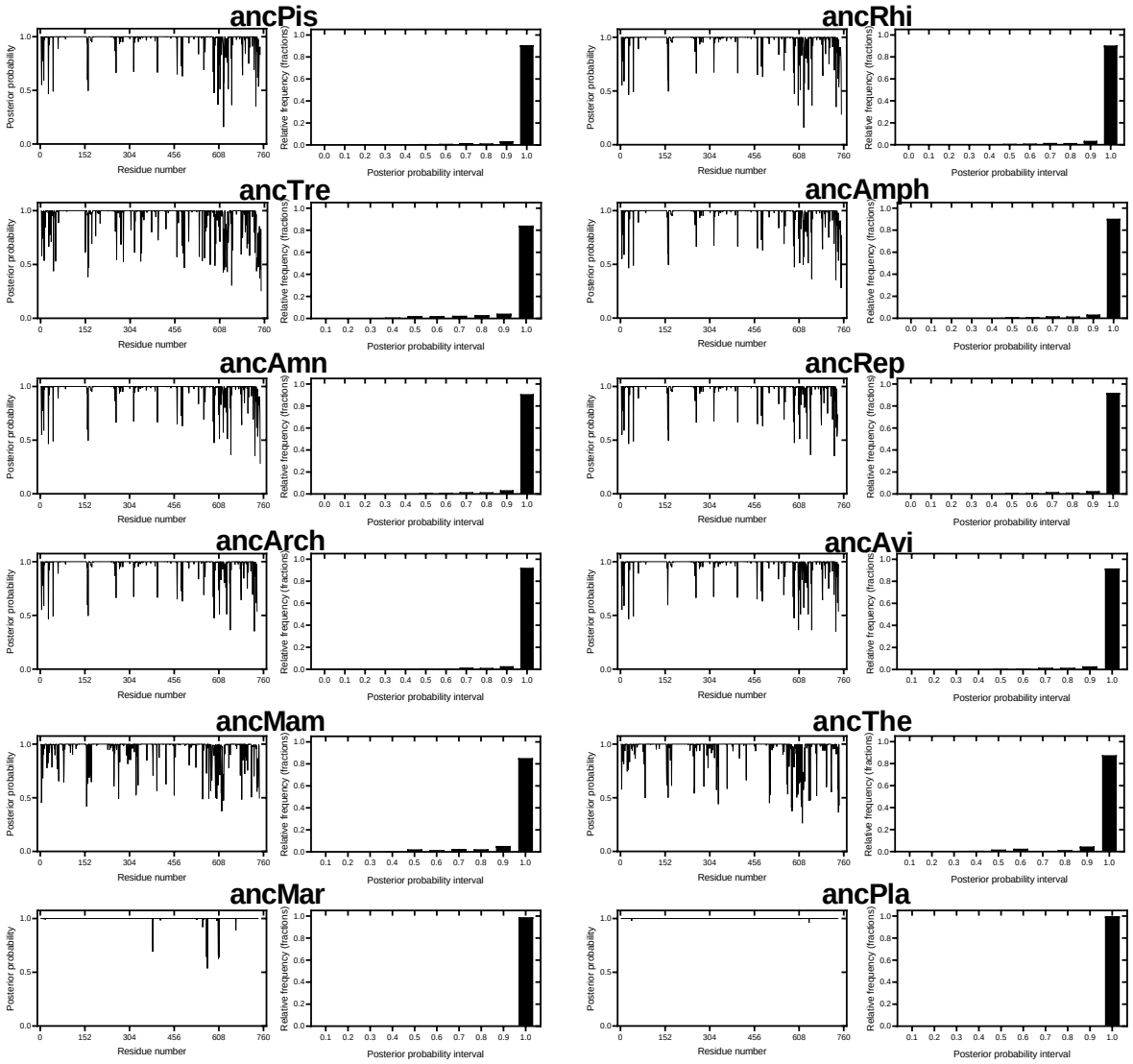

Figure S1. **Robustness of Ancestral Sequence Reconstruction.** The posterior probability for each sequence position (residue number) is shown on the right, while the posterior probability distribution for each ancestor is displayed on the left. Both the mean and median posterior probabilities exceed 0.95.

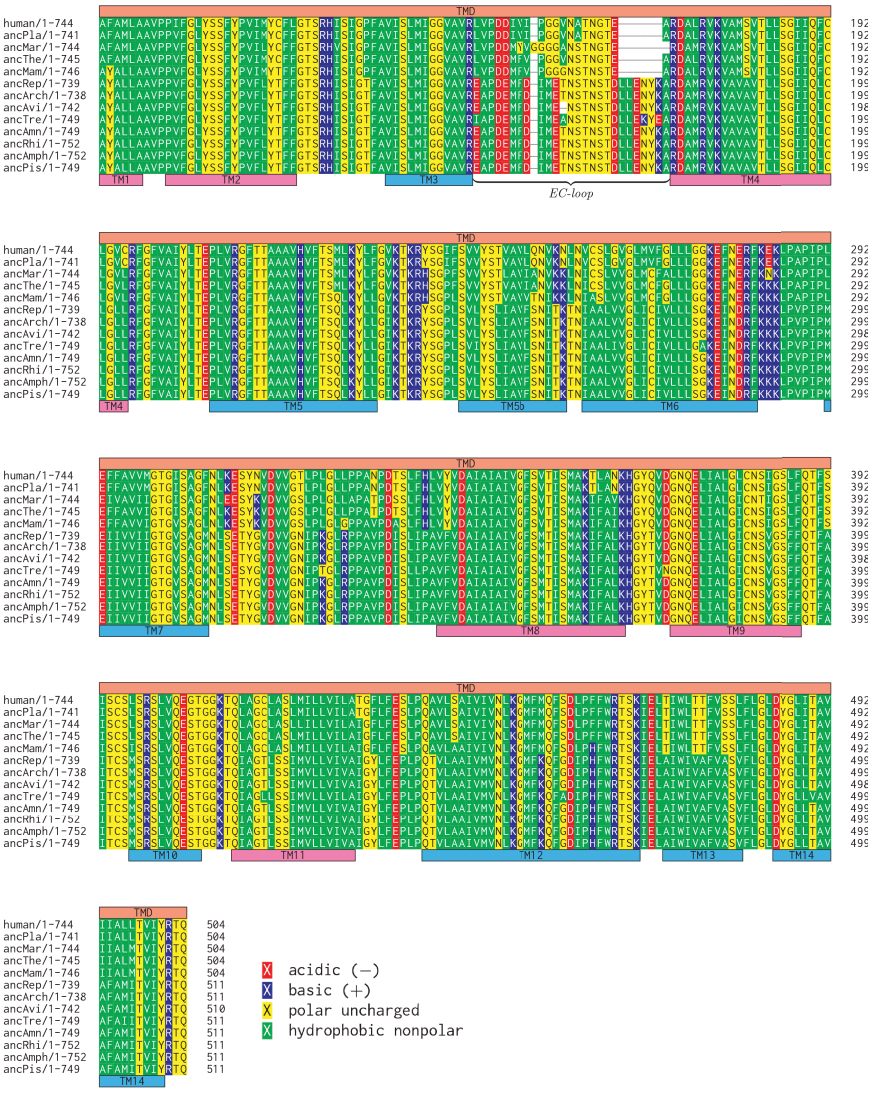

Figure S2. **Ancestral prestin sequences - TMD**Sequence alignment of the transmembrane domain (TMD) from inferred ancestral prestin sequences, along with human prestin (hsPres). Transmembrane (TM) helices are shown as boxes below the alignment. Helices from the transport domain (core) are depicted in magenta, and those from the scaffold domain (gate) are shown in light blue. The extracellular (EC) loop connecting TM3 and TM4 is indicated by a curly bracket. The positions of the TM helices and the EC-loop are based on the 7LGU human prestin structure. Residues are color-coded according to their hydropathy.

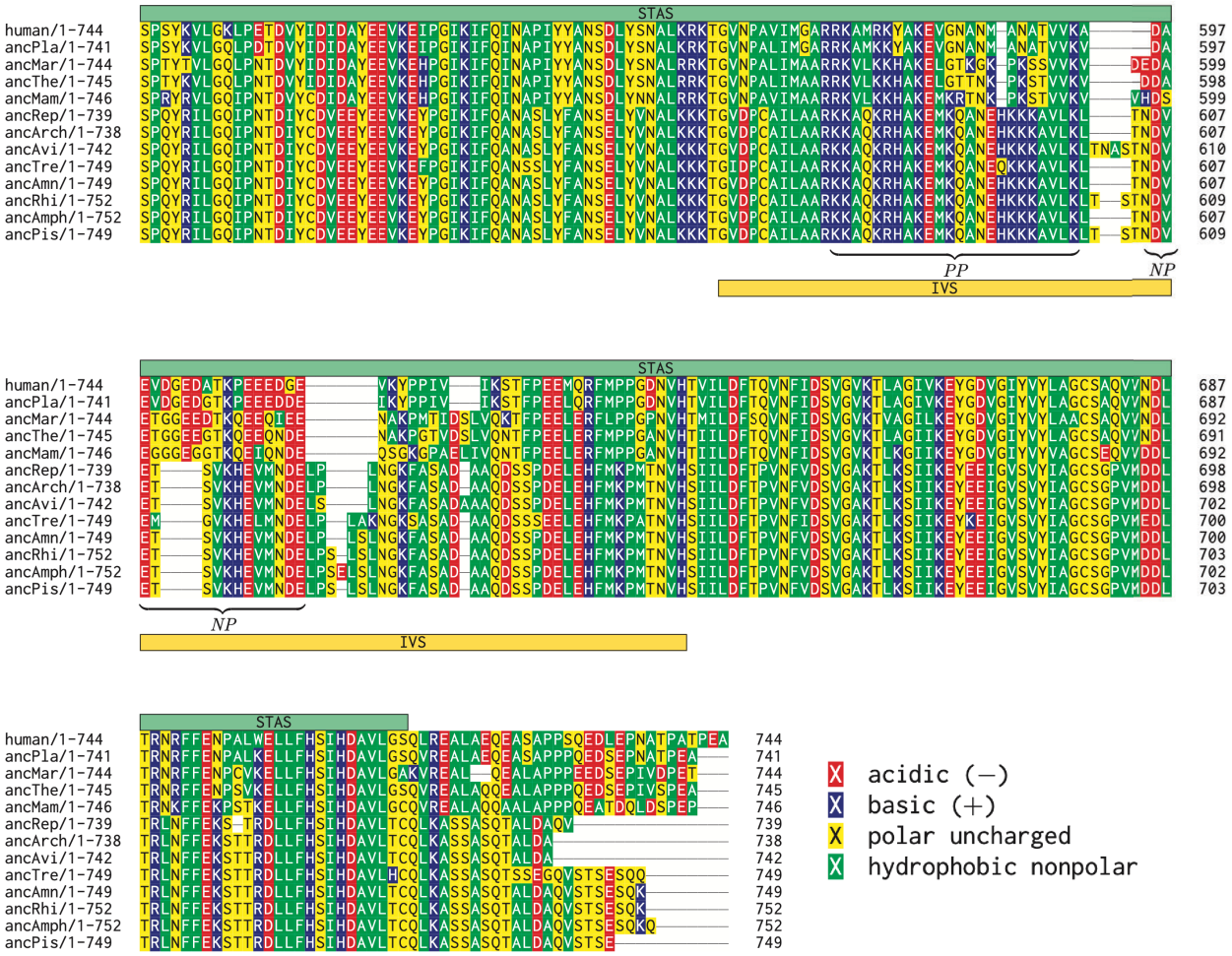

Figure S3. **Ancestral prestin sequences - STAS** Sequence alignment of the STAS domain from inferred ancestral prestin sequences, along with human prestin (hsPres). The IVS is represented as a golden box below the alignment. The positive patch (PP) and the negative patch (NP) are indicated by curly brackets. Residues are color-coded according to their hydropathy.

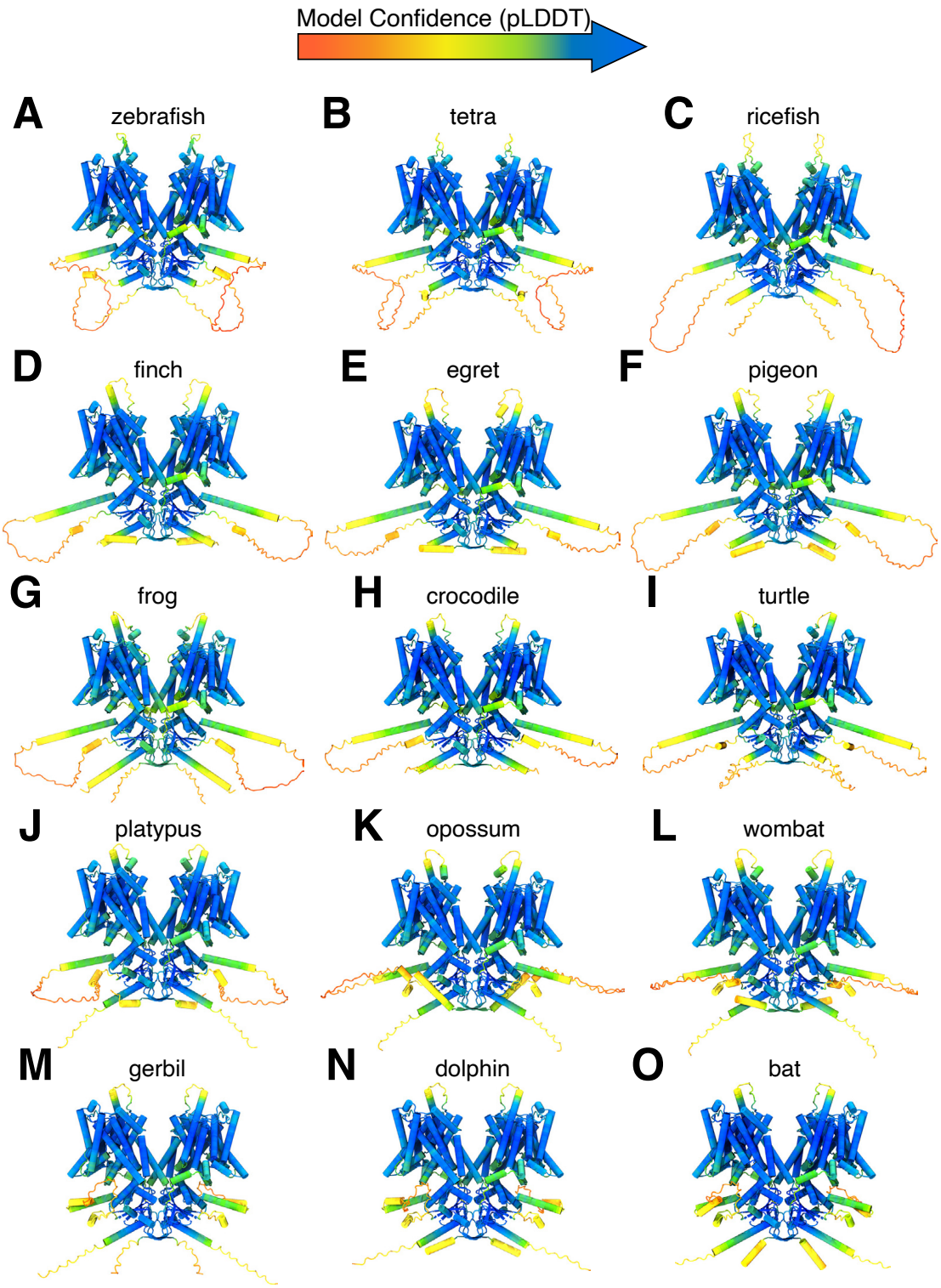

Figure S4. **Reconstruction of non-mammalian and mammalian representative prestin structures using AF2-multimer(V3)**. Models shown in cartoon representation and colored by model confidence (pLDDT).

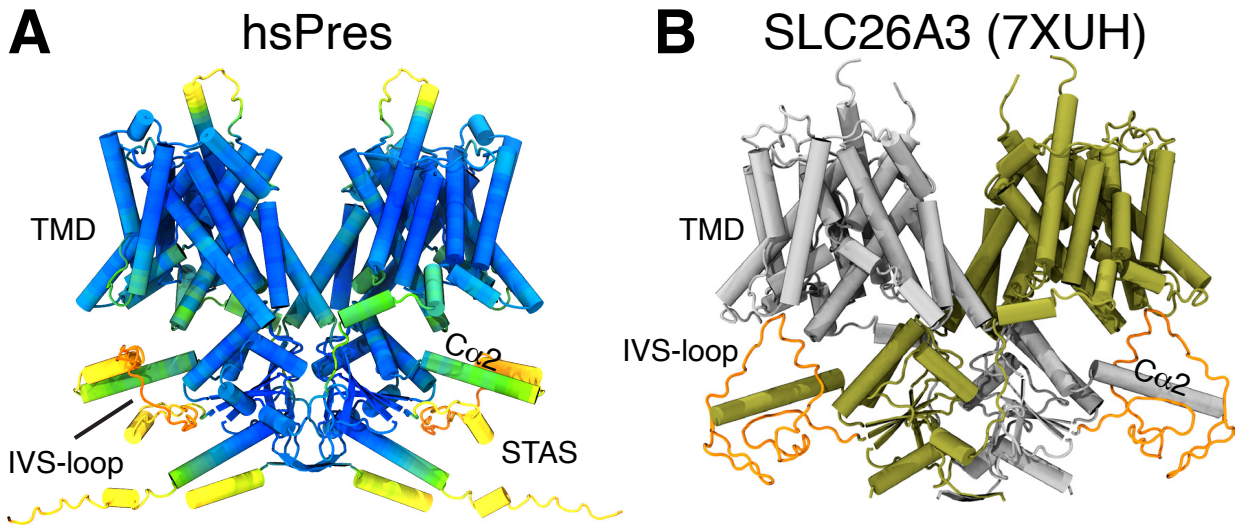

Figure S5. **Comparison between hsPres and SLC26A3.** (A) hsPres model presented and colored as in Figure 4. (B) Rendering of SLC26A3 (PDB-ID: 7XUH) shown in cartoon representation with one chain colored gray and the other ochre. In both chains, the IVS-loop is colored orange.

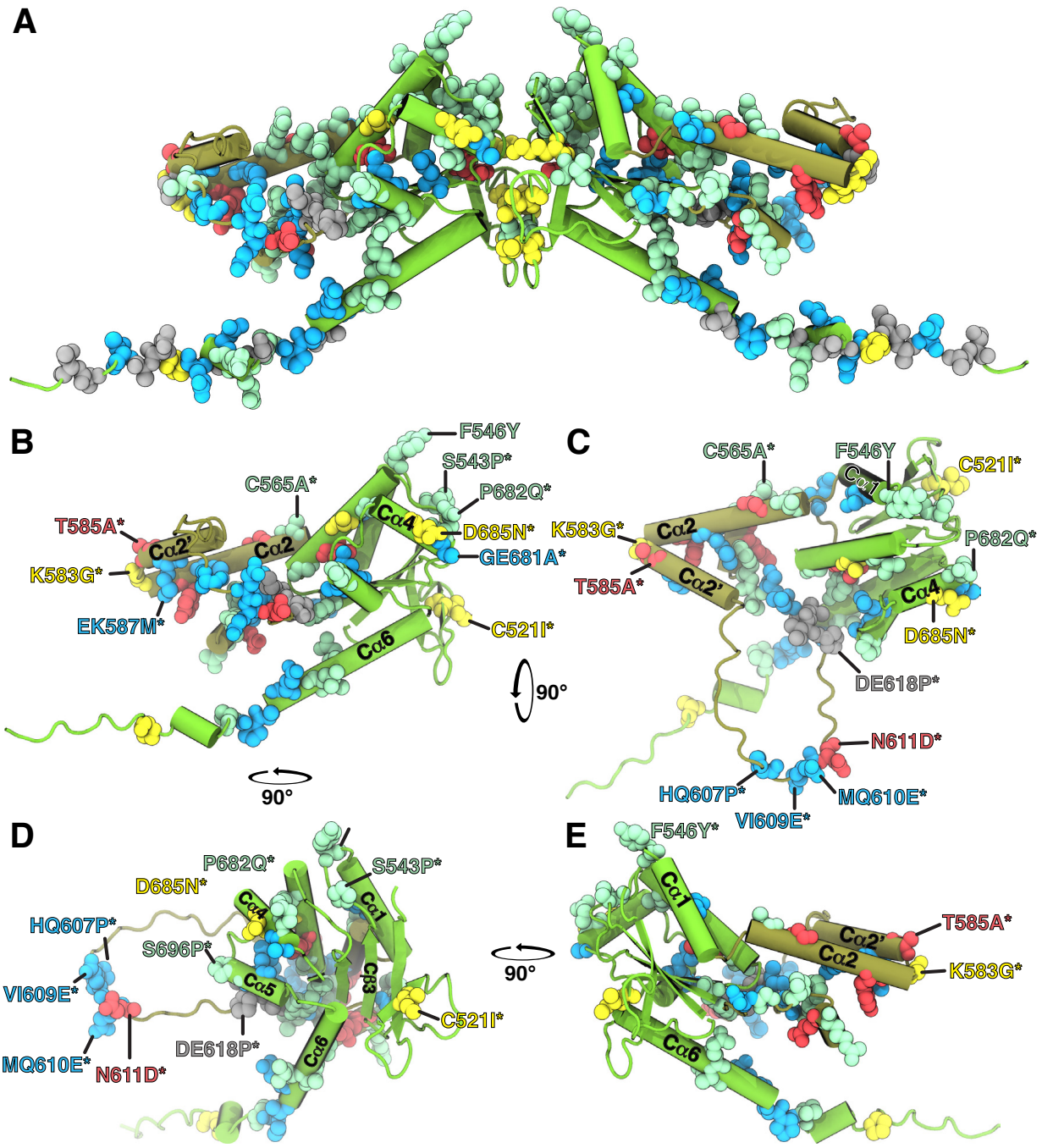

Figure 16. **Substitutions STAS domain** (A)All substitutions found on the STAS domain mapped into the hsPres model. (B-E) Detail of the STAS domain showing only the non-conservative substitutions. (B) Same orientation as in A showing only one monomer. (C) 90° rotation of B along the x-axis to show a top wiew (from the extracelular side). (D) 90° rotation of B along the z-axis. (E) 90° rotation of D along the z-axis. In all panels the protein is shown in cartoon representation and substituted residues in VdW representation. The same color scheme as in Figure 1 (BD) was used.

SUPPLEMENTARY TABLES

Table S1. Sequence identity between the inferred ancestral prestin sequences sequences and hsPres.

	ancPis	ancRhi	ancTre	ancAmph	ancAmn	ancRep	ancArch	ancAvi	ancMam	ancThe	ancMar	ancPla	HsPres
ancPis	-	99.60%	92.70%	98.80%	99.10%	98.50%	98.40%	98.00%	66.90%	63.90%	61.20%	60.60%	59.80%
ancRhi		-	93.00%	99.20%	99.50%	98.10%	98.00%	97.60%	67.10%	63.90%	61.20%	60.60%	59.80%
ancTre			-	93.10%	93.50%	92.40%	92.30%	91.60%	66.70%	63.90%	61.30%	61.00%	60.00%
ancAmph				-	99.60%	98.30%	98.10%	97.20%	67.00%	63.80%	61.10%	60.50%	59.80%
ancAmn					-	98.70%	98.50%	97.60%	67.20%	64.00%	61.30%	60.70%	60.00%
ancRep						-	99.60%	98.70%	66.90%	64.00%	61.40%	60.70%	60.10%
ancArch							-	99.10%	66.90%	64.00%	61.40%	60.70%	60.10%
ancAvi								-	66.90%	63.80%	61.20%	60.60%	59.90%
ancMam									-	90.40%	82.10%	82.60%	80.00%
ancThe										-	89.00%	89.10%	86.10%
ancMar											-	81.20%	79.00%
ancPla												-	96.00%
hsPres													-

Table S2. **Substitutions found in N-terminal domain.** **Pink higlight** Earliest substitutions in the evolutionary pathway. **Gray highlight:** Three-step substitutions in the evolutionary pathway. **Blue highlight:** Two-step substitutions in the evolutionary pathway. **Light green highlight:** Early substitutions in the evolutionary pathway (from the amniote ancestor to the mammalian ancestor). **Yellow highlight:** Intermediate substitutions in the evolutionary pathway (from the mammal ancestor to the therian ancestor). **Light red highlight:** Late substitutions in the evolutionary pathway (from the therian ancestor to the placental ancestor). Non-conservative substitutions are marked with an asterisk (*). The estimated time of existence for the ancestral proteins were extracted from TimeTree[75].

	HsA6 (Outgroup)	ancPis (415 MYA)	ancRhi (408.5 MYA)	ancTre (351.7 MYA)	ancAmn (318.9 MYA)	ancMam (180.1 MYA)	ancThe (160 MYA)	ancPla (99.2 MYA)	HsA5 (Extant)
N-Terminal	G2	E2	E2	E2	E2	D2	D2	D2	D2
	D5	Q5	Q5	S5	Q5	P5	Q5	E5	E5
	S7	H7	H7	H7	H7	N7	N7	N7	N7
	G8	E8	E8	E8	E8	-	-	-	-
	M22	C10	C10	C10	C10	V9	I9	I9	I9*
	L24	V12	V12	A12	V12	A11	A11	A11	A11
	R25	Q13	Q13	Q13	Q13	A12	A12	A12	A12
	D28	M16	M16	K16	M16	R15	R15	R15	R15*
	H30	C18	C18	C18	C18	C17	C17	Y17	Y17
	L35	V23	V23	V23	V23	I22	I22	I22	I22
	L36	Y24	Y24	Y24	Y24	F23	F23	F23	F23
	N37	N25	N25	N25	N25	S24	S24	S24	S24
Nα1	Q38	Q26	Q26	Q26	Q26	H25	H25	H25	H25
	E39	E27	E27	E27	E27	Q26	P26	P26	P26*
	H40	L28	L28	S28	L28	V27	V27	V27	V27
	E42	Q30	Q30	Q30	Q30	H29	H29	Q29	Q29
	E43	G31	G31	G31	G31	G30	G30	E30	E30*
	L44	Q32	Q32	Q32	Q32	R31	R31	R31	R31
Loop Nα1-Nα2	-	E37	E37	E37	E37	D36	D36	D36	D36
	-	T39	T39	K39	T39	V38	V38	V38	V38
	A50	Q41	Q41	Q41	Q41	E40	D40	D40	D40
	P51	P42	P42	P42	P42	P41	P41	S41	S41
Nα2	R52	L43	L43	L43	L43	I42	I42	I42	I42
	T53	S44	S44	G44	S44	G43	G43	G43	A43
	H54	Q45	Q45	Q45	Q45	D44	D44	D44	D44
	W56	I47	I47	I47	I47	I46	I46	L46	L46
	R57	A48	A48	A48	A48	K47	K47	K47	K47*
	T58	H49	H49	H49	H49	Q48	Q48	Q48	Q48
	W59	S50	S50	S50	S50	A49	A49	A49	A49*
	L60	C51	C51	C51	C51	L50	L50	F50	F50*
Loop Nα2-Nα3	Q61	R52	R52	R52	R52	S51	S51	T51	T51*
Nα3	S63	S54	S54	S54	S54	T53	T53	T53	T53
	R64	S55	S55	S55	S55	P54	P54	P54	P54*
	A67	A58	A58	A58	A58	V57	I57	I57	I57
	Y68	K59	K59	K59	K59	K58	R58	R58	R58
	A69	S60	S60	S60	S60	N59	N59	N59	N59
	L70	H61	H61	H61	H61	I60	I60	I60	I60*
	L71	L62	L62	L62	L62	I61	I61	I61	I61
	Q73	S64	S64	S64	S64	R63	M63	M63	M63
Nα4	L78	L69	L69	L69	L69	C68	C68	T68	T68*
Loop Nα4-TMD 1	R83	R74	R74	R74	R74	A73	A73	A73	A73*
	P85	P76	P76	P76	P76	K75	K75	K75	K75*

Table S3. **Substitutions found in the transmembrane domain (TMD).** Same legend as S2

	HsA6 (Outgroup)	ancPis (415 MYA)	ancRhi (408.5 MYA)	ancTre (351.7 MYA)	ancAmn (318.9 MYA)	ancMam (180.1 MYA)	ancThe (160 MYA)	ancPla (99.2 MYA)	HsA5 (Extant)
TM1	V86	V77	V77	V77	V77	P76	P76	F76	F76*
	L90	L81	L81	L81	L81	I80	V80	V80	V80
	L94	I85	I85	I85	I85	I84	F84	L84	L84
	L95	I86	I86	I86	I86	V85	V85	V85	V85
	L98	I89	I89	L89	I89	I88	I88	I88	I88
	M93	M94	M94	M94	M94	L93	L93	L93	L93
	Y111	Y102	Y102	Y102	Y102	Y101	F101	F101	F101*
	L113	L104	L104	L104	L104	L103	M103	M103	M103*
TM2	F131	F122	F122	F122	F122	I121	I121	I121	I121
	F134	T125	T125	T125	T125	T124	C124	C124	C124*
TM3	T146	T137	T137	T137	T137	P136	P136	P136	P136*
	L161	E152	E152	I152	E152	L151	L151	L151	L151*
	A162	A153	A153	A153	A153	V152	V152	V152	V152
EC-Loop	-	E156	E156	E156	E156	D155	D155	D155	D155
	-	M157	M157	M157	M157	M156	M156	I156	I156
	-	F158	F158	F158	F158	F157	F157	V157	V157*
	Q164	I160	I160	I160	I160	P159	P159	P159	P159*
	A165	M161	M161	M161	M161	G160	G160	G160	G160*
	L166	E162	E162	E162	E162	G161	G161	G161	G161*
	N167	-	-	A163	-	G162	V162	V162	V162
	S169	S165	S165	S165	S165	S164	S164	A164	A164*
	N172	S168	S168	S168	S168	G167	G167	G167	G167*
TM4	T174	D170	D170	D170	E169	E169	E169	E169	E169
	A179	M181	M181	M181	L174	L174	M174	L174	L174
	S185	V187	V187	V187	M180	M180	M180	M180	M180
	T186	A188	A188	A188	S181	S181	S181	S181	S181*
	I201	L203	L203	L203	L196	L196	L196	C196	C196*
TM5	Q230	Q232	Q232	Q232	Q232	M225	M225	M225	M225*
	F235	L237	L237	L237	L237	F230	F230	F230	F230*
TM5b	L246	L248	L248	L248	L248	L241	F241	F241	F241*
	I249	L251	L251	I251	L251	V244	V244	V244	V244
	V252	L254	L254	L254	L254	T247	T247	T247	T247*
	L253	I255	I255	I255	I255	V248	V248	V248	V248
	W257	S259	S259	S259	S259	T252	A252	Q252	Q252*
	L259	I261	I261	I261	I261	I254	V254	V254	V254
	P260	T262	T262	T262	T262	K255	K255	K255	K255*
TM6	S262	T264	T264	T264	T264	L257	L257	L257	L257
	G265	A267	A267	A267	A267	A260	C260	C260	C260*
	T266	A268	A268	A268	A268	S261	S261	S261	S261*
	V268	V270	V270	I270	V270	V263	V263	G263	G263*
	V272	I274	I274	I274	I274	M267	M267	M267	M267
	A273	C275	C275	C275	C275	C268	C268	V268	V268*
	G274	I276	I276	I276	I276	F269	F269	F269	F269*
	V275	V277	V277	V277	V277	G270	G270	G270	G270*
	V279	S281	S281	G281	S281	G274	G274	G274	G274*
	L283	I285	I285	I285	I285	F278	F278	F278	F278*
	D285	D287	D287	D287	D287	E285	E285	E285	E280
TM7	G297	M299	M299	M299	M299	L292	L292	L292	L292
	L299	I301	I301	I301	I301	F294	F294	F294	F294
	L300	I302	I302	I302	I302	F295	F295	F295	F295
	T301	V303	V303	V303	V303	A296	A296	A296	A296
	G304	I306	I306	I306	I306	I299	I299	M299	M299
	I308	V310	V310	V310	V310	V303	I303	I303	I303
	M312	M314	M314	M314	M314	L307	F307	F307	F307*
	K315	S317	S317	S317	S317	K310	K310	K310	K310*
	R317	T319	T319	S319	T319	S312	S312	S312	S312
TM8	I358	M360	M360	M360	M360	V353	V353	V353	V353
	I365	I367	I367	I367	I367	I360	I360	T360	T360*
	F366	F368	F368	F368	F368	F361	F361	L361	L361
	L368	L370	L370	L370	L370	I363	I363	N363	N363*
TM9	I389	V391	V391	V391	V391	I384	I384	I384	I384
	I392	F394	F394	F394	F394	L387	L387	L387	L387
TM10	M402	M404	M404	M404	M404	I397	L397	L397	L397
	S410	S412	S412	S412	S412	G405	G405	G405	G405*
TM11	A420	T422	T422	L422	T422	C415	C415	C415	C415*
	S422	S424	S424	S424	S424	A417	A417	A417	A417*
	L424	I426	I426	I426	I426	L419	L419	L419	L419
	I426	V428	V428	V428	V428	I421	I421	I421	I421
	V431	V433	V433	L433	V433	L426	L426	L426	L426
	L433	I435	I435	I435	I435	L428	L428	T428	T428
	E435	Y437	Y437	Y437	Y437	F430	F430	F430	F430*
TM12	A443	T445	T445	T445	T445	A438	A438	A438	A438*
	A446	A448	A448	A448	A448	A441	S441	S441	S441*
	I445	M452	M452	M452	M452	I445	I445	I445	I445
	R458	K460	K460	K460	K460	M453	M453	M453	M453*
	S461	G463	G463	A463	G463	S456	S456	S456	S456*
TM13	L475	A477	A477	A477	A477	T470	T470	T470	T470*
	V479	V481	V481	V481	V481	T474	T474	T474	T474*
	T480	A482	A482	A482	A482	T475	T475	T475	T475
	A483	A485	A485	A485	A485	S478	S478	S478	S478*
TM14	V494	L496	L496	L496	L496	I489	I489	I489	I489
	I498	A500	A500	A500	A500	I493	I493	I493	I493*
	F499	F501	F501	F501	F501	I494	I494	I494	I494*
	L501	M503	M503	I503	M503	L496	L496	L496	L496

Table S4. **Substitutions found in the STAS domain.** Same legend as S2

	HsA6 (Outgroup)	ancPis (415 MYA)	ancRhi (408.5 MYA)	ancTre (351.7 MYA)	ancAmn (318.9 MYA)	ancMam (180.1 MYA)	ancThe (160 MYA)	ancPla (99.2 MYA)	HsA5 (Extant)
Cβ1	S514	R516	R516	R516	R516	R509	K509	K509	K509
	V515	I517	I517	I517	I517	V510	V510	V510	V510
	V519	I521	I521	I521	I521	I514	L514	L514	L514
Cβ2	I524	I526	I526	I526	I526	V519	V519	V519	V519
	R526	C528	C528	C528	C528	C521	I521	I521	I521*
Cβ3	S545	A547	A547	A547	A547	I540	I540	I540	I540
Loop Cβ3-Cα1	T548	S550	S550	S550	S550	P543	P543	P543	P543*
	V549	L551	L551	L551	L551	I544	I544	I544	I544
	F551	F553	F553	F553	F553	Y546	Y546	Y546	Y546*
Cα1	E555	E557	E557	E557	E557	D550	D550	D550	D550
	S558	V560	V560	V560	V560	N553	S553	S553	S553*
Cα2	D570	C572	C572	C572	C572	A565	A565	A565	A565*
	F571	A573	A573	A573	A573	V566	L566	L566	V566
	I573	L575	L575	L575	L575	M568	M568	M568	M568
	S574	A576	A576	A576	A576	A569	A569	G569	G569*
	K577	K579	K579	K579	K579	R572	R572	R572	R572
	L580	Q582	Q582	Q582	Q582	L575	L575	M575	M575*
	Q583	H585	H585	H585	H585	H578	H578	Y578	Y578*
IVs Loop	L591	M589	M589	M589	M589	M582	L582	V582	V582
	Q592	K590	K590	K590	K590	K583	G583	G583	G583*
	K593	Q591	Q591	Q591	Q591	R584	T584	N584	N584
	E594	A592	A592	A592	A592	T585	T585	A585	A585*
	K596	E594	E594	E594	E594	K587	K587	M587	M587*
	L617	V613	V613	V611	V611	T607	T606	T605	T605
	D619	H615	H615	H613	H613	Q609	Q608	P607	P607*
	R621	V617	V617	L615	V615	I611	E610	E609	E609*
	S622	M618	M618	M616	M616	Q612	Q611	E610	E610*
	N623	N619	N619	N617	N617	N613	N612	D611	D611*
	Q647	D635	D635	D632	D632	E623	D622	P618	P618*
	E648	A636	A636	A633	A633	L624	S623	I619	I619*
	D649	A637	A637	A634	A634	I625	L624	V620	V620*
	S650	Q638	Q638	Q635	Q635	V626	V625	I621	I621*
	K651	-	-	-	-	Q627	Q626	K622	K622
	A652	D639	D639	D636	D636	N628	N627	S623	S623*
	P653	S640	S640	S637	S637	T629	T628	T624	T624*
	D654	S641	S641	S638	S638	F630	F629	F625	F625*
	K659	E646	E646	E643	E643	E635	E634	Q630	Q630*
	A660	H647	H647	H644	H644	R636	R635	R631	R631*
	L663	K650	K650	K647	K647	P639	P638	P634	P634*
	Q665	M652	M652	A649	M649	G641	G640	G636	G636*
	P666	T653	T653	T650	T650	A642	A641	D637	D637*
Cβ4	S670	S657	S657	S654	S654	T646	T645	T641	T641
	L671	I658	I658	I655	I655	I647	I646	V642	V645
Cα3	L686	A673	A673	A670	A670	V662	V661	V657	V657*
	A661	K677	K677	K674	K674	K666	A665*	A661	A661*
	N691	S678	S678	S675	S675	G667	G666	G662	G662
	F693	I680	I680	I677	I677	I669	I668	V664	V664*
	R697	E684	E684	K681	E681	G673	G672	G668	G668*
	E698	E685	E685	E682	E682	D674	D673	D669	D669
	I699	I686	I686	I683	I683	V675	V674	V670	V670
Cβ5	E702	S689	S689	S686	S686	Y678	Y677	Y673	Y673*
	M705	I692	I692	I689	I689	V681	L680	L676	L676
Cα4	S710	G697	G697	G694	G694	E686	A685	A681	A681*
	P711	P698	P698	P695	P695	Q687	Q686	Q682	Q682*
	V713	M700	M700	M697	M697	V689	V688	V684	V684
	S714	D701	D701	E698	D698	D690	N689	N685	N685*
Cα5	S725	S712	S712	S709	S709	P701	P700	P696	P696*
	I726	T713	T713	T710	T710	S702	S701	A697	A697*
	T727	T714	T714	T711	T711	T703	V702	L698	L698*
	K729	D716	D716	D713	D713	E705	E704	E700	E700
Cα6	H745	K732	K732	K729	K729	R721	R720	R716	R716
	P746	A733	A733	A730	A730	E722	E721	E717	E717*
	R747	S734	S734	S731	S731	A723	A722	A718	A718*
	P748	S735	S735	S732	S732	L724	L723	L719	L719*
	P750	S737	S737	S734	S734	Q726	Q725	E721	E721
	S752	T739	T739	T736	T736	A728	E727	E723	E723*

Table S5. **Substitutions found in the C-terminal tail.** Same legend as S2

	HsA6 (Outgroup)	ancPis (415 MYA)	ancRhi (408.5 MYA)	ancTre (351.7 MYA)	ancAmn (318.9 MYA)	ancMam (180.1 MYA)	ancThe (160 MYA)	ancPla (99.2 MYA)	HsA5 (Extant)
C-Terminal	V754	L741	L741	S738	L738	L730	L729	S725	S725
	S755	D742	D742	E739	D739	A731	A730	A726	A726*
	V756	A743	A743	G740	A740	P732	P731	P727	P727*
	T757	Q744	Q744	Q741	Q741	P733	P732	P728	P728*
	R758	V745	V745	V742	V742	P734	P733	P729	S729
	L759	S746	S746	S743	S743	Q735	Q734	Q730	Q730
	-	T747	T747	T744	T744	E736	E735	E731	E731
	-	S748	S748	S745	S745	A737	D736	D732	D732
	-	E749	E749	E746	E746	T738	S737	S733	L733
	-	-	S750	S747	S747	D739	E738	E734	E734
	-	-	Q751	Q748	Q748	Q740	P739	P735	P735*
	-	-	K752	Q749	K749	L741	I740	N736	N736
	-	-	-	-	-	D742	V741	A737	A737
	-	-	-	-	-	S743	S742	T738	T738
	-	-	-	-	-	E745	E744	E740	A740
	-	-	-	-	-	P746	A745	A741	T741

Table S6. List NCBI accession codes for the sequences used for the phylogenetic tree inference.

ncbi accession code	species	ncbi accession code	species
XP_036256283.1	Molothrus ater	XP_019065031.1	Fukomys damarensis
XP_005423579.1	Geospiza fortis	XP_004639300.1	Octodon degus
XP_005482704.1	Zonotrichia albicollis	XP_005377288.1	Chinchilla lanigera
XP_030093883.1	Serinus canaria	XP_005319521.1	Ictidomys tridecemlineatus
XP_021401801.1	Lonchura striata domestica	XP_003469853.1	Cavia porcellus
XP_038009369.1	Motacilla alba alba	XP_008590477.1	Galeopterus variegatus
XP_041261364.1	Onychostruthus taczanowskii	XP_045420227.1	Lemur catta
XP_041874339.1	Corvus kubaryi	XP_008061008.1	Carlito syrichta
XP_039579687.1	Passer montanus	XP_008828979.1	Nannospalax galili
XP_039918735.1	Hirundo rustica	XP_044991818.1	Jaculus jaculus
NXH49971.1	Dicaeum eximium	XP_027629197.1	Tupaia chinensis
XP_032912856.1	Catharus ustulatus	XP_004676967.1	Condylura cristata
XP_005039223.1	Ficedula albicollis	XP_048195806.1	Perognathus longimembris pacificus
XP_014724522.1	Sturnus vulgaris	XP_028746752.1	Peromyscus leucopus
NXA75387.1	Thryothorus ludovicianus	XP_021017527.1	Mus caroli
XP_023774723.1	Cyanistes caeruleus	XP_004479043.1	Dasypus novemcinctus
XP_027743972.1	Empidonax traillii	XP_037693397.1	Choloepus didactylus
XP_027535188.1	Neopelma chrysocephalum	XP_006882669.1	Elephantulus edwardii
XP_047924367.1	Anser cygnoides	XP_004390776.1	Trichechus manatus latirostris
XP_027315248.1	Anas platyrhynchos	XP_007942749.1	Orycteropus afer afer
NXL86920.1	Alectura lathami	NP_001129435.1	Sus scrofa
NXC50793.1	Penelope pileata	XP_037023606.1	Artibeus jamaicensis
NXJ03225.1	Odontophorus gujanensis	AIJ04812.1	Taphozous melanopogon
XP_021240219.1	Numida meleagris	XP_036748345.1	Manis pentadactyla
XP_042728216.1	Lagopus leucura	XP_060229914.1	Meriones unguiculatus
XP_029871101.1	Aquila chrysaetos chrysaetos	ADE75006.1	Murina leucogaster
XP_026696926.1	Athene cunicularia	XP_004320315.1	Tursiops truncatus
XP_009813276.1	Gavia stellata	EHB08960.1	Heterocephalus glaber
XP_009570914.1	Fulmarus glacialis	XP_042540604.1	Dipodomys spectabilis
XP_042648655.1	Tyto alba	XP_004602180.1	Sorex araneus

Table S7. Continuation of Table S1

ncbi accession code	species	ncbi accession code	species
XP_009326612.1	Pygoscelis adeliae	XP_037367109.1	Talpa occidentalis
XP_009461166.1	Nipponia nippon	XP_004702662.1	Echinops telfairi
XP_009693986.1	Cariama cristata	NP_001135733.1	Oryctolagus cuniculus
XP_009562019.1	Cuculus canorus	NP_945350.1	Homo sapiens
XP_010086019.1	Pterocles gutturalis	XP_045243469.1	Macaca fascicularis
XP_009986007.1	Tauraco erythrolophus	XP_004591411.1	Ochotona princeps
XP_009481643.1	Pelecanus crispus	XP_040827316.1	Ochotona curzoniae
XP_009945683.1	Leptosomus discolor	KAB1276031.1	Camelus dromedarius
XP_037246269.1	Falco rusticolus	ACI02071.1	Rhinolophus ferrumequinum
XP_009881210.1	Charadrius vociferus	XP_032944790.1	Rhinolophus ferrumequinum
XP_009639677.1	Egretta garzetta	AJF40182.1	Rhinolophus ferrumequinum
XP_010297045.1	Balearica regulorum gibbericeps	ACI02074.1	Hipposideros armiger
XP_009869992.1	Apaloderma vittatum	ACI02079.1	Rhinolophus luctus
KFZ67571.1	Podiceps cristatus	ACI02077.1	Cynopterus sphinx
XP_010199933.1	Colius striatus	XP_006925450.1	Pteropus alecto
XP_010002455.1	Chaetura pelagica	ADE75004.1	Eonycteris spelaea
XP_008947217.1	Merops nubicus	ACI02076.1	Rousettus leschenaultii
XP_010129337.1	Buceros rhinoceros silvestris	AEQ27771.1	Leptonycteris yerbabuenae
XP_009941858.1	Opisthocomus hoazin	ADO14483.1	Kogia breviceps
XP_010159229.1	Eurypyga helias	CAF5187026.1	Ranitomeya imitator
XP_014819266.1	Calidris pugnax	ADE75013.1	Physeter catodon
XP_010183410.1	Mesitornis unicolor	ADO14481.1	Mesoplodon densirostris
XP_009076534.1	Acanthisitta chloris	ADE75008.1	Balaenoptera physalus
XP_009900559.1	Dryobates pubescens	XP_036912978.1	Sturnira hondurensis
KQK74720.1	Amazona aestiva	XP_024413218.1	Desmodus rotundus
KGL74759.1	Tinamus guttatus	XP_036102144.1	Molossus molossus
XP_005148456.2	Melopsittacus undulatus	XP_016075078.1	Miniopterus natalensis
XP_030336665.1	Strigops habroptila	ADE75007.1	Rhinopoma hardwickii
XP_025953428.1	Dromaius novaehollandiae	XP_036285551.1	Pipistrellus kuhlii
XP_013799723.1	Apteryx mantelli mantelli	XP_008148033.1	Eptesicus fuscus
XP_030310668.1	Calypte anna	ADE75005.1	Megaderma lyra

Table S8. Continuation of Table S1

ncbi accession code	species	ncbi accession code	species
XP_021152991.1	Columba livia	AEQ27774.1	Pteronotus parnellii
XP_046763026.1	Gallus gallus	XP_005862958.1	Myotis brandtii
XP_003201941.2	Meleagris gallopavo	AEQ27772.1	Mormoops megalophylla
XP_015739933.1	Coturnix japonica	ADE75011.1	Megaptera novaeangliae
XP_025902176.1	Nothoprocta perdicaria	XP_024593872.1	Neophocaena asiaeorientalis asiaeorientalis
XP_009667752.1	Struthio camelus australis	XP_022410127.1	Delphinapterus leucas
XP_019364998.1	Gavialis gangeticus	XP_033295445.1	Orcinus orca
XP_019399134.1	Crocodylus porosus	XP_007177376.1	Balaenoptera acutorostrata scammoni
XP_006261320.1	Alligator mississippiensis	ADO14480.1	Hyperoodon ampullatus
XP_025070500.1	Alligator sinensis	XP_007467117.1	Lipotes vexillifer
XP_023964938.1	Chrysemys picta bellii	XP_007888429.1	Callorhinchus milii
XP_039373494.1	Mauremys reevesii	XP_047193271.1	Scophthalmus maximus
XP_024052722.2	Terrapene carolina triunguis	GCC28608.1	Chiloscyllium punctatum
XP_044836183.1	Mauremys mutica	XP_039657837.1	Perca fluviatilis
KAG6929882.1	Chelydra serpentina	KAG5845660.1	Anguilla anguilla
TFK06033.1	Platysternon megacephalum	XP_042191261.1	Callorhinchus milii
XP_037745122.1	Chelonia mydas	XP_033832208.1	Periophthalmus magnuspinnatus
XP_030410163.1	Gopherus evgoodei	XP_032895760.1	Amblyraja radiata
XP_048692359.1	Caretta caretta	XP_030209716.1	Gadus morhua
XP_038252898.1	Dermochelys coriacea	XP_005995355.1	Latimeria chalumnae
XP_006137601.1	Pelodiscus sinensis	XP_023697136.1	Paramormyrops kingsleyae
XP_032630412.1	Chelonoidis abingdonii	XP_047660243.1	Tachysurus fulvidraco
XP_028603177.1	Podarcis muralis	XP_030640173.1	Chanos chanos
XP_033017733.1	Lacerta agilis	XP_018585544.2	Scleropages formosus
XP_034983284.1	Zootoca vivipara	XP_028306182.1	Gouania willdenowi
XP_028603181.1	Podarcis muralis	BAE75795.1	Takifugu obscurus
XP_020645638.1	Pogona vitticeps	XP_042578750.1	Cyprinus carpio
XP_044279152.1	Varanus komodoensis	XP_041110887.1	Polyodon spathula
XP_008109741.1	Anolis carolinensis	XP_013884866.1	Austrofundulus limnaeus
XP_042323185.1	Sceloporus undulatus	XP_023651296.1	Paramormyrops kingsleyae
XP_015270544.1	Gekko japonicus	XP_026882432.2	Electrophorus electricus

Table S9. Continuation of Table S1

ncbi accession code	species	ncbi accession code	species
XP_048357987.1	Sphaerodactylus townsendi	XP_038667500.1	Scyliorhinus canicula
XP_032077595.1	Thamnophis elegans	XP_030220884.1	Gadus morhua
XP_026557454.1	Pseudonaja textilis	XP_034019476.1	Thalassophryne amazonica
XP_026532999.1	Notechis scutatus	XP_040028743.1	Gasterosteus aculeatus aculeatus
XP_034289543.1	Pantherophis guttatus	KAF6725554.1	Oryzias melastigma
XP_039219698.1	Crotalus tigris	XP_030628249.1	Chanos chanos
XP_015666174.1	Protobothrops mucrosquamatus	XP_015208044.1	Lepisosteus oculatus
XP_025021645.1	Python bivittatus	XP_020456913.1	Monopterus albus
XP_029471991.1	Rhinatrema bivittatum	XP_037099016.1	Syngnathus acus
XP_033814420.1	Geotrypetes seraphini	XP_010885587.2	Esox lucius
XP_030071983.1	Microcaecilia unicolor	XP_007577439.1	Poecilia formosa
XP_040199424.1	Rana temporaria	XP_029974339.1	Salarias fasciatus
XP_040268475.1	Bufo bufo	XP_028810143.1	Denticeps clupeoides
XP_044136099.1	Bufo gargarizans	XP_030575683.1	Archocentrus centrarchus
KAG8577153.1	Engystomops pustulosus	XP_018602967.2	Scleropages formosus
KAG8439951.1	Hymenochirus boettgeri	KAG9280146.1	Astyanax mexicanus
XP_002933151.3	Xenopus tropicalis	XP_029901453.1	Myripristis murdjan
XP_018111261.1	Xenopus laevis	KAG7455842.1	Megalops atlanticus
XP_018108161.1	Xenopus laevis	XP_032363930.1	Etheostoma spectabile
XP_007504130.1	Monodelphis domestica	KAF4079897.1	Ameiurus melas
XP_038597097.1	Tachyglossus aculeatus	XP_028279308.1	Parambassis ranga
XP_044535659.1	Gracilinanus agilis	XP_020785137.1	Boleophthalmus pectinirostris
XP_027694688.1	Vombatus ursinus	XP_043877147.1	Solea senegalensis
XP_036615877.1	Trichosurus vulpecula	XP_029017915.1	Betta splendens
XP_043823103.1	Dromiciops gliroides	XP_029912187.1	Myripristis murdjan
XP_025305222.1	Canis lupus dingo	XP_034529019.1	Notolabrus celidotus
XP_035923360.1	Halichoerus grypus	XP_046878718.1	Hypomesus transpacificus
XP_045876718.1	Meles meles	XP_035771674.1	Neolamprologus brichardi
XP_039095034.1	Hyaena hyaena	XP_034382164.1	Cyclopterus lumpus
XP_046525804.1	Equus quagga	XP_047245584.1	Girardinichthys multiradiatus
XP_004439369.1	Ceratotherium simum simum	XP_026183039.1	Mastacembelus armatus
XP_035003799.1	Hippoglossus stenolepis	XP_047431607.1	Mugil cephalus