Article

Discoveries

Accelerated evolution analysis uncovers *PKNOX2* as a key transcription factor in the mammalian cochlea

Anabella P. Trigila[1], Valeria C. Castagna[2], Lara Berasain[1], Dante Montini[1], Marcelo Rubinstein1[1,3], Maria Eugenia Gomez-Casati[2] & Lucía F. Franchini[1,*]

[1] Instituto de Investigaciones en Ingeniería Genética y Biología Molecular (INGEBI), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, C1428, Argentina.

[2] Instituto de Farmacología, Facultad de Medicina, Universidad de Buenos Aires, Buenos Aires, Argentina.

[3] Departamento de Fisiología, Biología Molecular y Celular, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, 1428 Buenos Aires, Argentina

*To whom correspondence should be addressed: franchini@dna.uba.ar; franchini.lucia@gmail.com

Abstract

The genetic bases underlying the evolution of morphological and functional innovations of the mammalian inner ear are poorly understood. Gene regulatory regions are thought to play an important role in the evolution of form and function. To uncover crucial hearing genes whose regulatory machinery evolved specifically in mammalian lineages, we mapped accelerated noncoding elements (ANCEs) in inner ear transcription factor (TF) genes and found that *PKNOX2* harbors the largest number of ANCEs within its transcriptional unit. Using reporter gene expression assays

in transgenic zebrafish we determined that four *PKNOX2*-ANCEs drive differential expression patterns when compared with ortholog sequences from close outgroup species. Because the functional role of *PKNOX2* in cochlear hair cells has not been previously investigated, we decided to study *Pknox2* null mice generated by CRISPR/Cas9 technology. We found that *Pknox2*$^{-/-}$ mice exhibit reduced distortion product otoacoustic emissions (DPOAEs) and auditory brainstem response (ABR) thresholds at high frequencies together with an increase in peak 1 amplitude, consistent with a higher number of IHCs-auditory nerve synapsis observed at the cochlear basal region. A comparative cochlear transcriptomic analysis of *Pknox2*$^{-/-}$ and *Pknox2*$^{+/+}$ mice revealed that key auditory genes are under *Pknox2* control. Hence, we report that *PKNOX2* plays a critical role in cochlear sensitivity at higher frequencies and that its transcriptional regulation underwent lineage-specific evolution in mammals. Our results provide novel insights about the contribution of *PKNOX2* to normal auditory function and to the evolution of high-frequency hearing in mammals.

Introduction

Mammals are characterized by a highly developed auditory system that includes the transformation of jaw joints into middle ear ossicles and the evolution of the cochlea (Manley 2000; Manley 2012; Fritzsch et al. 2013). The mammalian cochlea has two types of mechanosensory hair cells that have different and critical functions in hearing. Inner hair cells (IHCs) display an elaborate presynaptic apparatus, receive predominantly type I afferent innervation, signal to cochlear neurons communicating sound information to the brain and are considered as the true phonoreceptors. Outer hair cells (OHCs) are biological motors innervated predominantly by efferent fibers that amplify the sound through a mechanism known as somatic electromotility (Brownell 1990; J. Zheng et al. 2000; Liberman et al. 2002; Dallos et al. 2008). OHCs and their associated type II spiral ganglion neurons are a mammalian innovation as they are absent in the amniote and avian basilar papilla (Manley 2010; Zhang and Coate 2017). Although the organ of Corti, composed of IHCs and OHCs emerged before the split of monotremes and therian mammals (including marsupials and placentals), only the latter have fully developed high-frequency sensitivity. In fact, monotremes including platypus and echidnas show high-frequency limits around 15 kHz (Gates et al. 1974; Mills and Shepherd 2001) that are similar to those found in other amniotes lineages such as birds and lizards (Manley 2012). High-frequency hearing depends on the function of OHCs and its sound amplification mechanism named somatic electromotility mediated by the molecular motor prestin. This mechanism of sound amplification developed to an extreme in several therian lineages such as echolocators, which are capable of perceiving ultrasonic signals (Madsen et al. 2004; Li et al. 2008; Churchill et al. 2016).

Although a recent study has found that coding sequences from genes involved in hearing underwent positive evolution in particular mammalian lineages (Wang et al. 2020), the genetic bases underlying the evolution of high-frequency hearing in mammals remain largely unknown. In this work, we aimed to identify candidate genes differentially expressed in OHCs that could underlie the emergence of high-frequency hearing in mammals using signatures of accelerated evolution in noncoding regions, since it has been proposed that molecular evolution of non-coding regulatory regions dictate lineage-specific functional novelties (King and Wilson 1975; Prud'homme et al. 2007; Carroll 2008). In particular, we have focused on genes encoding transcription factors (TF) because lineage-specific evolution of TFs has contributed to the origin of morphological and functional innovations (Lynch and Wagner 2008; Kaessmann 2010; Nowick and Stubbs 2010; Wagner and Lynch 2010; Cheatle Jarvela and Hinman 2015). In this study, we sought to identify mammalian genes carrying lineage-specific accelerated noncoding elements (ANCEs) within their transcriptional units. ANCEs are conserved noncoding regions that accumulate nucleotide changes at a faster rate than neutral in a lineage-specific manner and are, therefore, a useful tool to identify putative regulatory regions underlying lineage-specific evolutionary changes. To this end, we used publicly available databases of bat accelerated regions (BARs) (Eckalbar et al. 2016), human accelerated regions (HARs) (Capra et al. 2013) and therian-specific accelerated regions (TSARs) (Holloway et al. 2016),) and found that *PKNOX2/Pknox2* (*PBX/Knotted 1 Homeobox 2*) accumulated the greatest number of ANCEs particularly within its

introns. To investigate the possibility that *PKNOX2*-ANCEs have contributed to mammalian inner ear evolution, here, we analyzed whether they drive reporter gene expression in transgenic zebrafish assays to novel territories in comparison to ortholog sequences from close outgroup species.

PKNOX2 encodes a transcription factor that is highly expressed in the inner ear, as reported by several transcriptomic studies (Liu et al. 2014; Scheffer et al. 2015; Li et al. 2016; Liu et al. 2018; Yamashita et al. 2018; Ranum et al. 2019), although its functional role in the auditory system remains unknown. To gain insight into *Pknox2* function in the mouse inner ear, we generated *Pknox2* deficient mice by CRISPR/Cas9 technology and found that *PKNOX2* plays a critical role in the regulation of cochlear sensitivity at higher frequencies in mammals.

Results

Seeking for mammalian-specific accelerated evolution signatures in inner ear genes

To gain insight into the molecular evolution underlying the unique features of mammalian hearing we sought to identify non coding accelerated elements present in the transcriptional units of TF genes expressed in IHCs and OHCs. To this end, we performed an intersection between a recently generated database of 1,643 TF mouse genes expressed in IHCs and/or OHCs (Li et al. 2016) and three publicly available databases of mammalian accelerated elements obtained by comparing distinct mammalian lineages carrying 2,148 Bat Accelerated Regions (BARs; Eckalbar et al., 2016); 4,797 Therian-Specific Accelerated Regions (TSARs; Holloway et al., 2016); and 2,745 Human Accelerated Regions (HARs; Capra et al., 2013). By crossing these databases we found 340 TF hair cells-expressed genes harboring mammalian ANCEs in their transcriptional units (Supp. Table 1), 18 of which are OHC-defining cluster genes as indicated by Ranum et al., 2019 (Supp. Table 2). Four of these 18 OHC TF genes (*Six4, Stat3, Rbfox2* and *Pknox2*) stood out from the rest by displaying the largest number of ANCEs accumulated in their transcriptional units (Supp. Table 2). The functional roles of *Six4* and *Stat3* in the development of the inner ear have been already established (Ozaki et al. 2001; Ozaki et al. 2004; Chen et al. 2017) whereas the RNA binding protein RBFOX2 has been primarily involved in the regulation of cell-specific alternative splicing (Gehman et al. 2012; Zhou et al., 2021-NAR). In turn, the developmental and physiological roles of *PKNOX2* in the inner ear remains unexplored to date. Interestingly, human *PKNOX2* accumulates 7 ANCEs in its transcriptional unit (chr11:125,164,751-125,433,389 (hg38); 3 BARs, 1 HAR and 3 TSARs; Fig. 1A).

Analysis of PKNOX2-ANCEs as putative transcriptional enhancers in transgenic zebrafish assays

To investigate whether the uncovered mammalian PKNOX2-ANCEs act as transcriptional enhancers in the auditory system we evaluated their ability to drive reporter gene expression in transgenic zebrafish, a validated strategy previously used to characterize several mammalian enhancers, even in the absence of conserved fish orthologs (Fisher, Grice, Vinton, Bessling, and McCallion

2006; Bessa et al. 2009; Domené et al. 2013; Kamm, López-Leal, et al. 2013; Liu et al. 2017; Caporale et al. 2019; Trigila et al. 2021). We first examined whether endogenous *pknox2* is expressed in the zebrafish auditory system. To this end we performed *in situ* hybridization studies along several stages of zebrafish development (Fig. 2 and Supp. Fig. 1). We found that *pknox2* is expressed in the developing otic capsules at 24 and 48 hpf (Supp. Fig.1). At 72 hpf *pknox2* expression was apparent in the otic capsule and in neuromasts of the lateral line, where it remained highly expressed up to 7 dpf, the last stage analyzed (Fig. 2 and Supp. Fig. 1). The otic capsule gives rise to the fish inner ear whereas the lateral line is a sensory system that allows fishes to detect weak water motions and pressure gradients (Whitfield 2002). The lateral line is composed of several hair cells-containing sensory units called neuromasts accommodated alongside the body and head and interconnected among them. Because the morphology and function of lateral line hair cells is very similar to those of the inner ear, many studies use the fish lateral line to better understand hair cells' physiology (Whitfield 2002). Moreover, numerous genes required for hair-cell function in the zebrafish have been recently associated with auditory defects in mice and humans, revealing their molecular and functional conservation (Nicolson 2017), and prompting the zebrafish as a valuable genetic model for the study of hearing and balance (Sheets et al. 2021). In addition to the auditory system, *pknox2* expression has been found in the developing brain and eyes at 24, 48 and 72 hpf (Supp. Fig. 1), and in the branchial arches starting at 48 hpf and up to 7 dpf.

We then sought to evaluate whether the mammalian *PKNOX2*-ANCEs are capable of driving eGFP expression in transgenic zebrafish (Fig. 1, B-I and K-S). In addition to the seven ANCEs found within *PKNOX2* introns, we also decided to study a BAR located in the proximal 5' flanking region of *PKNOX2*, making a total of eight *PKNOX2*-ANCEs (4 BARs, 1 HAR and 3 TSARs) that were tested in comparison with ortholog sequences taken from a near outgroup species (Fig. 1A). Each of these 16 sequences was subcloned upstream of a mouse *c-Fos* minimal promoter fused to the green fluorescent protein (eGFP) reporter gene and, together, flanked by Tol2 elements to maximize genomic integration, as we previously reported (Kamm, Pisciottano, et al. 2013; Caporale et al. 2019; Trigila et al. 2021). The 16 transgenes were individually microinjected into one-cell stage zebrafish embryos along with the Tol2 transposase mRNA, and eGFP activity was monitored 24, 48 and 72 hours post-fertilization (hpf) (Fig. 1, B-I and K-S).

BAR1156 from *Myotis lucifugus* (little brown bat) and its *Mus musculus* (mouse) ortholog drove similar expression patterns at 24 hpf (Fig. 1, B,C; Supp. Fig 2) in the eye, forebrain, hindbrain, somites, spinal cord and otic capsules. At 48 hpf, transgenic expression of both constructs remained in the nervous system, while expression strength at the somites diminished and expression in the heart became apparent. At this stage, bat and mouse ortholog sequences continued to drive expression to the developing otic capsule in all transgenic lines tested (3 lines for each transgene) and in the neuromasts of the lateral line (2 out of 3 and 1 out of 3

transgenic lines generated with mouse and bat BAR1156, respectively). At 72 hpf, eGFP expression in the neuromasts disappeared in most lines whereas it remained active in the eye, forebrain, midbrain, and the heart with only slight expression in somitic muscle (Supp. Fig. 2).

The mouse BAR1160 sequence drove strong eGFP expression to the otic capsule and various regions of the developing zebrafish nervous system including the eye, forebrain, midbrain, hindbrain and spinal cord at 24, 48 and 72 hpf (Fig. 1, H, I; Supp. Fig. 3), and also in somitic muscle in all 5 transgenic lines analyzed. In contrast, its bat BAR1160 ortholog sequence failed to drive reporter gene expression in all 3 transgenic lines analyzed, suggesting a lineage-specific loss of function.

The other two *PKNOX2*-BARs (BAR1157 and BAR1158), either from bat or mouse sequences, failed to drive reporter gene expression in all transgenic zebrafish lines generated at all developmental stages analyzed (Fig. 1, D-G, Q, R; Supp. Figs. 4 and 5), suggesting that these sequences do not act as transcriptional enhancers, at least in the zebrafish model.

The mouse 2xHAR.32 sequence elicited strong eGFP expression in the developing nervous system, eye, inner ear, pharyngeal arches and pronephric structures (Fig. 1, M; Supp. Fig. 6) in all transgenic lines generated. However, its human 2xHAR.32 ortholog failed to drive reporter gene expression at any of the stages analyzed in the 3 transgenic lines generated (Fig. 1, M, N; Supp. Fig. 6) suggesting lineage-specific loss of enhancer function.

Analyses of the three TSARs were performed comparing the expression patterns elicited by the *M. musculus* sequence, as a therian representative, and their *Gallus gallus* (chicken) orthologs, as a non-mammalian outgroup. Mouse TSAR.3236 did not drive reporter gene expression at any of the developmental stages analyzed in the 3 transgenic lines generated, while its ancestral chicken ortholog directed eGFP expression to the developing nervous system, eye, pharyngeal arches and otic vesicle at 24 hpf (Fig. 1, K, L; Supp. Fig. 7) and continuing at 48 and 72 hpf, with additional expression in the fin and heart at 72 hpf (Fig. 1, L; Supp. Fig. 7) in the 3 transgenic lines generated, suggesting mammalian-specific loss of enhancer function.

Mouse and chicken TSAR.0878 showed high reporter gene expression at 24 hpf in the developing nervous system and the otic capsule (Fig.1, O, P; Supp. Fig. 8). At 48 hpf, eGFP expression was also observed in the developing nervous system, the eyes and the heart of all transgenic lines generated with each transgene. At this latter stage, we observed strong expression in the developing otic capsule in the six transgenic lines carrying the chicken TSAR.0878 sequence that contrasted with a much less intense expression in the 3 transgenic lines carrying the mouse sequence (Supp. Fig. 8). At 72 hpf, eGFP expression was observed in the eyes, nervous system, heart and craniofacial structures in all chicken and mouse TSAR.0878 transgenic lines. Noticeably, in four out of six of the chicken TSAR.0878 transgenic zebrafish lines we observed eGFP expression in the neuromasts of the lateral line

(Fig. 1, P; Supp. Fig. 8) in quite contrast with the lack of eGFP expression in neuromasts from any of the four mouse TSAR.0878 transgenic zebrafish lines, suggesting that this expression territory was lost in therian mammals (Supp. Fig. 8). We also found strong eGFP expression in the auditory system of all transgenic lines carrying either mouse or chicken TSAR.0878 (Supp. Fig. 8). Further analysis of a chicken TSAR.0878 transgenic line performed at 7 dpf showed strong eGFP expression in the hair cells of the neuromasts of the head and the trunk (Fig. 2, B to M). This expression pattern coincides with that shown by the endogenous *pknox2* at this stage as evidenced in our *in situ* hybridization study (Fig. 2, N-P). Lastly, mouse and chicken TSAR.2216 drove no eGFP expression in all transgenic zebrafish lines generated (Fig. 1, Q, R).

Taking together, we have identified five *PKNOX2*-ANCEs that act as active enhancers in transgenic zebrafish and likely regulate *PKNOX2* expression during embryonic development in vertebrates. We have also found that four *PKNOX2*-ANCEs (TSAR.0878, BAR1160, TSAR.3236 and 2xHAR.32) display changes in expression patterns (including gain or loss of function) possibly as a consequence of the evolutionary process they underwent in the different mammalian lineages.

DNA methylation analysis of *PKNOX2*-ANCEs

To gain more insight into the regulation of *PKNOX2* expression in the inner ear, we searched for epigenetic signals indicative of regulatory activity in the developing mouse inner ear by using inner ear methylome data obtained at three developmental mouse stages (E16.5, postnatal day (P) 0 and P22) with Whole Genome Bisulfite Sequencing (WGBS) (Yizhar-Barnea et al. 2018) (Supp. Table 3). Tissue-specific differential methylation states of genomic regions at single-base resolution allow prediction of regulatory function (Stadler et al. 2011). In fact, active gene promoters have been associated with unmethylated regions (UMRs), defined as regions with average methylation rates lower than 10%. In turn, intergenic or intronic low-methylated regions (LMRs) displaying average methylation rates between 10% and 50% are commonly observed in transcriptional enhancers (Stadler et al. 2011). We found that BAR1156 was included in a LMRs at E16.5, BAR1157 displayed UMRs at all three stages (E16.5, P0, P22), BAR1158 had UMRs and LMRs, TSAR3236 presented LMRs at E16.5 and P0, 2xHAR.32 was contained in LMRs at the three developmental stages (E16.5, P0, P22), TSAR.2216 displayed LMRs at P0 (Supp. Table 3). This data complements our enhancer assay findings about the regulatory function of *PKNOX2*-ACNE sequences.

Transcription factor binding sites embedded in putative enhancers of *Pknox2*

A group of TFs expressed in the inner ear including ATOH1, SOX2, GFI1, SIX1 and POU4F3 are known to bind to enhancers of genes that play important roles in the development of the sensory epithelium of the organ of Corti (Xiang et al. 1997; Bermingham et al. 1999; Wallis et al. 2003; Zheng et al. 2003; Kiernan et al. 2005; Matern et al. 2020). We searched for canonical binding sites of these TFs, as defined by JASPAR2022 (mm10), within *PKNOX2*-ANCEs and other non-accelerated conserved noncoding

sequences present in the mouse *Pknox2* locus. We found that all *PKNOX2*-ANCEs contain canonical binding sites for SOX2, all *PKNOX2*-ANCEs except TSAR.3236 contain binding sites for ATOH1, and all *PKNOX2*-ANCEs but 2xHAR.32 contain GFI1 binding sites (Supp. Table 3). In addition, BAR1156, TSAR.0878 and TSAR.3236 contain binding sites for SIX1 (Supp.Table 3) and no *PKNOX2*-ANCEs contain binding sites for POU4F3 (Supp. Table. 3). We also found multiple binding sites for these TFs in predicted CRE regulatory sites (ENCODE) and conserved noncoding sequences (Supp. Table 3) that could participate in the regulation of *Pknox2* expression in the inner ear.

We then performed a comparative analysis of TF binding sites present in *PKNOX2*-ANCEs and their close outgroup orthologs to detect possible gains or losses in the accelerated elements. We found that TSAR.0878 from mice, which showed no expression in the zebrafish neuromasts, lacks one SIX1 and three GFI1 binding sites compared to the chicken (galGal5) sequence (Supp. Table 3). In addition, TSAR.3236 from mice, while showing no eGFP expression in transgenic assays, gained one SOX2, three GFI1 and two SIX1 binding sites compared to the galGal5 chicken sequence. The human 2xHAR.32, which is inactive as an enhancer in transgenic zebrafish at all developmental stages analyzed, lost three ATOH1 and one SOX2 sites compared to its mouse ortholog (Supp. Table 3). The bat BAR1160 lost one GFI1 site and gained two ATHO1 and one SOX2 sites. Although gains and losses of TF binding sites in *PKNOX2*-ANCEs may underlie spatio-temporal and/or quantitative changes in gene expression, their functional consequences will need to be experimentally tested in further studies.

PKNOX2 emergence in vertebrates and functional diversification

PKNOX2 and its paralog *PKNOX1* (PBX/Knotted 1 Homeobox 1; Fig. 3, A and B) are members of the TALE family of atypical homeodomain-containing TFs that also include PBX and MEIS. TALE TFs can heterodimerize with typical HOX TF adding DNA binding specificity and affinity to canonical binding sites within regulatory sequences of target genes (Merabet and Mann 2016).

To further understand *PKNOX2* evolutionary history in vertebrates and particularly in mammals, we performed a comparative analysis of gene paralogs. Using available vertebrate and chordate *PKNOX* sequences we built a phylogenetic tree depicting a duplication event that generated *PKNOX1* and *2* occurred at the origin of vertebrates, since just one ancestral *PKNOX* gene is found in chordates (Fig. 3, C). These data suggest that *PKNOX1* and *2* emerged at the time of the whole genome duplication event (WGD) at the origin of vertebrates (Fig. 3, C).

Based on their different spatial patterns of expression (Imoto et al. 2001) and distinct ability to form heterodimers with other TFs of the TALE family (Fognani et al. 2002), it has been suggested that *Pknox1* and *Pknox2* have functionally diversified along evolution. To test this hypothesis, we compared coding and noncoding evolutionary rates of the two paralogs and other members of the TALE family (Suppl. Material 1). We found no evidence for positive selection in PKNOX1 and PKNOX2 coding

sequences, following a general trend in the TALE family to evolve under strong purifying selection (Supp. Material 1). In fact, human PKNOX2 (ENSP00000298282) and PKNOX1 (ENSP00000291547) are 62% identical (Supp. Table 4), indicating that changes in protein sequence may not have been a major driver of functional divergence between the two paralogs. We then evaluated the distribution of conserved sequences (phastCons) in the genomic *loci* of all members of the TALE family (Fig. 3, D and E) and found that *PKNOX2* harbors one of the greatest number of phastCons in its noncoding sequence that actually doubles the number of phastCons found in noncoding *PKNOX1* (Fig. 3, E). In addition, other members of the TALE family such us *MEIS1*, *MEIS2*, *PBX1* and *PBX3* accumulate multiple phastCons in their noncoding regions, suggesting that their expression patterns could be also determined by multiple *cis*-regulatory elements (Fig. 3, D and E), as it has been found for many other developmental genes. Altogether, these results suggest that numerous noncoding elements in *PKNOX2* could have contributed to shape a functional diversification process relative to *PKNOX1*.

Despite the fact that *PKNOX1* and *2* share a common evolutionary origin and that both proteins display a relatively high sequence identity, the number of CNEs in the transcriptional unit of *PKNOX2* is much higher than in *PKNOX1*. Thus, we hypothesize that after duplication, the most likely outcome in the functional divergence of both paralos is a neofunctionalization process of *PKNOX2* that might have occurred by gaining novel regulatory non-coding elements along several time points of the vertebrate tree (Fig. 3, F). Alternatively, the ancestral gene already had multiple regulatory regions but *PKNOX1* could have lost some of them after duplication, restricting its expression territories. Further molecular evolution of PKNOX2 enhancers in mammals could have contributed to fine tuning and other functional features in the inner ear of these phylogenetic groups.

Pknox2 is highly expressed in hair cells of the mammalian inner ear

In order to analyze the expression patterns of *Pknox1* and 2 throughout mouse development, we explored anRNA-seq CAGE (Cap Analysis of Gene Expression) database performed by the RIKEN FANTOM5 project after studying several mouse cell types. We detected that *Pknox1* has a basal expression in all cell types and higher expression restricted to T-cells (Supp. Fig. 9, A), as previously described (Chen et al. 1997; Ferretti et al. 1999; Ferretti et al. 2006), and further supported by deviation in haematologic parameters involving B- and T-cell number found in *Pknox1*$^{-/-}$ mice (Dickinson et al. 2016). Moreover, *Pknox1* has been shown to play an essential role in hematopoiesis (Di Rosa et al. 2007) and *Pknox1* hypomorphic mutants die of anemia and angiogenic anomalies (Ferretti et al. 2006). *Pknox2* expression, in turn, is mainly restricted to cardiac muscle cells, various neuronal cell types and inner ear cells (Fig. 4, A-D; Supp. Fig. 9, A). According to the Gene Expression Database (GXD; http://www.informatics.jax.org/expression.shtml) *PKNOX2* is expressed in several tissues including skeletal muscle, cardiovascular, nervous, digestive and reproductive systems during development. It is also reported to be expressed in branchial

arches, craniofacial structures and the auditory system (Supp. Fig.9, A, C). By reanalyzing data from previous transcriptomics studies in the auditory system we found that *Pknox2* is mostly expressed in hair cells with higher expression levels in OHCs than in IHCs, contrasting with the low expression level of *Pknox1* (Fig. 4, A, B; Supp. Fig. 9, B; (Liu et al. 2014; Liu et al. 2018)). *Pknox2* transcripts are mainly detected in cochlear (hair-cell enriched) *Atoh1*-GFP+ cells starting at postnatal day 4 (P4), while *Pknox1* levels remained low (Fig. 4, A-C; Supp. Fig 9, B; (Scheffer et al. 2015)). These later stages correspond to the maturation of mechanosensitivity in inner ear hair cells. Another comparative study between the expression of TFs in mouse OHCs and IHCs reported *Pknox2* as one of the top OHCs differentially expressed transcripts (Li et al. 2016). Similarly, a single-cell RNA-Seq study performed with OHCs showed that *Pknox2* was one of the top three featured genes in this cell type, a result that was further confirmed by bulk RNA-Seq and single cell qPCR (Yamashita et al. 2018). Single cell RNA-Seq from manually isolated OHCs, IHCs and Deiters' cells also confirmed that *Pknox2* is mainly expressed in OHCs at P15 (Ranum et al. 2019) and is a cluster defining gene at this stage. In addition, transcriptomic studies show that *Pknox2* is also expressed in neurons of the spiral ganglion at P25-27 (Shrestha et al. 2018). To explore in more detail the role of *Pknox2* in the mouse inner ear, we characterized the protein localization at P8 by immunofluorescence in the organ of Corti (Fig. 4, D) and found that PKNOX2 is mainly restricted to the nuclei of both OHCs and IHCs (Fig. 4, D). In contrast, we did not detect PKNOX1 in the nuclei of OHCs or IHCs at P8 (Supp. Fig 9, B).

Pknox2 ablation affects hearing in mutant mice

Given the high expression level of *Pknox2* in OHCs and IHCs, we sought to investigate its functional role in the mouse auditory system. To this end we generated a novel null allele mutant mouse strain by using CRISPR/Cas9 technology. A single guide RNA targeted to *Pknox2* coding exon 1 (Fig. 4, E) led to a 20-bp deletion predicting a truncated protein (Fig. 4, F). Homozygous *Pknox2*$^{-/-}$ mice showed undetectable levels of PKNOX2 in the brain, in quite contrast to their *wild-type* siblings that showed robust expression in this tissue (Fig. 4, F, G; Supp. Fig. 9, D). We then compared auditory function in adult *Pknox2*$^{+/+}$ and *Pknox2*$^{-/-}$ littermates. To assess the integrity of OHC function *in vivo* we performed a distortion product otoacoustic emissions (DPOAEs) assay by using a microphone in the external auditory canal (Shera and Guinan 1999) and found a slight decrease in DPOAEs thresholds in P60 *Pknox2*$^{-/-}$ mice compared to *Pknox2*$^{+/+}$ controls that was statistically significant only at 45.25 kHz (Mann-Whitney test: df = 1, p = 0.002898) (Fig. 5, B). Hearing sensitivity was evaluated by recording auditory brainstem responses (ABR) which are sound evoked potentials generated by neuronal circuits in the ascending auditory pathway. We observed significant reductions in ABR thresholds in *Pknox2*$^{-/-}$ mice at 22.65 kHz (Mann-Whitney test: df = 1, p = 0.02087), 32 kHz (Mann-Whitney test: df = 1, p = 0.032) and 45.25 kHz (Mann-Whitney test: df = 1, p = 0.01931) (Fig. 5, A). Interestingly, evoked

potential amplitudes in the ABR wave I were higher in *Pknox2*$^{-/-}$ mice than in their *wild-type* siblings at 80 dB SPL (Mann-Whitney, p = 0.044, df=1 at 22.65 kHz, p = 0.001224, df=1 at 32 kHz, and p= 0.002111, df=1. at 45.25 kHz) (Fig. 5, C). This first peak refers to the first synapse between the IHC and type I afferent terminals of the auditory pathway. To establish if the hearing phenotype in these mice is based on sensory-neural hearing gain from synaptic communication or altered function of the spiral ganglion neurons, we performed a histological analysis. Whole-mount organs of Corti were immunostained with antibodies against CtBP2-Ribeye, a critical protein present at the presynaptic ribbon (Khimich et al. 2005), and GluA2 AMPA-type glutamate receptors, which are expressed at the postsynaptic afferent terminal (Matsubara et al. 1996; Liberman et al. 2002; Maison et al. 2013) (Fig. 5, D). IHC–afferent synapses were identified by colocalization of CtBP2 and GluA2 puncta at the base of the IHC (Liberman et al. 2011). We counted puncta at three different cochlear locations: apical, medial, and basal. The number of prelocalized, postlocalized, or colocalized synaptic markers per IHC was counted in each imaged cochlear section (3-5 animals/genotype) to calculate the synaptic density per IHC in *Pknox2*$^{-/-}$ and *Pknox2*$^{+/+}$ mice (Fig. 5, D). The number of presynaptic ribbon, postsynaptic afferent terminal and synaptic counts were similar in the apical and medial cochlear regions of both genotypes. However, a significant increase of ~ 28% was detected in the basal (high-frequency) cochlear region of *Pknox2*$^{-/-}$ mice (Mann-Whitney, $p<0.05$, for the three measurements) (Fig. 5, E-G). In addition, we found no differences in the number of CtBP2 puncta (Supp. Fig. 9, E) or the overall morphology of the organ of Corti and the cochlea assessed by prestin immunolabelling in OHCs from *Pknox2*$^{-/-}$ and *Pknox2*$^{+/+}$ mice (Supp. Fig. 9, E). Further immunofluorescence analysis using Myosin VIIa and neurofilament heavy chain (NFH) antibodies (Boero et al. 2020; Hickman et al. 2021) confirmed normal cochlear morphology in mice of both genotypes (Supp. Fig. 10). Taken together, *Pknox2*$^{-/-}$ OHCs are functionally normal *in vivo* while displaying a gain of sensitivity at high frequencies. The lower threshold for sound intensity and greater electrical responses of primary auditory neurons to a sound stimulus (ABR peak 1 amplitude) at high frequencies found in *Pknox2*$^{-/-}$ mice suggest that *Pknox2* participates in the amplification process normally occurring in OHCs in the high-frequency zone, the most relevant to mammalian hearing (Heffner and Heffner 2018). Our data shows that IHC/auditory nerve synapses at high frequencies are also affected, in line with the observation that Pknox2 is found in both OHCs and IHCs during development (Fig. 4, D). Considering the normal hearing capacity of *Pknox1*$^{-/-}$ mice (Dickinson et al. 2016), it is tempting to speculate that *Pknox2* plays a non-redundant functional role during hair-cell cochlear development and hearing capacity in mammals.

Pknox2 **deficiency alters gene expression in the mouse cochlea**

To investigate the genetic bases underlying the peculiar hearing phenotype observed in mice lacking *Pknox2* we performed a comparative RNA-seq study in cochleas obtained from eight-day-old (P8) *Pknox2*$^{-/-}$ mice and their *wild-type* littermates. We found

690 downregulated and 334 upregulated genes in cochleas taken from *Pknox2*$^{-/-}$ mice (fold change > 1.5, p-value < 0.01 with FDR correction) (Supp. Table 5). The 20 genes showing highest increases or decreases in cochleas from *Pknox2*$^{-/-}$ mice are depicted in Fig. 6 (A). To identify molecular pathways and multiple gene functional associations we performed a Gene Ontology (GO) term analysis using cut-off FDR-corrected P < 0.05 values to assess the enrichment of differentially expressed genes. As illustrated in Supp. Table 5, there was a frequent association with cell-cell signalling, sensory perception of mechanical stimulus, cell projection organization, vesicle-mediated transport in synapse, cell development, cellular localization and neurotransmitter transport terms, related to the biological processes these differentially expressed genes are involved (Supp. Table. 5). In line with this finding, the cellular components terms were related to cell junction, extracellular matrix, myelin sheath, membrane protein complex and apical dendrite (Supp. Table 5) while terms related to glutamatergic and GABA-ergic synapses were also involved. Finally, the terms related to molecular function revealed an enrichment in structural constituent of ribosome, structural molecule activity, transporter activity, binding, calcium ion binding, inorganic solute uptake transmembrane transporter activity, laminin-1 binding, protein binding, antiporter activity and enzyme binding (Supp. Table 5).

One of the top downregulated genes in *Pknox2*$^{-/-}$ mice was *Ceacam16*, which encodes for a mammal-specific secreted glycoprotein highly expressed in the inner ear (Kammerer et al. 2012). Ceacam16 is present in the tip of the tallest stereocilia in cochlear OHCs and in the tectorial membrane, where its function seems to be critical for successful hearing over an extended frequency range (Zheng et al. 2011). *Ceacam16* is also strongly expressed in Deiter´s cells and interdental cells of the cochlear limbus, and to a lesser level in phalangeal, border, pillar, and inner hair cells (Kammerer et al. 2012). As the majority of GO terms in downregulated genes were related to synapsis (Fig. 6, B), we explored the expression of gene families relevant to neurotransmission and neurite guidance (Fig. 6, C). Differential downregulation was observed in genes encoding for ion channels such as sodium (*Scn4b*), chloride (*Clic5*, *Clcn5*, *Clcn6*, *Lrrc8b*), calcium (*Cacng5*, *Cacng2*, *Cacnb4*, *Ryr2*) and potassium (*Kcnj16*, *Kcnq5*, *Kcnc2*, *Kcnq4*, *Kcna2*, *Kcnc1*, *Kcnk9*, *Kcnc3*, *Kcnj12*, *Kcna1*, *Kcna3*). There was also downregulation of genes encoding for cholinergic nicotinic (*Chrnb3*, *Chrna6*) and metabotropic receptors (*Chrm5*, *Chrm3*), and glutamate (*Grid2ip*) and GABA ionotropic receptors (*Gabbr2*). Other genes included thrombospondin 4 (*Thbs4*), an extracellular-matrix glycoprotein that controls synaptogenesis and neurite growth that has been involved in the particular synaptic organization of the human brain (Cáceres et al. 2007).

Interestingly, the most upregulated transcript we found in the cochleas of *Pknox2*$^{-/-}$ mice encodes for beta-tectorin (*Tectb*), a glycoprotein necessary to upkeep the structure of the tectorial membrane, an extracellular matrix that covers the neuroepithelium of the cochlea and controls the bundles of stereocilia present in sensory hair cells. Mice lacking *Tectb* showed disrupted tectorial

membrane and exhibited low-frequency hearing loss (Russell et al. 2007). Based on these findings, *Pknox2* could regulate the expression levels of structural genes in the organ of Corti such as *Ceacam16* and *Tectb* which, when greatly modified, may explain the hearing differential phenotypes found in *Pknox2*$^{-/-}$ mice.

To analyze whether *Pknox2* might regulate *Ceacam16* and/or *Tectb* expression we searched for potential PKNOX2 binding sites in the *Ceacam16* and *Tectb* mouse loci. We found 12 PKNOX2 binding sites in noncoding regions of the *Ceacam16 locus* (chr7:19,822,844-19,875,216 Mouse Dec. 2011 (GRCm38/mm10)) (Supp. Table 6; Supp. Fig. 11, A). Three of these sites are predicted by ENCODE as Conserved Regulatory Elements and display enhancer signatures: EM10E0843135, EM10E0843115 and EM10E0843142, located in chr7:19864187-19864521, chr7:19831862-19832077 and chr7:19872242-19872553, respectively. In the *Tectb locus* (chr19:55,127,810-55,197,115) we found 25 PKNOX2 binding sites (Supp. Table 6; Supp. Fig. 11, B), one of which is predicted by ENCODE as a CRE present in a distal enhancer (EM10E0682480 chr19:55,190,076-55,190,419) and another one is located in a moderately conserved sequence next to the beginning of a coding exon (chr19:55,183,920-55,183,931).

Finally, we were interested in identifying TF genes that were up- or down-regulated in the cochlea of *Pknox2*$^{-/-}$ mice and that carry PKNOX2 binding sites in their non-coding sequences (Supp. Table 7). *Hes1, Bcl6* and *Id1* stand out in this group because they have been involved in the development of the inner ear (J.L. Zheng et al. 2000; Morrill and He 2020). In addition, we used manually curated databases of human and mouse transcriptional regulatory networks (TRRUST; (Han et al. 2018)) to identify TFs that are likely to participate in regulatory interactions together with PKNOX2 (Supp. Table 8). Although these analyses shed light on PKNOX2 targets, further molecular and functional studies will be necessary to better understand the role of PKNOX2 in the healthy and diseased mammalian hearing system.

Discussion

In this work we performed an accelerated evolution analysis of non-coding elements to uncover TF genes underlying the emergence of morphological and/or functional features of the mammalian inner ear. We found several TF genes that accumulated accelerated noncoding sequences along different mammalian lineages and decided to focus on *PKNOX2*, because this gene accumulated the greatest number of accelerated noncoding elements (ANCEs) within its transcriptional unit and its functional role in the inner ear remained completely unknown. To investigate whether these accelerated changes might have contributed to anatomical and/or functional novelties, we performed molecular and expression analysis of the eight *PKNOX2*-ACNEs detected. Methylation signatures found in genomic DNA taken from mouse inner ear sensory epithelium (Yizhar-Barnea 2018) showed that five out of eight *PKNOX2*-ANCEs are located within low methylation regions (LMRs; BAR1156, BAR1158, TSAR.3236, 2xHAR.32, TSAR.2216) and two of them in unmethylation regions (UMRs; BAR1157, BAR1158) suggesting that *PKNOX2*-ANCEs are likely to act as transcriptional enhancers (LMRs) or promoters (UMRs).

In addition, our reporter gene expression assay in transgenic zebrafish showed that five out of the eight tested *PKNOX2*-ACNEs act as transcriptional enhancers during development of the fish auditory system (BAR1156, BAR1160, 2xHAR.32, TSAR.0878, TSAR.3236). Our results show that each active cis-regulatory element determines a complex expression pattern including the developing nervous system, pharyngeal arches, somites, inner ear and the lateral line system. Besides, we identified several regulatory elements showing partially redundant expression territories, as it has been reported for many other regulatory regions controlling the expression of developmental genes (Cannavò et al. 2016; Osterwalder et al. 2018). It is interesting to note that predictive epigenetic methylation marks do not completely coincide with the results obtained in the reporter transgenic zebrafish assay for most of the analyzed sequences suggesting that assessment of putative regulatory sequences is more informative when tested in *in vivo* animal models. Most importantly, we found that four out of five *PKNOX2*-ACNEs acting as transcriptional enhancers in the auditory system of transgenic zebrafish drove differential reporter gene expression patterns when the accelerated sequence was compared with ancestral ortholog sequences (BAR1160, 2xHAR.32, TSAR.0878, TSAR.3236) suggesting that lineage-specific molecular evolution of *PKNOX2*-ANCEs could have shaped its expression in the mammalian hearing system. However, other reasons may explain differential reporter gene expression patterns driven by ortholog conserved elements such as: (a) a critical TF from the model species does not bind to the accelerated or ancestral ortholog sequence; (b) the enhancer sequence is active at a developmental stage that was not analyzed in this study; (c) positional effects due to persistent integration of one of the orthologs in silent heterochromatin, although by using the Tol2 system as we did in our study to generate transgenic zebrafish, all transgenes are inserted in hundreds of different genomic regions within in the original founder line (Kawakami et al. 2004; Kawakami 2007) addition, several lines for the same transgene are produced to overcome the possibility of insertion in silent chromatin. ; (d) functional incompatibility between a putative enhancer sequences and the heterologous minimal promoter used in

the transgene. All these caveats call for a careful interpretation of the results obtained with different animal models. Nevertheless, we and others have successfully used this methodology to identify changes in enhancer activity in ortholog sequences (Kamm, Pisciottano, et al. 2013; Oksenberg et al. 2013; Erwin et al. 2014; Caporale et al. 2019) and also to identify mutations in noncoding elements involved in human genetic diseases (Mann and Bhatia 2019). In a broader picture, the accelerated evolution process that *Pknox2* underwent in different mammalian lineages could have led to the acquisition of a differential expression pattern along its history. Our phylogenetic analysis indicated that the duplication event that generated *PKNOX1* and *2* occurred at the origin of vertebrates since just one PKNOX ancestral gene is found in urochordates, coinciding with the main source of gene duplication for vertebrates: i.e. the whole genome duplication event (WGD) at the stem of this lineage. How do some members of a transcription factor family acquire a new function or a different expression pattern? Briefly, this could be due to changes in coding or noncoding regions of their transcriptional units (Hoekstra and Coyne 2007). We established that the genomic sequences encoding for amino acids of most members of the TALE family including *PKNOX1* and *2*, were under high selective constraint throughout the history of vertebrates and particularly in mammals. Our analysis of gene expression data in mice indicates that *Pknox1* and *Pknox2* have different expression domains and that *Pknox2* shows a more restricted expression pattern. Although there is some overlap in the expression patterns of these two paralogs, the inner ear seems to be a tissue where these two transcription factors clearly differ in the pattern and level of expression in vertebrates. In fact, our results and data from other sources indicated that *Pknox2* is strongly expressed in hair cells and spiral ganglion neurons in the mouse inner ear in clear contrast to the lack of expression of *Pknox1*. This differential pattern of expression is also observed in other vertebrates, such as chicken, where *Pknox1* is expressed in the posterior neural tube, in the eye and in the branchial arches, whereas *Pknox2* is observed in the anterior areas of the neural tube, including the brain, eye and otic vesicle (Coy and Borycki 2010). Thus, the putative diversification in the expression profile of *PKNOX1* and *2* could be due to the acquisition of a heterogeneous regulatory landscape produced by the modification of noncoding regulatory sequences. We detected that, despite the fact that *PKNOX1* and *2* share a common origin and their level of protein similarity is quite high, the proportion of CNEs in their intronic sequences is very different. In *Pknox1*, vertebrate conserved sequences mainly concentrate in the coding portions of the gene, comprising approximately half of the total conserved sequences. In contrast, *Pknox2* has more than 90% of its total vertebrate conserved sequences in noncoding regions. Our results could likely indicate that the increased presence of retained noncoding conserved elements helped to shape the *PKNOX2* expression pattern in the nervous system and its associated sensory organs. We hypothesize that, after duplication, *PKNOX2* retained its ancestral expression domains but gained regulatory non-coding regions such as enhancers along several time points of the vertebrate tree acquiring new expression domains. Then, in mammals, accelerated evolutionary changes in these enhancers could underlie fine tuning of *PKNOX2* inner ear expression in different lineages to serve particular functional features displayed in these phylogenetic groups.

Given earlier literature indicating expression of *Pknox2* in the developing mammalian cochlea (Li et al. 2016; Yamashita et al. 2018; Ranum et al. 2019) and the fact that this gene was reported as one of the top differentially expressed genes in OHCs, we characterized its function in the inner ear though the generation of mutant mice lacking *Pknox2*. We observed that, in contrast to the previously indicated by transcriptomic data, the Pknox2 protein is highly expressed at the P8 mouse cochlea, in both IHCs and OHCs. Besides, we found that mice lacking *Pknox2* expression show a significant reduction in auditory thresholds together with an increase in ABR peak 1 amplitude at high frequencies. ABR peak 1 amplitude represents the summed sound-evoked spike activity at the synapse between IHCs and afferent nerve fibers. Notably, the increase in suprathreshold ABR peak 1 amplitude correlates with a higher number of colocalized synaptic puncta (CtBP2/GluA2) at the high-frequency (basal-end) region of the cochlea.

Thus, our electrophysiological observations pinpoint to a role of *Pknox2* in regulating the expression of genes that influence synaptic function and/or auditory nerve subtype identities. The correct specification of neuronal subtypes and the presence of a functionally diverse pool of afferent type I neurons contacting IHCs is thought to be critical for sound encoding in the auditory nerve, contributing particularly to the wide dynamic range of the auditory periphery and hearing in background noise (Shrestha et al. 2018). We can speculate that Pknox2 is an important regulator of afferent type I identity (via signals from hair cells or expression in the SGN) and its lack of expression can lead to an expansion of the low-threshold, high-spontaneous rate (high-SR) auditory nerve fiber subpopulation, potentially explaining why there is an increased ABR peak I amplitude response in *Pknox2*$^{-/-}$ mice with an increase in colocalized synaptic puncta. However, since the observed phenotype seems to be exclusive of the basal regions, we could expect that there are *Pknox2* expression differences along the tonotopic axis in afferent type I neurons or hair cells, but unfortunately the evidence does not support this hypothesis so far. There are a few studies that evaluate transcriptomic differences of inner ear cells along the tonotopic axis, however, in a comparison of IHCs between the basal and apical regions at P30-P40, there was not a significant difference in *Pknox2* expression between this regions (Tang et al. 2019).. While *Pknox2* is natively expressed in SG neurons, there was no difference in *Pknox2* expression levels across all neuronal type I subtypes and between apical, basal and medial regions evaluated in mice at P25-P27 (Shrestha et al. 2018). Therefore, we speculate that *Pknox2* is exerting its transcriptional function in IHCs or in neurons during development and that its role could be related to the correct establishment of cochlear synapses. In the case that *Pknox2* is influencing signals released by the hair cell, there are two described mechanisms by which hair cells can affect the specification of neuronal subtypes: (i) by disruption of mechanotransduction or (ii) by the blockage of glutamatergic signaling (Sun et al. 2018). Integrating our evolutionary data with the mechanotransduction mechanism described earlier, we can hypothesize that *Pknox2* could be finely regulating genes involved in very typical mammalian hearing processes, such as frequency selectivity and sensitivity particularly at higher frequencies.

An additional hypothesis is that *Pknox2* is regulating some components of the basal OHCs which, in turn, generate an exacerbated stimulation to the IHCs. OHCs hair bundles are embedded in the tectorial membrane and do also provide feedback influencing the mechanical interactions. This cross-communication between the OHCs and the tectorial membrane is essential to provide the characteristics of sensitivity and selectivity of the mammalian cochlea. Indeed, some structural components of the hair cells and tectorial membrane seem to be altered in its expression patterns in *Pknox2*$^{-/-}$mice, as exemplified by *Ceacam16* and *Tectb*. Particularly interesting is the top down regulated gene *Ceacam16* which is a mammalian-specific secreted glycoprotein expressed in the tip of the tallest stereocilia in OHCs and in the tectorial membrane, where its function seems to be critical for successful hearing over an extended frequency range (Zheng et al. 2011; Kammerer et al. 2012). For instance, in lizards which lack a tectorial membrane, their auditory nerve afferents properties are poorer (Manley 2017). In addition, the gene *Tectb* shows the highest upregulation in *Pknox2*$^{-/-}$mice. Tectb$^{-/-}$ mice show a tectorial membrane with altered wave properties, unable to properly coordinate OHCs to determine sensitivity and frequency selectivity (Ghaffari et al. 2010). The study of tectorial membrane waves showed that it functions as a complementary system in which increased gain sharpens tuning (Ghaffari et al. 2010). This mechanism could possibly explain how, in a context where *Tectb* is upregulated, a high sensitivity without an increase in frequency selectivity is achieved, reconciling the mammalian hallmarks of sensitivity, frequency selectivity and temporal resolution (Ghaffari et al. 2010). If properties in the TM somehow affect mechanotransduction (by disrupting the normal function of an ion channel complex), the altered properties in the IHC could lead to an enhanced formation of synapses. In this sense, there is evidence that increased IHC synapse density increases ABR amplitudes as it has been recently observed in a mouse model overexpressing Ntf3 (Ji et al. 2022).

Although these hypotheses about the role of *Pknox2* regulating high frequency hearing are interesting, further characterization on the molecular mechanism behind the auditory phenotype is needed to confirm these speculations. It is very likely that this transcription factor is having a pleiotropic role in both hair cells, sensory neurons and cochlear nuclei, since our TF regulatory network analyses indicate that *Pknox2* could be controlling several downstream genes that are key to the development and physiology of the auditory system. In any case, the *Pknox2*$^{-/-}$ mouse is a unique genetic model to study the functional consequences of decreased ABR thresholds and increased ABR P1 amplitude.

In sum, our data provide evidence suggesting that *PKNOX1* and *2* genes underwent functional diversification acquiring novel expression patterns and functions after the duplication event that originated them. The hearing phenotype that we found in the *Pknox2* mutant is in clear contrast to what has been found analyzing the *Pknox1* mutant mice, which shows no abnormal auditory function (Dickinson et al. 2016). The *Pknox2* mutant mouse strain may serve as a new model for further studying the effect of impaired gene function and, to our knowledge, this is the first report of mutant mice that increases hearing sensitivity as a consequence of a gene silencing. This could open the door to new gene therapies involving the regulation of transcription factors,

as has been demonstrated with *Atoh1*, which regenerates hair cells and improves hearing in deaf mice (Izumikawa et al. 2005). Furthermore, we found that *PKNOX2* underwent lineage-specific evolution along different mammalian lineages that probably shaped its expression pattern and made it a key player in the mammalian-specific morphological and functional evolution of the inner ear. Thus, we add a new member to the growing list of genes (Franchini and Belén Elgoyhen 2006; Li et al. 2008; Li et al. 2010; Liu et al. 2010; Elgoyhen and Franchini 2011; Cortese et al. 2017; Pisciottano et al. 2019; Trigila et al. 2021) that underwent lineage-specific evolution and that possibly played a role in the evolution of the inner ear in mammals. Finally, our approach suggests that evolutionary analysis could lead to uncovering previously overlooked genes that are key for the physiology or development of a particular organ or system.

Methods

Identification of Inner Ear Transcription Factors Displaying Noncoding Accelerated Sequences

To identify clusters of genomic-accelerated elements in the mammalian genome, we used three publicly available databases containing a total of 9690 accelerated elements including 2148 bat accelerated regions (BARs; (Eckalbar et al. 2016)), 4797 therian-specific accelerated regions (TSARs; (Holloway et al. 2016)), and 2745 human accelerated elements (HARs; (Capra et al. 2013)). The resulting hg19 intervals from the four selected databases were concatenated into one data set where overlapping elements in two or more original data sets were converted into a single one element. Inner ear transcription factors (n = 1643) were retrieved from the dataset of Li et a. 2016. This set was intersected with cluster-specific gene sets from a single-cell study on OHCs (n = 705), IHCs (n = 285) and Deiter cells (n = 558) from Ramun et al. 2019, which led to three subsets of transcription factors specifically expressed in each cell type. We then intersected and arranged the number of accelerated elements (either TSARs, BARs, HARs or all) for each of these cell-type specific transcription factors. We also searched for transcription factor binding sites for Atoh1, Sox2, Gfi1, Six1, and Pou4f3 in the genomic locus of Pknox2, using the JASPAR 2022 database (score: >300).

In Vivo Enhancer Assays in Zebrafish

The function of noncoding accelerated elements located in the *PKNOX2* genomic region were explored using in vivo enhancer assays in transgenic zebrafish. The original multiple alignment files (MAFs) used in the generation of each corresponding accelerated element database (e.g. human accelerated sequences) contain the conserved block including sequences corresponding to the ortholog outgroup species (e.g. other mammals, such as mouse). A visualization of these multiple alignments can be found in Supp. Figures 3-5, section A. Each lineage-specific PKNOX2 accelerated element was studied in comparison with an ortholog sequence taken from a representative species of the corresponding nearest ancestral group. For example, the ability of HARs

(*Homo sapiens*) and BARs (represented by *M. lucifugus*) to drive reporter gene expression in transgenic zebrafish was compared with ortholog mouse (*M. musculus*) sequences, whereas TSARs (represented by *M. musculus*) were compared with their chicken (*Gallus gallus*) orthologs. The genomic location of a corresponding ortholog sequence for each accelerated sequence was obtained via liftingOver (https://genome.ucsc.edu/cgi-bin/hgLiftOver) the accelerated sequence (e.g. mouse; *M. musculus*, mm10) to the target outgroup genome (e.g. chicken; *G. gallus*, galGal6; bat; *M. lucifugus*). Genomic regions containing each conserved and accelerated ortholog element were amplified by proofreading PCR using the primers described (Supp. Table 9) from human, mouse, chicken or bat genomic samples and cloned individually in the vector pXIG_cFos containing the minimal promoter cFos fused to the reporter gene eGFP (Fisher, Grice, Vinton, Bessling, Urasaki, et al. 2006) that was kindly donated by Andy McCallion.

Transgenic zebrafish were produced as originally described by (Fisher, Grice, Vinton, Bessling, Urasaki, et al. 2006). Briefly, each accelerated elements-cFos construct was co-injected with transposase mRNA in one- to two-cell zebrafish embryos. For the generation of stable transgenic lines, injected larvae were raised to adulthood and screened for stable germ-line insertion. Information about transgenic lines analyzed for each construct are included in Supp. Table 9. Transgenic lines not expressing the reporter gene GFP were confirmed positive for the transgene through PCR. When necessary, 0.1 mM of 1-phenyl-2-thiourea was added to the E3 medium to prevent pigment formation. Microscopy was carried out on tricaine-anaesthetized embryos mounted in 3% methyl cellulose. Whole-mount images were taken on an Olympus BX41 fluorescence microscope with an Olympus DP71 digital camera. All experiments including zebrafish were performed in wild-type AB strain from the Zebrafish International Resource Center from the University of Oregon, according to approved protocols by the institutional Animal Studies Committee. Adult zebrafish were maintained at 28 °C in a 14/10 h light/dark cycle in a completely automatic Aquatic Habitats aquarium.

Regulatory elements in the *Pknox2* genomic locus

We used data from previous studies identifying *cis*-regulatory elements, such as promoters and enhancers, identified in whole genome bisulfite sequencing (WGBS) data through the detection of unmethylated and low-methylated regions (UMR/LMRs) in the mouse auditory sensory epithelium (Yizhar-Barnea et al. 2018). The authors defined UMRs as regions with an average methylation lower than 10%, and LMRs as regions with an average methylation between 10% and 50% (Yizhar-Barnea et al. 2018). In addition, we used a recent whole-genome study identifying H3K27ac marks in the mouse inner ear (Li et al. 2020). We intersected accelerated elements in the *Pknox2* locus with these regions, searching for evidence of regulatory function of these non-coding regions. Hi-C heat maps were generated in the virtual Hi-C browser (http://promoter.bx.psu. edu/hi-c/view.php) using the human fetal cortical plate data from (Won et al. 2016). Additional TADs were obtained from (Schmitt et al. 2016).

Coding Sequence Evolutionary Analysis and Gene Expression Comparison

To analyze the evolution of the TALE family coding region, we obtained the coding sequences of a selected group of available species (see accompanying Supplementary Material) with 1-to-1 orthology in Ensembl.v95 (GRCh38.p12). We aligned the sequences using the OMM_MACSE framework (Scornavacca et al. 2019) and the evolutionary analysis was carried out using HyPhy abSREL (Smith et al. 2015). Phylogenetic trees of the PKNOX transcription factor family were constructed using an alignment of 214 sequences obtained from two hierarchical orthologous groups (HOGs) (HOG:0449946, HOG:0460462) derived from OMA Browser (https://omabrowser.org/) (Altenhoff et al. 2020). A JTT model was used to calculate neighbors joining optimal phylogenies, using standard methods in MEGA7 (Kumar et al. 2018). Heatmaps depicting RNA-Seq cell expression were calculated using data from CAGE-Seq RIKEN FANTOM5 retrieved from the Expression Atlas release 36 (https://www.ebi.ac.uk/gxa/home) (Papatheodorou et al. 2020). Inner ear RNA-Seq gene expression plots of *Pknox1* and *Pknox2* were computed with public data from (Liu et al. 2014; Scheffer et al. 2015; Liu et al. 2018).

Zebrafish In Situ Hybridization Assay and Neuromast Labelling

For whole-mount *in situ* hybridization, zebrafish larvae were fixed overnight in 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS), dehydrated in methanol 100% and stored at -20°C until use. Then zebrafish larvae were rehydrated in graded methanol concentrations (75%, 50% and 25%) in PBT (1X PBS, 0.1% Tween 20) and treated for 1 h (5/7 dpf), 45 min (72 hpf), 25 min (48 hpf) and 15 min (24 hpf) with 10 mg/ml proteinase K in PBT, postfixed for 20 min in 4% PFA in PBS and washed in PBT. Prehybridizations were performed for 4 h at 70°C in a hybridization solution (50% formamide, 5X SSC, 500 mg/ml torula tRNA, 50 mg/ml heparin, 0.1% Tween 20, 5 mM citric acid). Hybridizations were performed overnight at 70°C in a fresh hybridization solution containing 1 mg/ml antisense digoxigenin-labeled riboprobes. Probes for in situ hybridization were amplified by PCR from genomic DNA, using primers for the probe eu244 (ZFIN): Forward 5'-TTGATGAAACCCTGCTGTAG-3' and Reverse 5'-GGATCCATTAACCCTCACTAAAGGGAATTGCATGGTGATGAGTAAGAG-3'. The PCR product was cloned in a pBluescript KS+ plasmid, that was digested with SalI or NotI to synthesize using T3 or T7 RNA polymerases the sense and antisense probes, respectively. Larvae were then washed with decreasing concentrations of formamide in 2X SSC (50% and 25%), 2X SSC and twice with 0.2X SSC at 70ºC. Extra washes were performed at room temperature with decreasing concentrations of 0.2X SSC in PBT (75%, 50% and 25%). To perform anti-DIG antibody incubation, samples were first incubated in a blocking solution (10% normal goat serum) for 4 h at room temperature. Anti-DIG antibody (Roche) was incubated at 1/5,000 overnight at 4°C, washed in PBT and alkaline Tris buffer (100 mM NaCl, 50 mM $MgCl_2$, 0.1% Tween 20, and 100 mM Tris-HCl pH 9.5). Staining was performed with NBT/BCIP (Roche) in alkaline Tris buffer at room temperature and, when the desired staining intensity was reached, the reaction was stopped in 1X PBT pH 5.5, 1mM EDTA and mounted in glycerol 70%. For labeling of neuromast hair cells, 7 dpf larvae were immersed in a 140 µM solution of N-(3-trimethylammonium propyl)-4-(6-(4-

(diethylamino)phenyl) hexatrienyl)pyridinium dibromide (FM4-64; Thermo Fisher) for 2 min at room temperature in the dark. Then, they were anaesthetized and mounted in 3% methyl cellulose. Microscopy was carried out on whole-mount embryos/larvae on an Olympus BX41 fluorescence microscope with an Olympus DP71 digital camera.

Mutant mice generation

Mouse strains carrying deleted (knockout) alleles were generated using a modified CRISPR/Cas9 protocol (Wang et al. 2013). Briefly, sgRNA recognition sequence targeting the PKNOX2 coding region (5'-GTGGCCATCATTGTCAGAGC TGG-3', where TGG is the PAM) was designed using CRISPR Design Tool (http://crispr.mit.edu/) aiming towards the initial translated methionines. The T7 promoter was added to the recognition sequence, and the whole sgRNA was generated by a PCR with a reverse primer (5'-aaaagcaccgactcggtgcc-3') from the pX330 plasmid. The T7-sgRNA product was used as a template for *in vitro* transcription using the MEGAshortscript T7 kit (ThermoFisher Scientific). The Cas9 mRNA was *in vitro* transcribed from pMLM3613 plasmid using the mMESSAGE mMACHINE T7 kit (Thermo Fisher Scientific) and polyadenylated using Poly(A) Tailing Kit (Thermo Fisher Scientific #AM1350). Transgenic knockout mice were generated by injecting a mix of Cas9 mRNA (final concentration of 100 ng/ul) and sgRNA (50 ng/ul) into the cytoplasm of FVB blastomeres in accordance with standard procedures approved by the INGEBI-CONICET Laboratory Animal Welfare and Research Committee. Pseudopregnant female mice of FvB strain were used as foster mothers. The animals were bred to homozygosity. Mice were genotyped using the PCR primers: CAGCAGGGATCTCCCAAATA and TCCAGGTGTTCCAGGTTAGG followed by sequencing. In order to evaluate several behavioral parameters in *Pknox2* mutant mice we performed an open field test. We found that locomotor activity, exploration, spatial memory, and anxiety-like behavior of the $Pknox2^{+/+}$ and $Pknox2^{-/-}$ mice were not significantly different. This suggests that the absence of a Pknox2 functional protein did not have a significant impact on overall behavioral performance in mutant mice (Supp. Fig. 12).

Cochlear function tests

Inner ear physiology was performed in mice of either sex anesthetized with xylazine (20 mg/kg , i.p.) and ketamine (100 mg/kg , i.p.) and placed in soundproof chamber maintained at 30 °C, where auditory brainstem responses (ABRs) and distortion-product otoacoustic emissions (DPOAEs) were recorded. The responses were performed on postnatal day 60 (P60) mice ($Pknox2^{+/+}$ n=7, $Pknox2^{-/-}$ n=9). In order to measure sound pressure near the eardrum, sound stimuli were transmitted by a custom acoustic setup with two dynamic earphones used as sound sources (CDMG15008-03A; CUI) and an electret condenser microphone (FG- 23329-PO7; Knowles) connected to a probe tube. Digital stimulus generation and response processing was done by digital I-O boards (National Instruments) powered by LabVIEW-written custom software. For ABRs, needle electrodes were inserted into the skin at

the dorsal midline near the neural crest and pinna with a ground electrode near the tail. Stimuli were 5 ms tone pips (0.5 ms rise-fall, with a cos2 envelope, at 40 / s) delivered to the eardrum at log-spaced frequencies from 5.6 to 45.25 kHz. The response was amplified to 10,000X with a 0.3-3 kHz passband. The sound level was increased from 20 to 80 dB sound pressure level (SPL) in 5 dB stages. At each stage, 1024 responses were averaged, alternating with stimulus polarity. The DPOAEs in response to two primary tones of frequency f1 and f2 were recorded at 2f1-f2, with f2/f1=1.2, and the f2 level 10 dB lower than the f1 level. At intervals of 4 μs, ear-canal sound intensity was amplified and digitally sampled. The DPOAE threshold was defined as the lowest f2 level in which the signal to noise floor ratio is >1.

Cochlear processing and immunostaining

Cochleae from two months old mice (P60) and eight postnatal days (P8) were extracted, perfused intralabyrinthly with 4% PFA in PBS, post-fixed with 4% PFA overnight and decalcified in 0,12M EDTA for five days. Cochlear tissues were then microdissected and permeabilized by freeze/thawing in 30% sucrose (for CtBP2/GluA2 immunostaining) or directly blocked (for prestin, *Pknox1*, *Pknox2*, neurofilament heavy chain (NFH) and Myosin VIIa immunostaining). The microdissected pieces were blocked in 5% normal goat serum (for CtBP2/GluA2, NFH and Myosin VIIa immunostaining) or 5% normal donkey serum (for prestin, *Pknox1* and *Pknox2* immunostaining) with 1% Triton X-100 in PBS for 1 h, followed by incubation in primary antibodies (diluted in blocking buffer) at 37°C for 16 h (for CtBP2/GluA2 immunostaining) or 4°C for 16 h (for prestin, *Pknox1*, *Pknox2*, NFH and Myosin VIIa immunostaining). The primary antibodies used in this study were:1) goat anti-prestin antibody (Santa Cruz biotechnology inc. sc22692;1:700); 2) anti-C-terminal binding protein2 (mouse anti-CtBP2 IgG1; catalog #612044, BD Biosciences; RRID:AB_399431; 1:200) to label the presynaptic ribbon; 3) anti-glutamate receptor 2 (mouse anti-GluA2 IgG2a;1:2000; MAB397, Millipore; RRID:AB_11212990) to label the postsynaptic receptor plaques; 4) rabbit anti-PKNOX1 (#PA5-66065, ThermoFisher; 1:25); 5) rabbit anti-PKNOX2 (#PA5-65946, ThermoFisher; 1:50) 5) mouse anti-Myosin VIIa (#E3018, Santa Cruz; 1:50) and 6) chicken anti neurofilament heavy chain (#AB5539, Millipore, 1:1000). Tissues were then incubated with the appropriate Alexa Fluor-conjugated fluorescent secondary antibody (1:1000 in blocking buffer; Invitrogen) for 2 h at room temperature. Finally, tissues were mounted on microscope slides in FluorSave mounting media (Millipore) for P60 or VectaShield media with DAPI (Vector Laboratories) for P8. For IHC synaptic counts, NFH and IHC area confocal z-stacks (0.3 μm step size) of the apical, medial, and basal regions from each cochlea were taken using a Leica TCS SPE Microscope equipped with 63 (1.5 digital zoom) oil-immersion lens. Image stacks were imported to Fiji software (RRID:SCR_002285; Schindelin et al., 2012). For synaptic quantification IHCs were identified based on their CtBP2-stained nuclei. Each image usually contained 10–20 IHCs. For each stack, a custom Fiji plugin was developed to automate the quantifications of synaptic ribbons, glutamate receptor patches, and colocalized synaptic puncta. Additionally, maximum projections were generated to draw the different ROIs that correspond to each IHC taking the CtBP2-stained nuclei as a reference. Automatic counting of the number of particles on each

ROI was performed. A similar approach was used to analyze the area of each IHC and NFH , different ROIs were drawn in each cell then the threshold was adjusted and an automatic counting of surfaces was performed in each channel.

Western blot analyses

To analyze *Pknox2* expression in mutated and wild type mice, we extracted protein from brain tissue with a protein extraction buffer [50 mM Tris-HCl pH 7.5, 2 mM EDTA; 1% Triton X100; 150 mM NaCl; 0.05% SDS; Halt™ Protease and Phosphatase Inhibitor Cocktail (100X) (Thermo Scientific 78440)]. Then, samples were homogenized and added with 150 mM NaCL, 0,2 % glycerol, 2% bromophenol blue and β-Mercaptoethanol with heating to 100°C during 5min. We separated the samples by size through SDS-PAGE using 12% SDS-Polyacrylamide gels and to make the proteins accessible to detection by antibodies, we transferred them to a nitrocellulose membrane through an electric field (BIO-RAD). We blocked free binding sites with 5% (w/v) nonfat dry milk, 0.05% v/v Tween 20 in TBS (milk/1xTBS-T) for 1 hour. After blocking, membranes were incubated overnight at 4°C with a polyclonal anti-human PKNOX2 antibody produced in rabbit (PA5-65946, Thermo Fisher) at a dilution of 1:500. After washing 3 times in TBS containing 0.05% v/v Tween 20, blots were incubated with the secondary antibody donkey anti-rabbit HRP conjugate (1:2000, Fisher Scientific) for 3 hours at room temperature. The loading control was a mouse monoclonal anti-human beta Actin antibody (dilution 1:10000, BSA 0.5%) (MA5-15739-HRP, Thermo Fisher). Proteins were visualized using ECL detection (Cell Signaling Technology SignalFire™ ECL Reagent #6883) on the GeneGnomeXRQ (Syngene).

RNA sequencing and analysis

Entire cochleae were extracted from *Pknox2*$^{+/+}$ and *Pknox2*$^{-/-}$ mice at eight days of age (P8) in 5 independent biological samples for each genotype. Total RNA was obtained using the Direct-zol RNA miniprep kit (Zymo Research, Irvine, CA), following the CRC protocol (Vikhe Patil et al. 2015). Each sample consisted of pooled RNA from two cochleae derived from one single mouse. Messenger RNA (mRNA) sequencing was performed at Novogene (https://www.novogene.com/amea-en/) using Illumina NovaSeq platforms (paired-end 150 bp sequencing strategy). Sequenced reads were aligned to the mouse genome (mm10) using HISAT2 (Kim et al. 2019). Gene expression levels were calculated using featureCounts (Liao et al. 2014). Differential expression analysis was done using limma with the voom method (Law et al. 2014). Library sizes were normalized with the TMM methods. Weights were applied to samples. eBayes was used with robust settings (robust=TRUE). Lowly expressed genes were filtered on minimum CPM = 0.5 and at least 3 samples, resulting in 10531 genes filtered out for low expression. The minimum fold change was set at 1.5 and p-value threshold at 0.01. P-values were corrected by the Benjamini and Hochberg 1995 method (Benjamini and Hochberg 1995). RNA-seq data was deposited in the Gene Expression Omnibus database (accession ID GSE171921). Differentially expressed genes were submitted to gProfiler for gene ontology (GO) analysis of the biological pathways and processes that these genes are involved in (Raudvere et al. 2019). Network interactions were calculated with STRING (v.11) with

default settings (Szklarczyk et al. 2019). Top 500 up- and down- regulated genes were submitted to TRRUST V2 (https://www.grnpedia.org/trrust) to find target key regulators.

Ethics Approval

All the experiments involving animals were carried out following the Guide for the Care and Use of Laboratory Animals and were approved by the local institutional animal care and use committee.

Availability of Data and Materials

All data generated or analyzed during this study are included in this article and its Supplementary Material. RNA-seq data was submitted to the Gene Expression Omnibus database (accession ID GSE171921).

Acknowledgements

We thank Marta Treimun for excellent technical assistance with mouse experiments. This work was supported by Agencia Nacional de Promoción Científica y Tecnológica, Argentina (PICT2018-02216, PICT2019-1549). AT, VCC and LB had fellowships from CONICET. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. The authors declare no competing interests

Author Contributions

LFF designed and supervised the project. APT, VCC, LB, DM, MECG and LFF conducted the experiments and analyzed data. APT performed most bioinformatics analyses. APT and LFF wrote the manuscript. LFF, MR and MEGC provided reagents, analytical tools and discussed experiments. All authors edited and approved the final version of this report.

Figure Legends

Figure 1. Comparative functional characterization of *PKNOX2* accelerated elements using transgenic zebrafish.

(A) *PKNOX2* locus in chromosome 11 of the human genome showing the location of accelerated noncoding elements (ANCEs). (B-I and K-S) Fluorescent microphotographs showing the eGFP expression pattern driven by the accelerated or conserved ortholog sequence of each *PKNOX2*-ANCEs. Fluorescent microphotographs of BAR1156 mouse (B) and bat (C); BAR1157 mouse (D) and bat (E); BAR1158 mouse (F) and bat (G); BAR1160 mouse (H) and bat (I); TSAR.3236 mouse (K) and chicken (L); 2xHAR.32 mouse (M) and human (N); TSAR.0878 mouse (O) and chicken (P); TSAR.2216 mouse (Q) and chicken (S) transgenic zebrafish at 24, 48 and 72 hours post-fertilization (hpf). Only one representative transgenic line for each sequence is shown. All transgenic lines for each sequence are shown in Supp. Figs. 2 to 7. Bright-field (J) and fluorescent images (S) of wild-type zebrafish are also shown. Note that the yolk sac (yolk) is autofluorescent. Scale bar: 0.5 mm. Abbreviations: nm, neuromasts of the lateral line; oc, otic capsule.

Figure 2. TSAR.0878-chicken expression analysis in the hearing and lateral line systems in transgenic zebrafish.

(A) At the top: schematic of the transgene containing the TSAR.0878-chicken sequence cloned upstream of the cFos murine minimal promoter and the reporter gene eGFP. Below: detail of the middle part of the TSAR.0878 sequence alignment including the tested sequences (mouse and chicken) and other mammals representative sequences. (B to M) Fluorescent microphotographs of one representative transgenic zebrafish line at 7 days post fertilization (dpf) carrying the TSAR.0878-chicken transgene showing eGFP expression (B, E, H and K), the hair cell-specific marker FM4-64 (C, F, I and L) and the overlay (D, G, J and M). The expression of eGFP coincides with the hair cell marker in neuromasts of the lateral line and in the otic capsule (E, G and F) of the zebrafish. (F) shows a magnification of the otic capsule region. (N, O and P) show *PKNOX2* expression by in situ hybridization at 7 dpf. Arrows indicate the location of neuromasts of the hearing and balance systems expressing *PKNOX2*. Ba, branchial arches.

Figure 3. Functional Diversification of *PKNOX* genes in vertebrates

(A) *PKNOX1* locus in chromosome 21 and (B) *PKNOX2* locus in chromosome 11 of the human genome (GRCh37/hg19). (C) Phylogenetic gene tree reconstruction of *PKNOX* proteins across vertebrate evolution, using hierarchical orthologous groups (HOGs) from the OMA Browser (https://omabrowser.org/). The evolutionary history was inferred using the Neighbor-Joining method. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the JTT matrix-based method and are in the units of the number of amino acid substitutions per site. The analysis involved 218 amino acid sequences. **(**D) Proportion and (E) total number of coding and non-coding conserved elements (PhastCons) in TALE proteins. (F). Functional diversification hypothesis that could explain the regulatory domain gain and loss that resulted into the vertebrate *PKNOX* genes expression pattern.

Figure 4. *Pknox2* expression in the inner ear and gene editing strategy.

(A) Graph built using microarray data from manually collected OHCs and IHCs show *Pknox2* and *Pknox1* expression in adult mice (Liu et al. 2014). (B) Graph built using RNA-Seq data for four types of cochlear cells showing *Pknox2* and *Pknox1* expression (Liu et al. 2018). (C) *Pknox2* and *Pknox1* expression is depicted from RNA-Seq data for GFP+ (hair-cell enriched) cochlear samples at four stages in developing mice (Scheffer et al. 2015). (D) Photomicrographs of immunofluorescence assays showing Pknox2 (green) and Myosin 7 a (red) expression in the inner ear of P8 wild type mice. (E) Schematic of the PKNOX2 gene structure and the strategy developed to generate the mutant mice pedigree lacking *Pknox2*. The site of priming of the RNA guide on exon 4 and the STOP codon generated in exon 5 are indicated. Black boxes indicate 5'UTR exons. (F) Chromatograms of $Pknox2^{+/+}$ and $Pknox2^{-/-}$ loci sequencing showing the deletion induced by Cas9 in the site of priming of the sgRNA guide. (G)

Western blot quantification showing strong *Pknox2* expression in wild type mouse brain samples, a tissue where *Pknox2* is strongly expressed. We observed absence of expression in *Pknox2*$^{-/-}$ mice.

Figure 5. *Pknox2* mutants display hearing impairment.

Hearing assessment of *Pknox2* mutant mice ABRs (A) and DPOAEs (B) threshold measurements in 2-month-old *Pknox2*$^{+/+}$ and *Pknox2*$^{-/-}$ mice at different frequencies (from 5.6 to 45.25 kHz). (C) ABR peak I amplitude at 80 dB. Statistical analysis: non parametric Mann-Whitney test *P<0.05;**P<0.01; and ***P<0.001. (D) Representative confocal images of IHCs synapses from the basal turn of the cochleae immunolabeled for pre-synaptic ribbons (CtBP2-red) and postsynaptic receptor patches (GluA2-green) in *Pknox2*$^{-/-}$ and *Pknox2*$^{+/+}$ mice. AntiCtBP2 antibody also weakly stains IHC nuclei. Scale bar, 7 µm. (E-G) Puncta per IHC. Quantitative data obtained from *Pknox*$^{+/+}$ and *Pknox2*$^{-/-}$mice. For each IHC, we analyzed the number of CtBP2 puncta (E), postsynaptic GluA2 receptor patches (F), and putative ribbon synapses(G). In *Pknox2*$^{-/-}$ mice, an increase in the number of CtBP2 puncta, GluA2 receptor patches and synapses on basal region is observed (*Pknox2*$^{+/+}$ n = 105 IHCs at the apical, 126 IHCs at the medial, and 139 IHCs at the basal from 3 animals; *Pknox2*$^{-/-}$ n = 130 IHCs at the apical, 153 IHCs at the medial, and 137 IHCs at the basal region from 5 animals).

Figure 6. *Pknox2* controls multiple key genes in the inner ear.

(A) Heatmap of top upregulated and downregulated genes by p-value and log2FC, grouped based on their pattern of gene expression according to the *Pknox2*$^{+/+}$ and *Pknox2*$^{-/-}$ samples. Each row represents a gene and each column represents a sample. (B) Scheme depicting $-\log_{10}$(padj) values of GO Terms in g:Profiler for differentially expressed genes in *Pknox2*$^{+/+}$ and *Pknox2*$^{-/-}$. Top terms are numbered and displayed in the table below. (C) Functional association of 57 proteins comprising the Biological Process "transmission of nerve impulse" in STRING (V11). Circle colors indicate whether the protein is up-regulated (red) or down-regulated (blue) in *Pknox2*$^{-/-}$ mice.

References

Altenhoff AM, Train C-M, Gilbert KJ, Mediratta I, Mendes de Farias T, Moi D, Nevers Y, Radoykova H-S, Rossier V, Warwick Vesztrocy A, et al. 2020. OMA orthology in 2021: website overhaul, conserved isoforms, ancestral gene order and more. *Nucleic Acids Res.* [Internet]. Available from: http://dx.doi.org/10.1093/nar/gkaa1007

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* [Internet]. Available from: https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1995.tb02031.x

Bermingham NA, Hassan BA, Price SD, Vollrath MA, Ben-Arie N, Eatock RA, Bellen HJ, Lysakowski A, Zoghbi HY. 1999.

Math1: an essential gene for the generation of inner ear hair cells. *Science* 284:1837–1841.

Bessa J, Tena JJ, de la Calle-Mustienes E, Fernández-Miñán A, Naranjo S, Fernández A, Montoliu L, Akalin A, Lenhard B, Casares F, et al. 2009. Zebrafish enhancer detection (ZED) vector: a new tool to facilitate transgenesis and the functional analysis of cis-regulatory regions in zebrafish. *Dev. Dyn.* 238:2409–2417.

Boero LE, Castagna VC, Terreros G, Moglie MJ, Silva S, Maass JC, Fuchs PA, Delano PH, Elgoyhen AB, Gómez-Casati ME. 2020. Preventing presbycusis in mice with enhanced medial olivocochlear feedback. *Proc. Natl. Acad. Sci. U. S. A.* 117:11811–11819.

Brownell WE. 1990. Outer hair cell electromotility and otoacoustic emissions. *Ear Hear.* 11:82–92.

Cáceres M, Suwyn C, Maddox M, Thomas JW, Preuss TM. 2007. Increased cortical expression of two synaptogenic thrombospondins in human brain evolution. *Cereb. Cortex* 17:2312–2321.

Cannavò E, Khoueiry P, Garfield DA, Geeleher P, Zichner T, Gustafson EH, Ciglar L, Korbel JO, Furlong EEM. 2016. Shadow Enhancers Are Pervasive Features of Developmental Regulatory Networks. *Curr. Biol.* 26:38–51.

Caporale AL, Gonda CM, Franchini LF. 2019. Transcriptional Enhancers in the FOXP2 Locus Underwent Accelerated Evolution in the Human Lineage. *Mol. Biol. Evol.* [Internet]. Available from: http://dx.doi.org/10.1093/molbev/msz173

Capra JA, Erwin GD, McKinsey G, Rubenstein JLR, Pollard KS. 2013. Many human accelerated regions are developmental enhancers. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368:20130025.

Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134:25–36.

Cheatle Jarvela AM, Hinman VF. 2015. Evolution of transcription factor function as a mechanism for changing metazoan developmental gene regulatory networks. *Evodevo* 6:3.

Chen H, Rossier C, Nakamura Y, Lynn A, Chakravarti A, Antonarakis SE. 1997. Cloning of a novel homeobox-containing gene, PKNOX1, and mapping to human chromosome 21q22.3. *Genomics* 41:193–200.

Cortese M, Papal S, Pisciottano F, Elgoyhen AB, Hardelin J-P, Petit C, Franchini LF, El-Amraoui A. 2017. Spectrin βV adaptive mutations and changes in subcellular location correlate with emergence of hair cell electromotility in mammalians. *Proc. Natl. Acad. Sci. U. S. A.* 114:2054–2059.

Coy SE, Borycki A-G. 2010. Expression analysis of TALE family transcription factors during avian development. *Dev. Dyn.* 239:1234–1245.

Dallos P, Wu X, Cheatham MA, Gao J, Zheng J, Anderson CT, Jia S, Wang X, Cheng WHY, Sengupta S, et al. 2008. Prestin-based outer hair cell motility is necessary for mammalian cochlear amplification. *Neuron* 58:333–339.

Dickinson ME, Flenniken AM, Ji X, Teboul L, Wong MD, White JK, Meehan TF, Weninger WJ, Westerberg H, Adissu H, et al. 2016. High-throughput discovery of novel developmental phenotypes. *Nature* 537:508–514.

Di Rosa P, Villaescusa JC, Longobardi E, Iotti G, Ferretti E, Diaz VM, Miccio A, Ferrari G, Blasi F. 2007. The homeodomain transcription factor Prep1 (pKnox1) is required for hematopoietic stem and progenitor cell activity. *Dev. Biol.* 311:324–334.

Domené S, Bumaschny VF, de Souza FSJ, Franchini LF, Nasif S, Low MJ, Rubinstein M. 2013. Enhancer turnover and conserved regulatory function in vertebrate evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368:20130027.

Eckalbar WL, Schlebusch SA, Mason MK, Gill Z, Parker AV, Booker BM, Nishizaki S, Muswamba-Nday C, Terhune E, Nevonen KA, et al. 2016. Transcriptomic and epigenomic characterization of the developing bat wing. *Nature Genetics* [Internet] 48:528–536. Available from: http://dx.doi.org/10.1038/ng.3537

Elgoyhen AB, Franchini LF. 2011. Prestin and the cholinergic receptor of hair cells: positively-selected proteins in mammals. *Hear. Res.* 273:100–108.

Erwin GD, Oksenberg N, Truty RM, Kostka D, Murphy KK, Ahituv N, Pollard KS, Capra JA. 2014. Integrating diverse datasets improves developmental enhancer prediction. *PLoS Comput. Biol.* 10:e1003677.

Ferretti E, Schulz H, Talarico D, Blasi F, Berthelsen J. 1999. The PBX-regulating protein PREP1 is present in different PBX-complexed forms in mouse. *Mech. Dev.* 83:53–64.

Ferretti E, Villaescusa JC, Di Rosa P, Fernandez-Diaz LC, Longobardi E, Mazzieri R, Miccio A, Micali N, Selleri L, Ferrari G, et al. 2006. Hypomorphic mutation of the TALE gene Prep1 (pKnox1) causes a major reduction of Pbx and Meis proteins and a pleiotropic embryonic phenotype. *Mol. Cell. Biol.* 26:5650–5662.

Fisher S, Grice EA, Vinton RM, Bessling SL, McCallion AS. 2006. Conservation of RET regulatory function from human to zebrafish without sequence similarity. *Science* 312:276–279.

Fisher S, Grice EA, Vinton RM, Bessling SL, Urasaki A, Kawakami K, McCallion AS. 2006. Evaluating the biological relevance of putative enhancers using Tol2 transposon-mediated transgenesis in zebrafish. *Nat. Protoc.* 1:1297–1305.

Fognani C, Kilstrup-Nielsen C, Berthelsen J, Ferretti E, Zappavigna V, Blasi F. 2002. Characterization of PREP2, a paralog of PREP1, which defines a novel sub-family of the MEINOX TALE homeodomain transcription factors. *Nucleic Acids Res.* 30:2043–2051.

Franchini LF, Belén Elgoyhen A. 2006. Adaptive evolution in mammalian proteins involved in cochlear outer hair cell electromotility. *Molecular Phylogenetics and Evolution* [Internet] 41:622–635. Available from: http://dx.doi.org/10.1016/j.ympev.2006.05.042

Fritzsch B, Pan N, Jahan I, Duncan JS, Kopecky BJ, Elliott KL, Kersigo J, Yang T. 2013. Evolution and development of the tetrapod auditory system: an organ of Corti-centric perspective. *Evol. Dev.* 15:63–79.

Gates GR, Saunders JC, Bock GR, Aitkin LM, Elliott MA. 1974. Peripheral auditory function in the platypus, Ornithorhynchus anatinus. *J. Acoust. Soc. Am.* 56:152–156.

Ghaffari R, Aranyosi AJ, Richardson GP, Freeman DM. 2010. Tectorial membrane travelling waves underlie abnormal hearing in Tectb mutant mice. *Nat. Commun.* 1:96.

Han H, Cho J-W, Lee S, Yun A, Kim H, Bae D, Yang S, Kim CY, Lee M, Kim E, et al. 2018. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res.* 46:D380–D386.

Heffner HE, Heffner RS. 2018. The evolution of mammalian hearing. *AIP Conf. Proc.* 1965:130001.

Hickman TT, Hashimoto K, Liberman LD, Liberman MC. 2021. Cochlear Synaptic Degeneration and Regeneration After Noise: Effects of Age and Neuronal Subgroup. *Front. Cell. Neurosci.* 15:684706.

Hoekstra HE, Coyne JA. 2007. The locus of evolution: evo devo and the genetics of adaptation. *Evolution* 61:995–1016.

Holloway AK, Bruneau BG, Sukonnik T, Rubenstein JL, Pollard KS. 2016. Accelerated Evolution of Enhancer Hotspots in the Mammal Ancestor. *Mol. Biol. Evol.* 33:1008–1018.

Imoto I, Sonoda I, Yuki Y, Inazawa J. 2001. Identification and characterization of human PKNOX2, a novel homeobox-containing gene. *Biochem. Biophys. Res. Commun.* 287:270–276.

Izumikawa M, Minoda R, Kawamoto K, Abrashkin KA, Swiderski DL, Dolan DF, Brough DE, Raphael Y. 2005. Auditory hair cell replacement and hearing improvement by Atoh1 gene therapy in deaf mammals. *Nat. Med.* 11:271–276.

Ji L, Martel DT, Borges BC, Wu C, Charles Liberman M, Shore SE, Corfas G. 2022. Inner hair cell synapse density influences auditory processing. *bioRxiv* [Internet]:2022.05.02.490340. Available from: https://www.biorxiv.org/content/biorxiv/early/2022/05/02/2022.05.02.490340

Kaessmann H. 2010. Origins, evolution, and phenotypic impact of new genes. *Genome Res.* 20:1313–1326.

Kammerer R, Rüttiger L, Riesenberg R, Schäuble C, Krupar R, Kamp A, Sunami K, Eisenried A, Hennenberg M, Grunert F, et al. 2012. Loss of mammal-specific tectorial membrane component carcinoembryonic antigen cell adhesion molecule 16 (CEACAM16) leads to hearing impairment at low and high frequencies. *J. Biol. Chem.* 287:21584–21598.

Kamm GB, López-Leal R, Lorenzo JR, Franchini LF. 2013. A fast-evolving human NPAS3 enhancer gained reporter expression in the developing forebrain of transgenic mice. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368:20130019.

Kamm GB, Pisciottano F, Kliger R, Franchini LF. 2013. The developmental brain gene NPAS3 contains the largest number of accelerated regulatory sequences in the human genome. *Mol. Biol. Evol.* 30:1088–1102.

Kawakami K. 2007. Tol2: a versatile gene transfer vector in vertebrates. *Genome Biol.* 8 Suppl 1:S7.

Kawakami K, Takeda H, Kawakami N, Kobayashi M, Matsuda N, Mishina M. 2004. A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish. *Dev. Cell* 7:133–144.

Khimich D, Nouvian R, Pujol R, Tom Dieck S, Egner A, Gundelfinger ED, Moser T. 2005. Hair cell synaptic ribbons are essential for synchronous auditory signalling. *Nature* 434:889–894.

Kiernan AE, Pelling AL, Leung KKH, Tang ASP, Bell DM, Tease C, Lovell-Badge R, Steel KP, Cheah KSE. 2005. Sox2 is required for sensory organ development in the mammalian inner ear. *Nature* 434:1031–1035.

Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37:907–915.

King MC, Wilson AC. 1975. Evolution at two levels in humans and chimpanzees. *Science* 188:107–116.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* 35:1547–1549.

Law CW, Chen Y, Shi W, Smyth GK. 2014. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* 15:R29.

Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30:923–930.

Liberman LD, Wang H, Liberman MC. 2011. Opposing gradients of ribbon size and AMPA receptor expression underlie sensitivity differences among cochlear-nerve/hair-cell synapses. *J. Neurosci.* 31:801–808.

Liberman MC, Gao J, He DZZ, Wu X, Jia S, Zuo J. 2002. Prestin is required for electromotility of the outer hair cell and for the cochlear amplifier. *Nature* 419:300–304.

Li G, Wang J, Rossiter SJ, Jones G, Cotton JA, Zhang S. 2008. The hearing gene Prestin reunites echolocating bats. *Proc. Natl. Acad. Sci. U. S. A.* 105:13959–13964.

Li J, Zhang T, Ramakrishnan A, Fritzsch B, Xu J, Wong EYM, Loh Y-HE, Ding J, Shen L, Xu P-X. 2020. Dynamic changes in cis-regulatory occupancy by Six1 and its cooperative interactions with distinct cofactors drive lineage-specific gene expression programs during progressive differentiation of the auditory sensory epithelium. *Nucleic Acids Res.* 48:2880–2896.

Liu H, Chen L, Giffen KP, Stringham ST, Li Y, Judge PD, Beisel KW, He DZZ. 2018. Cell-Specific Transcriptome Analysis Shows That Adult Pillar and Deiters' Cells Express Genes Encoding Machinery for Specializations of Cochlear Hair Cells. *Front. Mol. Neurosci.* 11:356.

Liu H, Leslie EJ, Carlson JC, Beaty TH, Marazita ML, Lidral AC, Cornell RA. 2017. Identification of common non-coding variants at 1p22 that are functional for non-syndromic orofacial clefting. *Nat. Commun.* 8:14759.

Liu H, Pecka JL, Zhang Q, Soukup GA, Beisel KW, He DZZ. 2014. Characterization of transcriptomes of cochlear inner and outer hair cells. *J. Neurosci.* 34:11085–11095.

Liu Y, Cotton JA, Shen B, Han X, Rossiter SJ, Zhang S. 2010. Convergent sequence evolution between echolocating bats and dolphins. *Curr. Biol.* 20:R53–R54.

Li Y, Liu H, Barta CL, Judge PD, Zhao L, Zhang WJ, Gong S, Beisel KW, He DZZ. 2016. Transcription Factors Expressed in Mouse Cochlear Inner and Outer Hair Cells. *PLoS One* 11:e0151291.

Li Y, Liu Z, Shi P, Zhang J. 2010. The hearing gene Prestin unites echolocating bats and whales. *Current Biology* [Internet] 20:R55–R56. Available from: http://dx.doi.org/10.1016/j.cub.2009.11.042

Lynch VJ, Wagner GP. 2008. Resurrecting the role of transcription factor change in developmental evolution. *Evolution* 62:2131–2154.

Maison SF, Usubuchi H, Liberman MC. 2013. Efferent feedback minimizes cochlear neuropathy from moderate noise exposure. *J. Neurosci.* 33:5542–5552.

Manley GA. 2000. Cochlear mechanisms from a phylogenetic viewpoint. *Proc. Natl. Acad. Sci. U. S. A.* 97:11736–11743.

Manley GA. 2010. The origin and evolution of high-frequency hearing in (most) mammals. *Hear. Res.* 270:2–3.

Manley GA. 2012. Evolutionary paths to mammalian cochleae. *J. Assoc. Res. Otolaryngol.* 13:733–743.

Manley GA. 2017. Comparative Auditory Neuroscience: Understanding the Evolution and Function of Ears. *J. Assoc. Res. Otolaryngol.* 18:1–24.

Mann A, Bhatia S. 2019. Zebrafish: A Powerful Model for Understanding the Functional Relevance of Noncoding Region Mutations in Human Genetic Diseases. *Biomedicines* [Internet] 7. Available from: http://dx.doi.org/10.3390/biomedicines7030071

Matern MS, Milon B, Lipford EL, McMurray M, Ogawa Y, Tkaczuk A, Song Y, Elkon R, Hertzano R. 2020. GFI1 functions to repress neuronal gene expression in the developing inner ear hair cells. *Development* [Internet] 147. Available from: http://dx.doi.org/10.1242/dev.186015

Matsubara A, Laake JH, Davanger S, Usami S, Ottersen OP. 1996. Organization of AMPA receptor subunits at a glutamate synapse: a quantitative immunogold analysis of hair cell synapses in the rat organ of Corti. *J. Neurosci.* 16:4457–4467.

Merabet S, Mann RS. 2016. To Be Specific or Not: The Critical Relationship Between Hox And TALE Proteins. *Trends Genet.* 32:334–347.

Mills DM, Shepherd RK. 2001. Distortion product otoacoustic emission and auditory brainstem responses in the echidna (Tachyglossus aculeatus). *J. Assoc. Res. Otolaryngol.* 2:130–146.

Morrill S, He D. 2020. The effects of Bcl6 KO in mouse model of hearing. *FASEB J.* 34:1–1.

Nicolson T. 2017. The genetics of hair-cell function in zebrafish. *J. Neurogenet.* 31:102–112.

Nowick K, Stubbs L. 2010. Lineage-specific transcription factors and the evolution of gene regulatory networks. *Brief. Funct. Genomics* 9:65–78.

Oksenberg N, Stevison L, Wall JD, Ahituv N. 2013. Function and regulation of AUTS2, a gene implicated in autism and human evolution. *PLoS Genet.* 9:e1003221.

Osterwalder M, Barozzi I, Tissières V, Fukuda-Yuzawa Y, Mannion BJ, Afzal SY, Lee EA, Zhu Y, Plajzer-Frick I, Pickle CS, et al. 2018. Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* [Internet] 554:239–243. Available from: http://dx.doi.org/10.1038/nature25461

Papatheodorou I, Moreno P, Manning J, Fuentes AM-P, George N, Fexova S, Fonseca NA, Füllgrabe A, Green M, Huang N, et al. 2020. Expression Atlas update: from tissues to single cells. *Nucleic Acids Res.* 48:D77–D83.

Pisciottano F, Cinalli AR, Stopiello JM, Castagna VC, Elgoyhen AB, Rubinstein M, Gómez-Casati ME, Franchini LF. 2019. Inner Ear Genes Underwent Positive Selection and Adaptation in the Mammalian Lineage. *Mol. Biol. Evol.* 36:1653–1670.

Prud'homme B, Gompel N, Carroll SB. 2007. Emerging principles of regulatory evolution. *Proc. Natl. Acad. Sci. U. S. A.* 104 Suppl 1:8605–8612.

Ranum PT, Goodwin AT, Yoshimura H, Kolbe DL, Walls WD, Koh J-Y, He DZZ, Smith RJH. 2019. Insights into the Biology of Hearing and Deafness Revealed by Single-Cell RNA Sequencing. *Cell Rep.* 26:3160–3171.e3.

Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J. 2019. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* 47:W191–W198.

Russell IJ, Legan PK, Lukashkina VA, Lukashkin AN, Goodyear RJ, Richardson GP. 2007. Sharpened cochlear tuning in a mouse with a genetically modified tectorial membrane. *Nat. Neurosci.* 10:215–223.

Scheffer DI, Shen J, Corey DP, Chen Z-Y. 2015. Gene Expression by Mouse Inner Ear Hair Cells during Development. *J. Neurosci.* 35:6366–6380.

Schmitt AD, Hu M, Jung I, Xu Z, Qiu Y, Tan CL, Li Y, Lin S, Lin Y, Barr CL, et al. 2016. A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. *Cell Rep.* 17:2042–2059.

Scornavacca C, Belkhir K, Lopez J, Dernat R, Delsuc F, Douzery EJP, Ranwez V. 2019. OrthoMaM v10: Scaling-Up Orthologous Coding Sequence and Exon Alignments with More than One Hundred Mammalian Genomes. *Mol. Biol. Evol.* 36:861–862.

Sheets L, Holmgren M, Kindt KS. 2021. How Zebrafish Can Drive the Future of Genetic-based Hearing and Balance Research. *J. Assoc. Res. Otolaryngol.* 22:215–235.

Shera CA, Guinan JJ Jr. 1999. Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs. *J. Acoust. Soc. Am.* 105:782–798.

Shrestha BR, Chia C, Wu L, Kujawa SG, Liberman MC, Goodrich LV. 2018. Sensory Neuron Diversity in the Inner Ear Is Shaped by Activity. *Cell* 174:1229–1246.e17.

Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K, Kosakovsky Pond SL. 2015. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol. Biol. Evol.* 32:1342–1353.

Sun S, Babola T, Pregernig G, So KS, Nguyen M, Su S-SM, Palermo AT, Bergles DE, Burns JC, Müller U. 2018. Hair Cell Mechanotransduction Regulates Spontaneous Activity and Spiral Ganglion Subtype Specification in the Auditory System. *Cell* 174:1247–1263.e15.

Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, et al. 2019. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47:D607–D613.

Tang F, Chen X, Jia L, Li H, Li J, Yuan W. 2019. Differential Gene Expression Patterns Between Apical and Basal Inner Hair Cells Revealed by RNA-Seq. *Front. Mol. Neurosci.* 12:332.

Trigila AP, Pisciottano F, Franchini LF. 2021. Hearing loss genes reveal patterns of adaptive evolution at the coding and non-coding levels in mammals. *BMC Biol.* 19:244.

Vikhe Patil K, Canlon B, Cederroth CR. 2015. High quality RNA extraction of the mammalian cochlea for qRT-PCR and transcriptome analyses. *Hear. Res.* 325:42–48.

Wagner GP, Lynch VJ. 2010. Evolutionary novelties. *Curr. Biol.* 20:R48–R52.

Wallis D, Hamblen M, Zhou Y, Venken KJT, Schumacher A, Grimes HL, Zoghbi HY, Orkin SH, Bellen HJ. 2003. The zinc finger transcription factor Gfi1, implicated in lymphomagenesis, is required for inner ear hair cell differentiation and survival. *Development* 130:221–232.

Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R. 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153:910–918.

Wang H, Zhao H, Sun K, Huang X, Jin L, Feng J. 2020. Evolutionary Basis of High-Frequency Hearing in the Cochleae of Echolocators Revealed by Comparative Genomics. *Genome Biol. Evol.* 12:3740–3753.

Whitfield TT. 2002. Zebrafish as a model for hearing and deafness. *J. Neurobiol.* 53:157–171.

Won H, de la Torre-Ubieta L, Stein JL, Parikshak NN, Huang J, Opland CK, Gandal MJ, Sutton GJ, Hormozdiari F, Lu D, et al. 2016. Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature* 538:523–527.

Xiang M, Gan L, Li D, Chen Z-Y, Zhou L, O'Malley BW, Klein W, Nathans J. 1997. Essential role of POU–domain factor Brn-3c in auditory and vestibular hair cell development. *Proceedings of the National Academy of Sciences* 94:9445–9450.

Yamashita T, Zheng F, Finkelstein D, Kellard Z, Carter R, Rosencrance CD, Sugino K, Easton J, Gawad C, Zuo J. 2018. High-resolution transcriptional dissection of in vivo Atoh1-mediated hair cell conversion in mature cochleae identifies Isl1 as a co-reprogramming factor. *PLoS Genet.* 14:e1007552.

Yizhar-Barnea O, Valensisi C, Jayavelu ND, Kishore K, Andrus C, Koffler-Brill T, Ushakov K, Perl K, Noy Y, Bhonker Y, et al. 2018. DNA methylation dynamics during embryonic development and postnatal maturation of the mouse auditory sensory epithelium. *Sci. Rep.* 8:17348.

Zhang KD, Coate TM. 2017. Recent advances in the development and function of type II spiral ganglion neurons in the mammalian inner ear. *Semin. Cell Dev. Biol.* 65:80–87.

Zheng JL, Shou J, Guillemot F, Kageyama R, Gao WQ. 2000. Hes1 is a negative regulator of inner ear hair cell differentiation. *Development* 127:4551–4560.

Zheng J, Miller KK, Yang T. 2011. Carcinoembryonic antigen-related cell adhesion molecule 16 interacts with α-tectorin and is mutated in autosomal dominant hearing loss (DFNA4). *Proceedings of the* [Internet]. Available from: https://www.pnas.org/content/108/10/4218.short

Zheng J, Shen W, He DZZ, Long KB, Madison LD, Dallos P. 2000. Prestin is the motor protein of cochlear outer hair cells. *Nature* [Internet] 405:149–155. Available from: http://dx.doi.org/10.1038/35012009

Zheng W, Huang L, Wei Z-B, Silvius D, Tang B, Xu P-X. 2003. The role of Six1 in mammalian auditory system development. *Development* 130:3989–4000.

Khimich D, Nouvian R, Pujol R, Tom Dieck S, Egner A, Gundelfinger ED,Moser T (2005) Hair cell synaptic ribbons are essential for synchronous auditory signalling. *Nature* 434:889 – 894.

Liberman LD, Wang H, Liberman MC (2011) Opposing gradients of ribbonsize and AMPA receptor expression underlie sensitivity differences amongcochlear-nerve/hair-cell synapses. *J Neurosci* 31:801– 808.

Maison SF, Usubuchi H, Liberman MC (2013) Efferent feedback minimizes cochlear neuropathy from moderate noise exposure. *J Neurosci* 33:5542–5552.

Matsubara A, Laake JH, Davanger S, Usami S, Ottersen OP (1996) Organi-zation of AMPA receptor subunits at a glutamate synapse: a quantitativeImmunogold analysis of hair cell synapses in the rat organ of Corti. *J Neurosci* 16:4457– 4467.

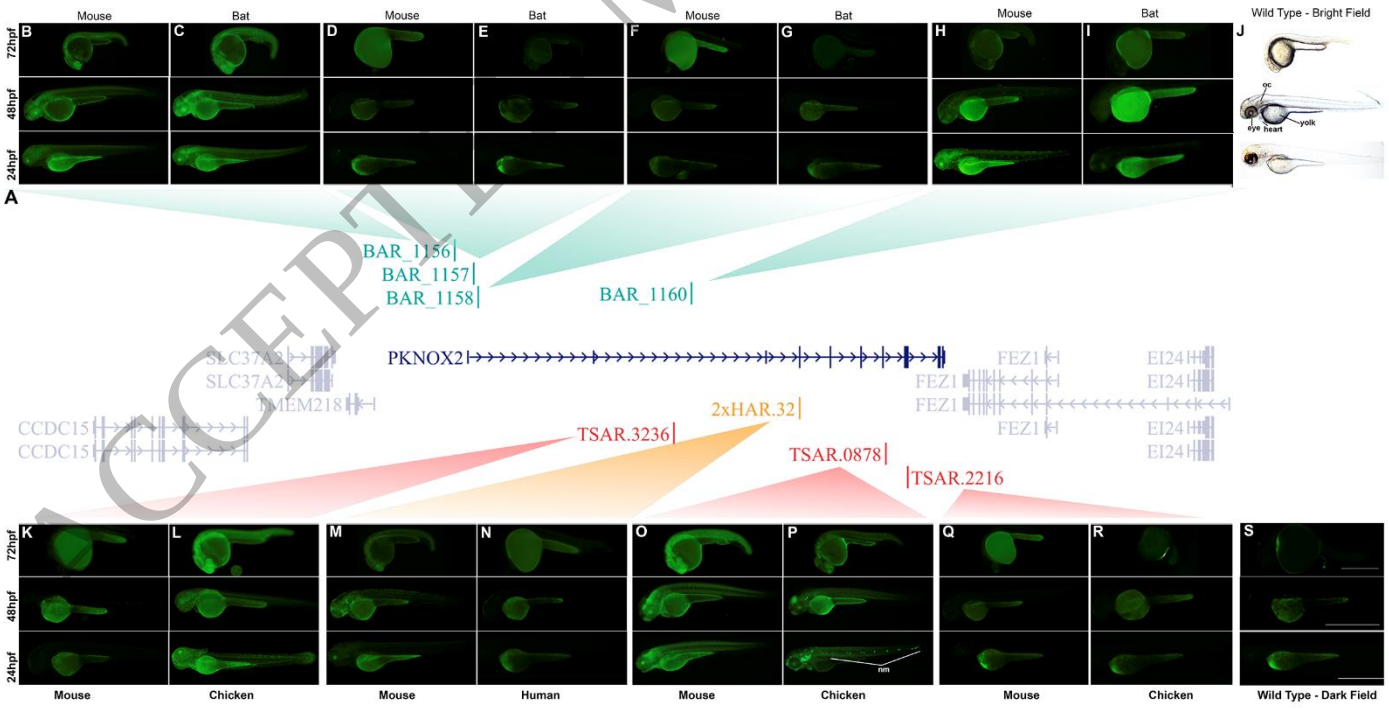

Figure 1
210x106 mm (x DPI)

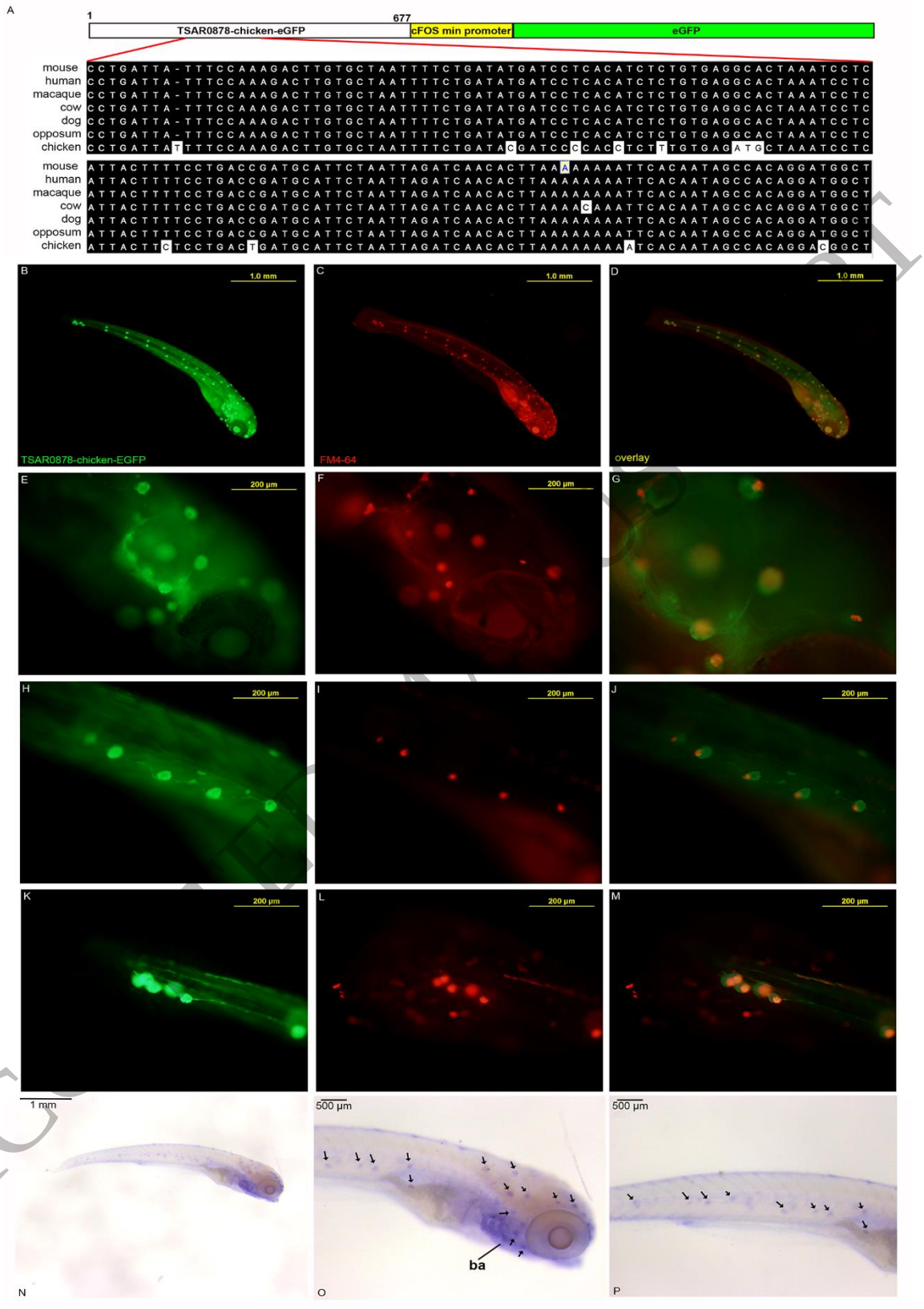

Figure 2
150x233 mm (x DPI)

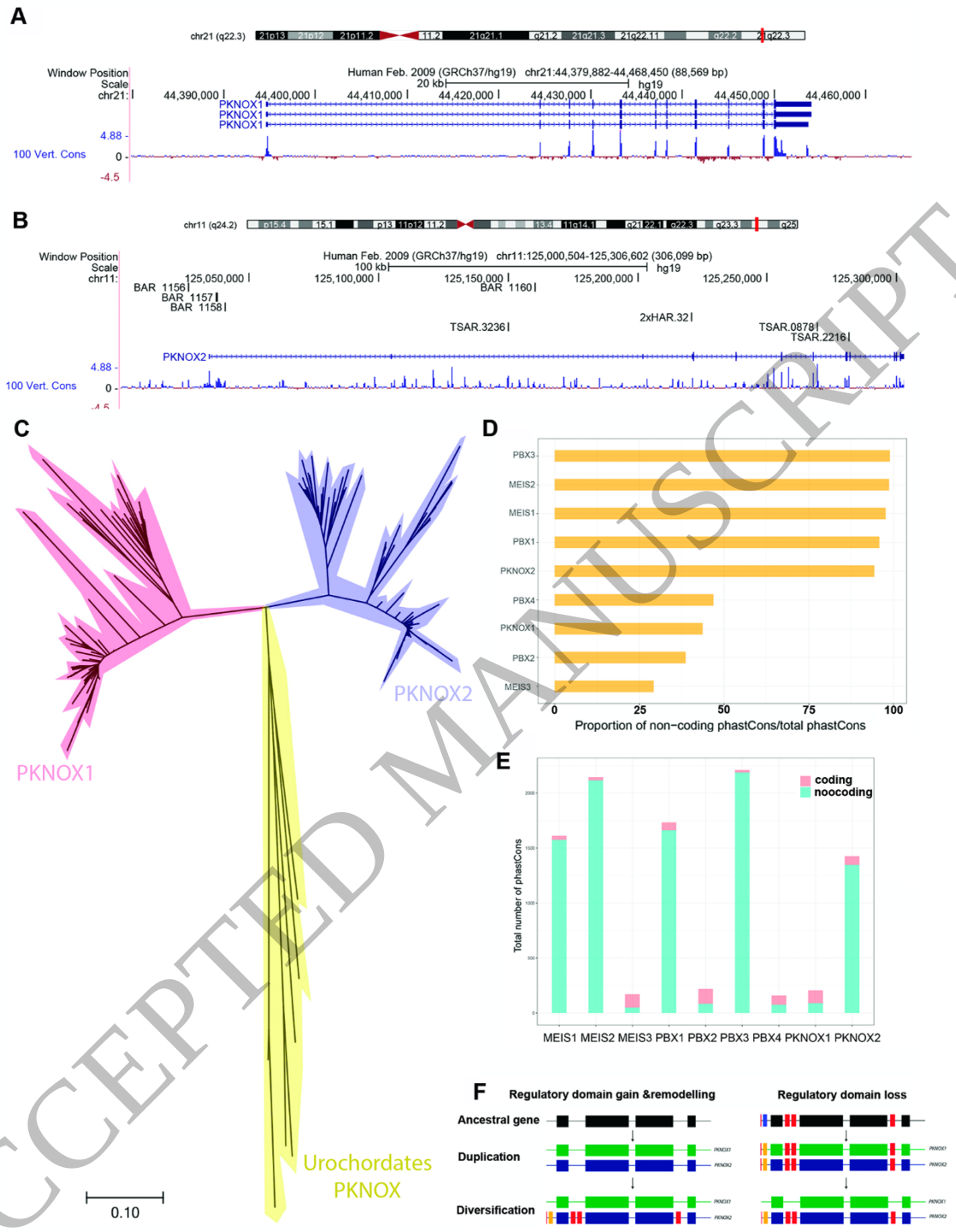

Figure 3
150x200 mm (x DPI)

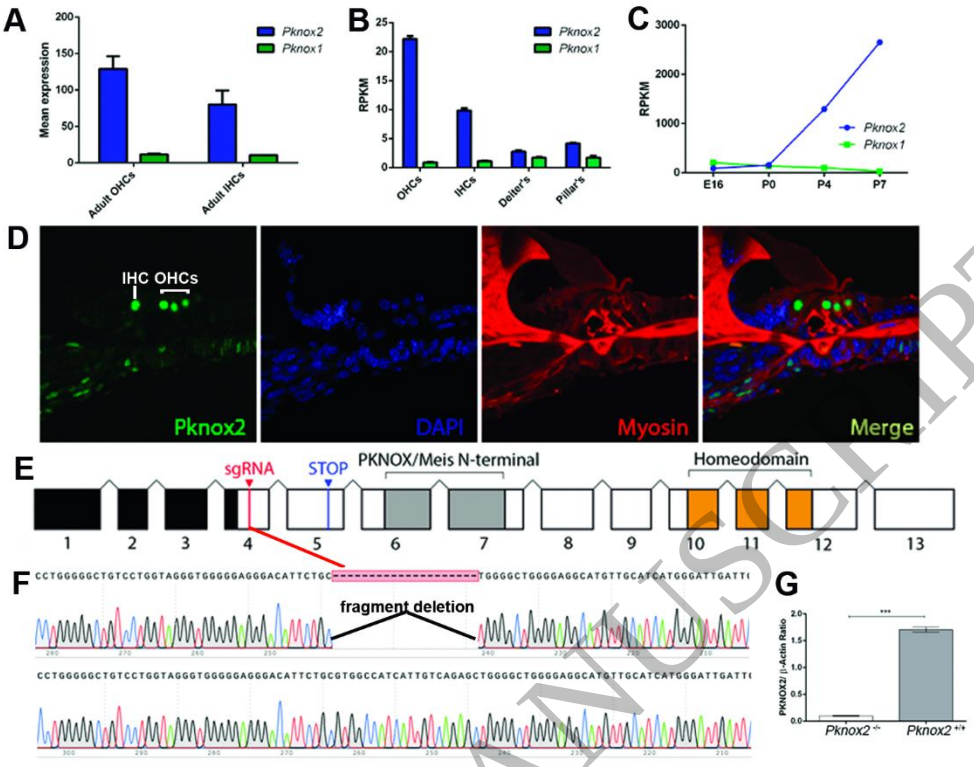

Figure 4
130x108 mm (x DPI)

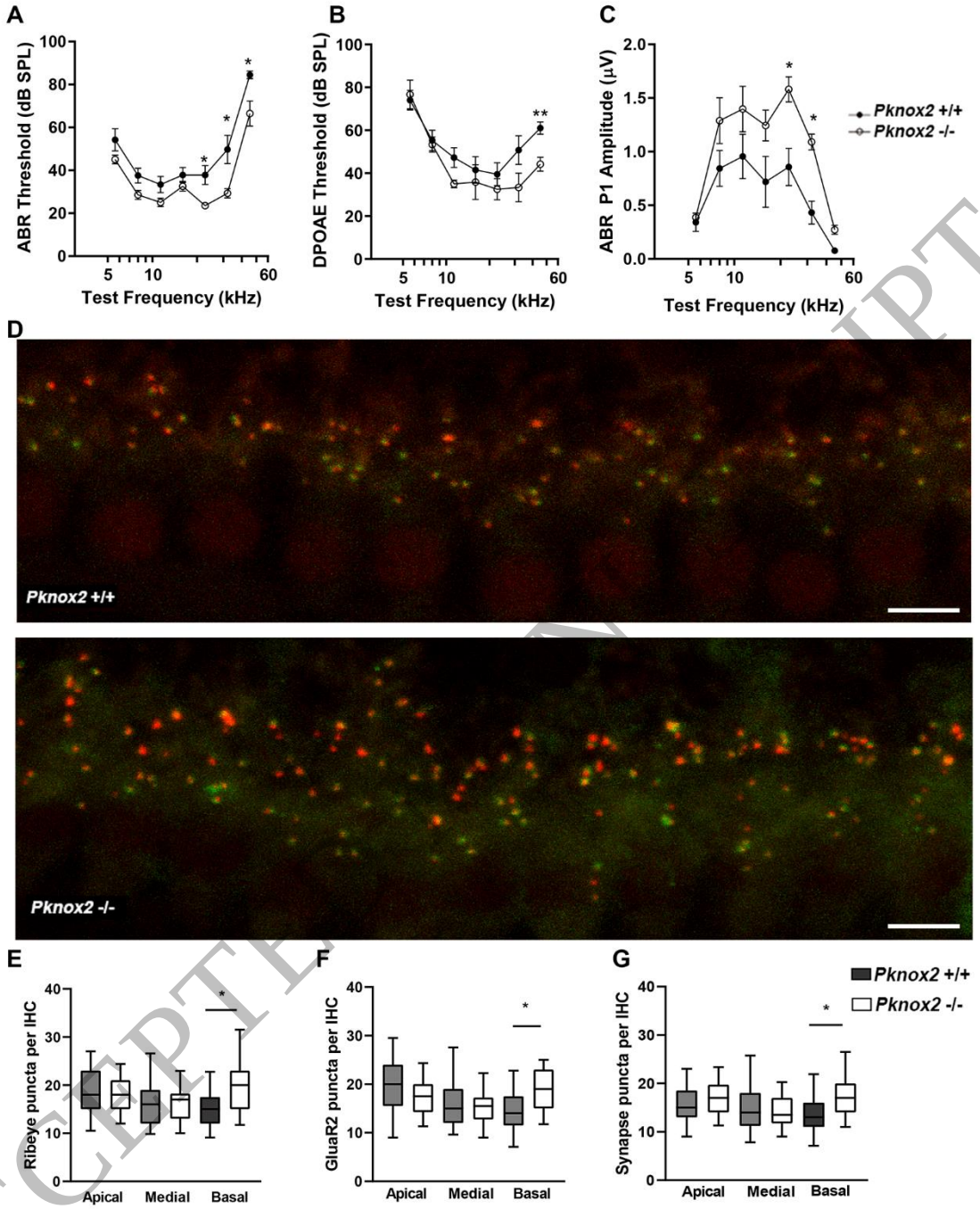

Figure 5
150x184 mm (x DPI)

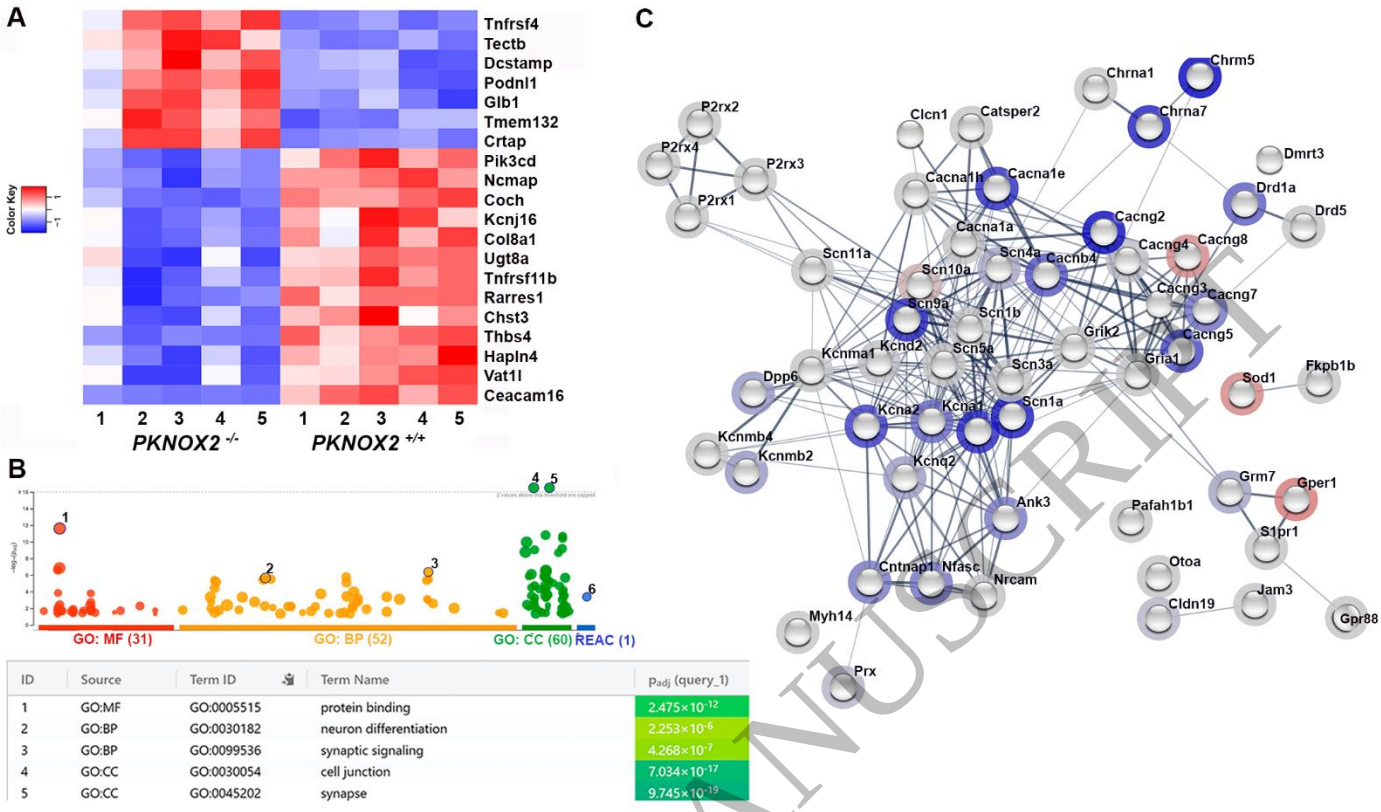

Figure 6
190x110 mm (x DPI)s