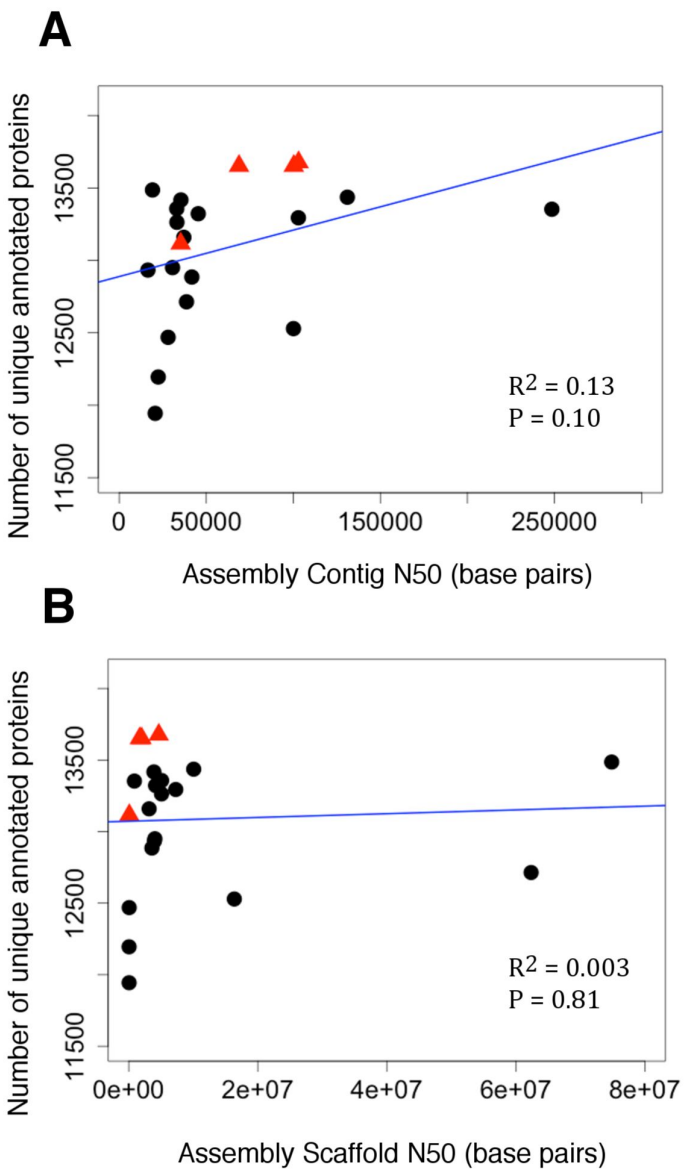

Fig. S1. Contig and scaffold N50 are poor predictors of genome completeness.
(A) Neither contig N50 nor **(B)** scaffold N50 predicts the total number of unique proteins in a given genome assembly. We used 14,052 high quality chicken proteins to annotate avian genomes. Blue line is the linear regression model. Genomes reported in this study are red triangles and other published avian genomes are black circles (see table S2 for full details).

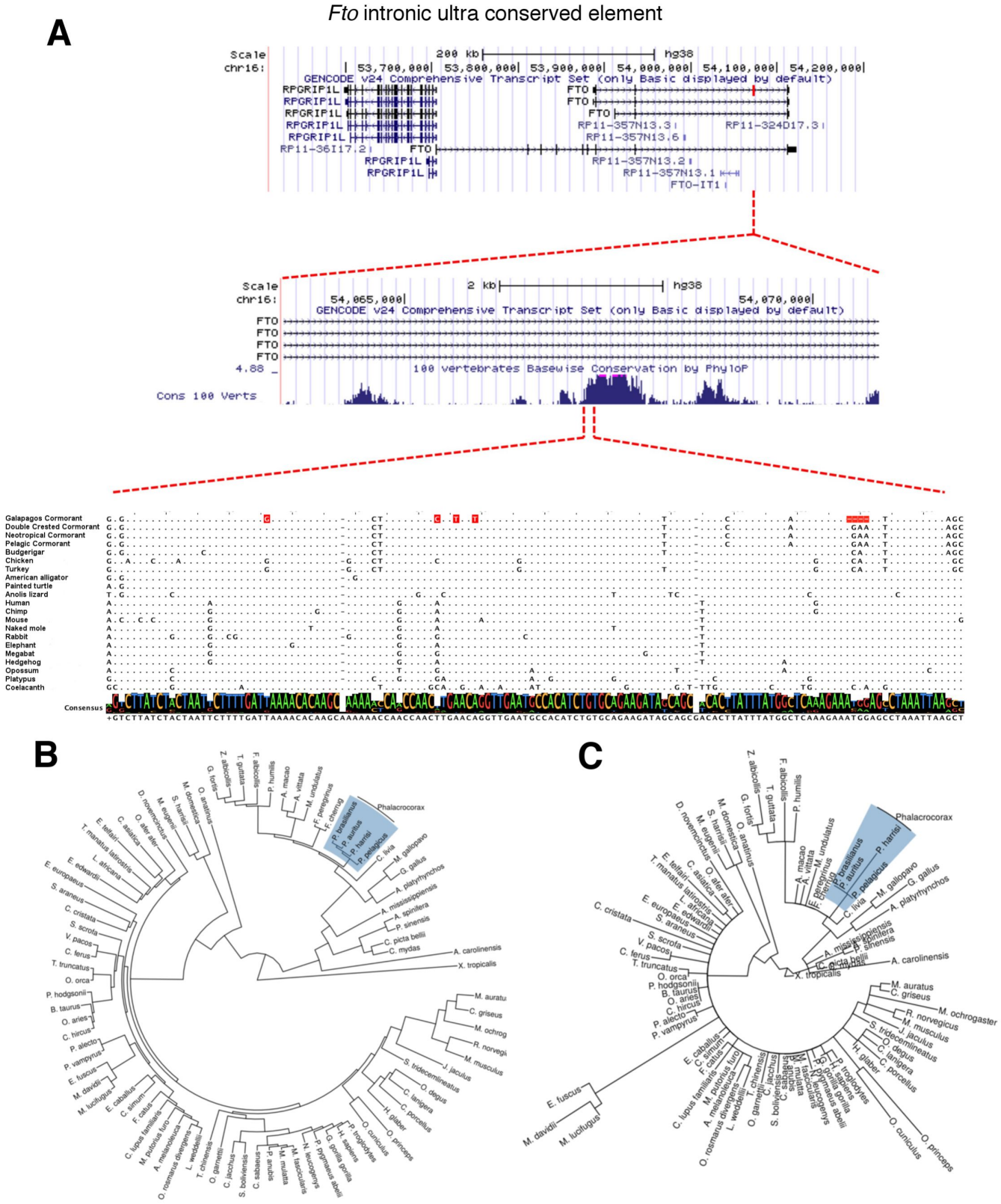

Fig. S2. *FTO* intronic element shows accelerated evolution in *P. harrisi*

(A) Localization of *FTO* in the human genome (Chr16, hg38, UCSC genome browser). The accelerated ultra-conserved element is located with the last intron (8/8) of *FTO*. Alignment of *FTO* alignment includes a representative set of tetrapods and includes also the Coelacanth that was not included in the analysis for accelerated regions. Nucleotides which are different in *P. harris*i compared to the flighted relatives are highlighted in red. Phylogenetic trees estimated from **(B)** all conserved elements and **(C)** only the *FTO* intronic element accelerated in *P. harrisi*. The four sequenced cormorants are highlighted in blue.

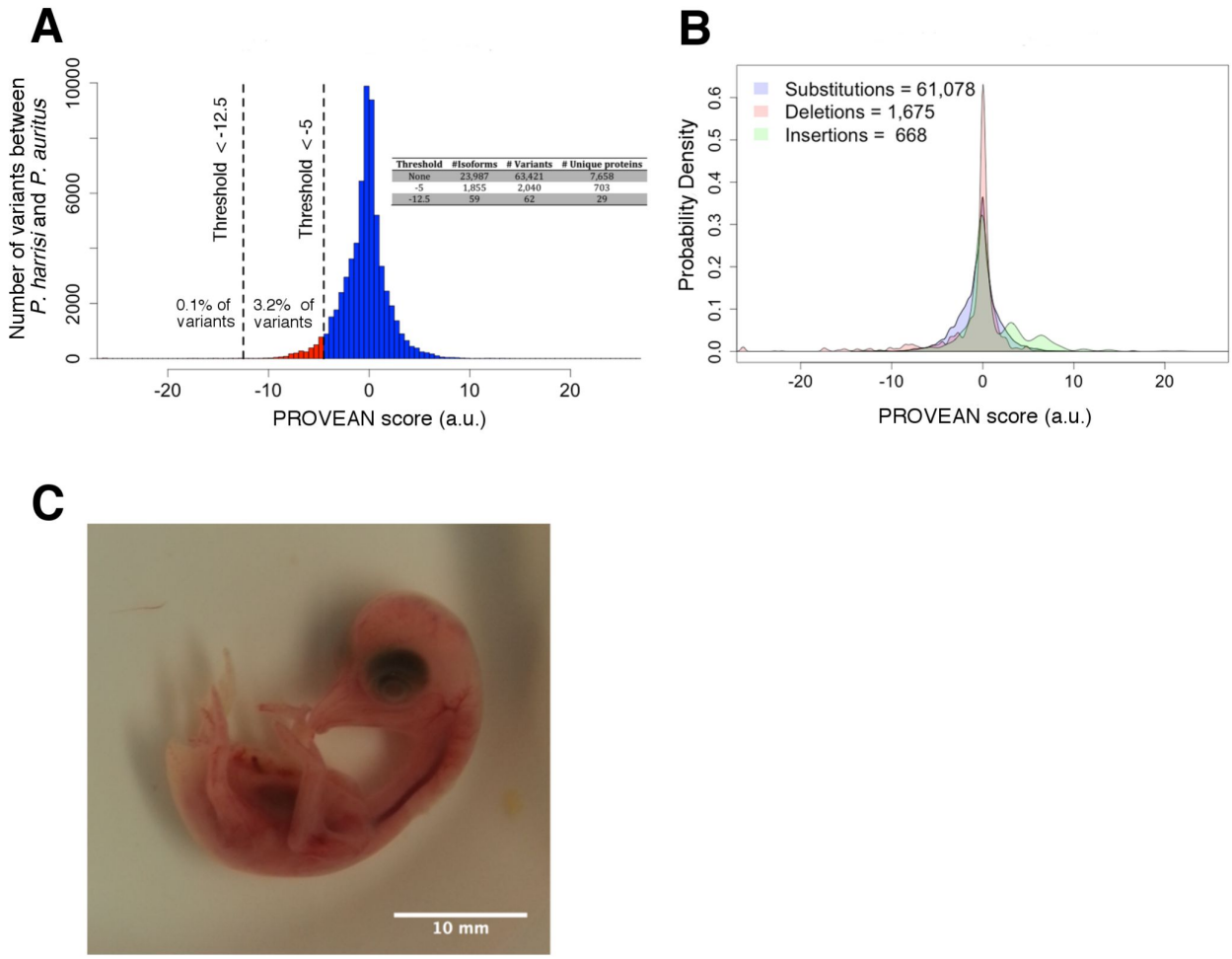

Fig. S3. Distribution of the effect of variants between *P. auritus* and *P. harrisi* from transcriptome predictions.

(A) We used PROVEAN to predict the effect on protein function of 63,421 variants contained in 23,987 orthologous isoforms (7,658 unique proteins) between *P. auritus* and *P. harrisi*. These isoforms were predicted from a de novo transcriptome assembly from the wing of a developing embryo. PROVEAN score thresholds used in this study are drawn as vertical dashed lines. Number of proteins and variants found for each threshold are presented in inset table. **(B)** Density of PROVEAN scores for each class of variant. The same variants presented in (A) were classified as single amino acid substitutions, deletions, and insertions. Number of variants in each class is indicated in the legend. **(C)** The double-crested cormorant embryo used to generate the reference genome and wing transcriptome.

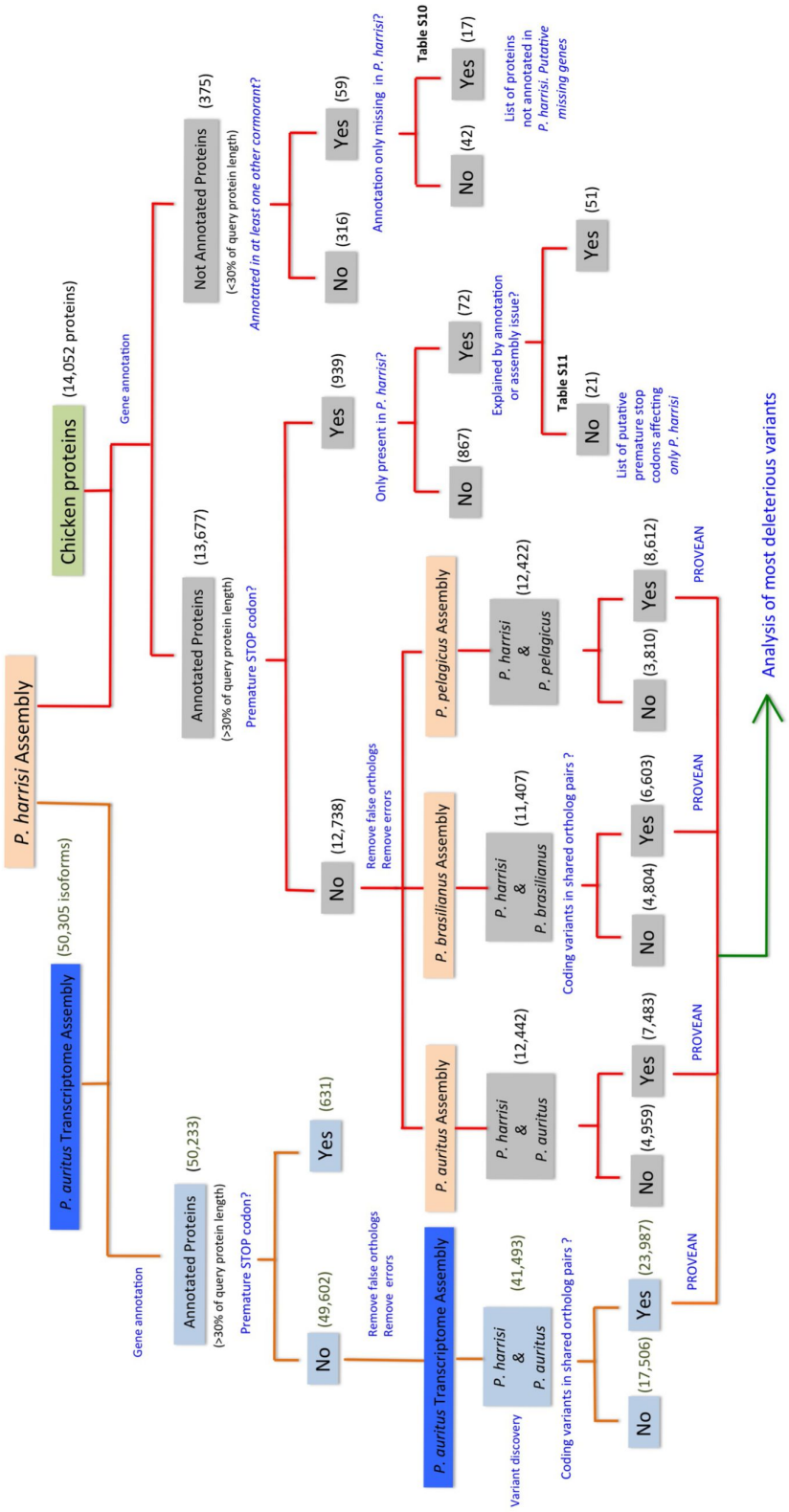

Fig. S4. Flowchart for the discovery and characterization of coding variants in *P. harrisi*.
Detailed flowchart of the annotation and characterization of coding variants in the Galapagos Cormorant that incorporates information from four cormorant genome assemblies (*P. auritus, P. harrisi, P. brasilianus, P. pelagicus*) and a de novo wing transcriptome assembly for *P. auritus*.

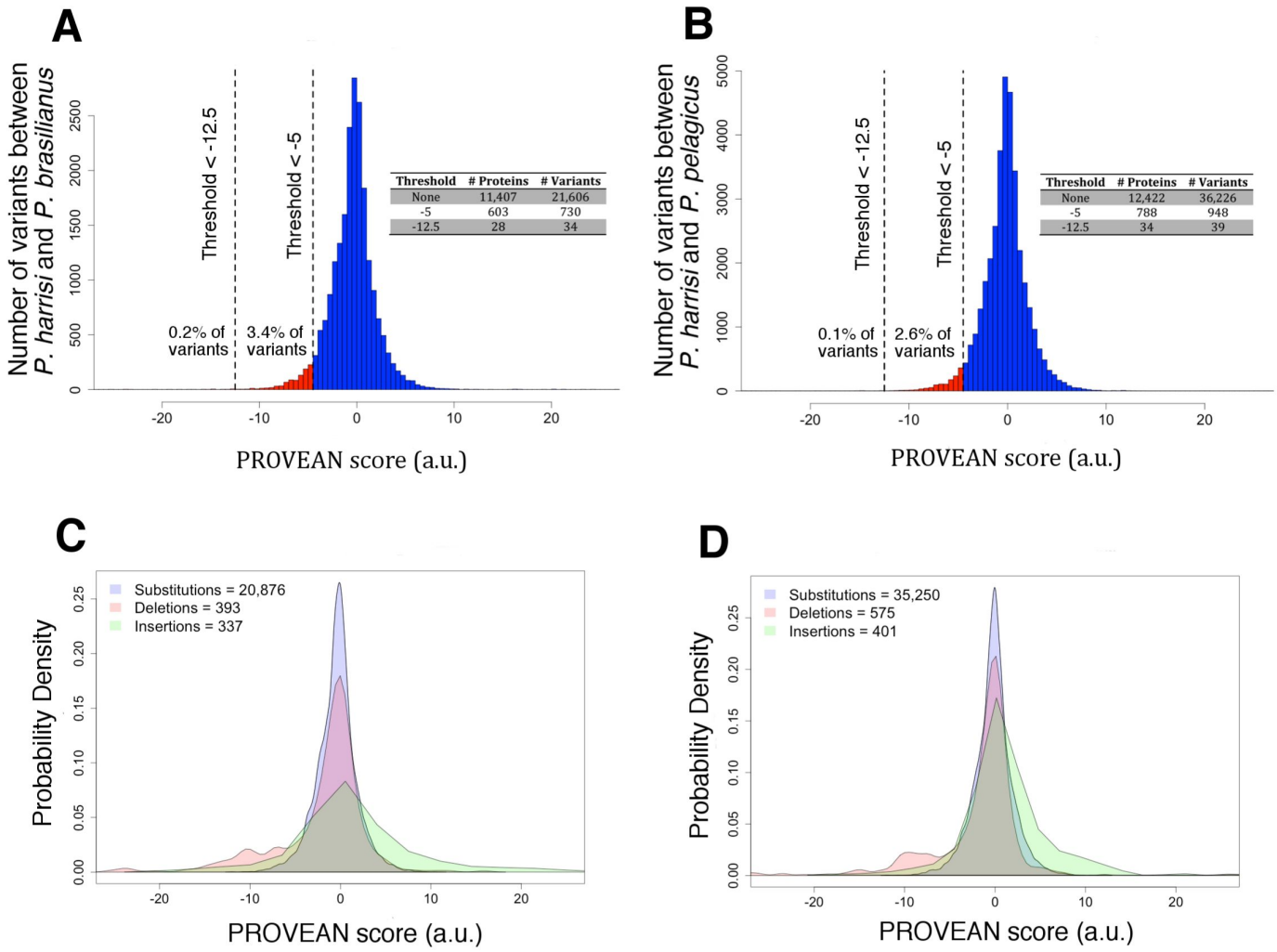

Fig. S5. Distribution of the effect of variants between *P. brasilianus* and *P. harrisi*, and between *P. pelagicus* and *P. harrisi*.

(A) We used PROVEAN to predict the effect on protein function of 21,606 variants contained in 11,407 orthologous pairs between *P. brasilianus* and *P. harrisi*. 4,804 pairs contained no variants. **(B)** Analogous to a, we used PROVEAN to predict the effect on protein function of 36,226 variants contained in 12,422 orthologous pairs between *P. pelagicus* and *P. harrisi*. 3,810 pairs contained no variants. PROVEAN score thresholds used in this study are drawn as vertical dashed lines. Number of proteins and variants found for each threshold are presented in inset table. **(C)** and **(D)** Density of PROVEAN scores for each class of variant. The same variants presented in (A) and (B) were classified as single amino acid substitutions, deletions, and insertions. Number of variants in each class is indicated in the legend.

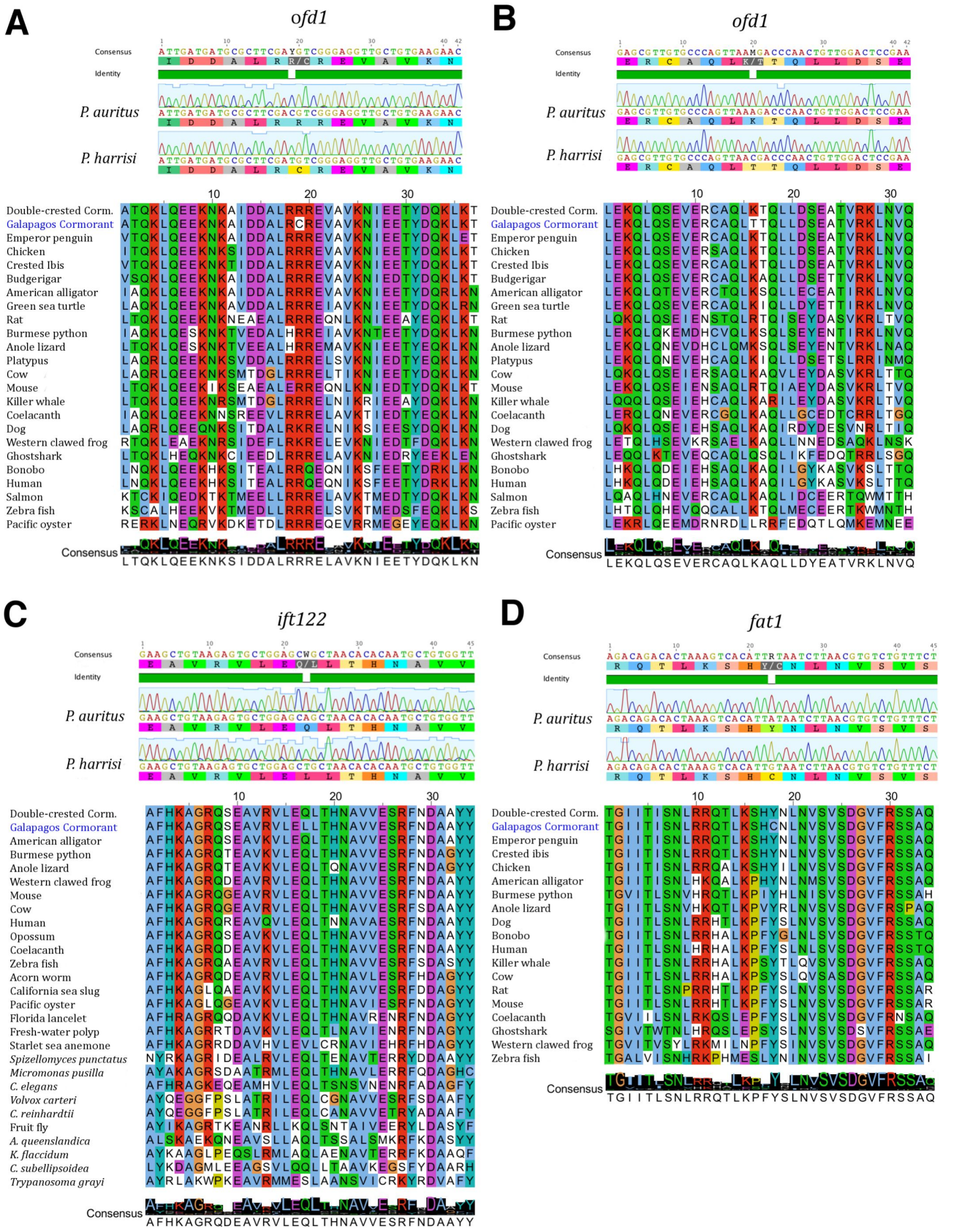

Fig. S6. Sanger sequencing confirmation and evolutionary conservation of function-altering variants in *P. harrisi*.

Sanger sequence confirmation and protein alignment for the following variants: (A), OFD1 (R325C) (B), OFD1 (K517T) (C), IFT122 (Q691L) and (D), FAT1 (Y2462C).

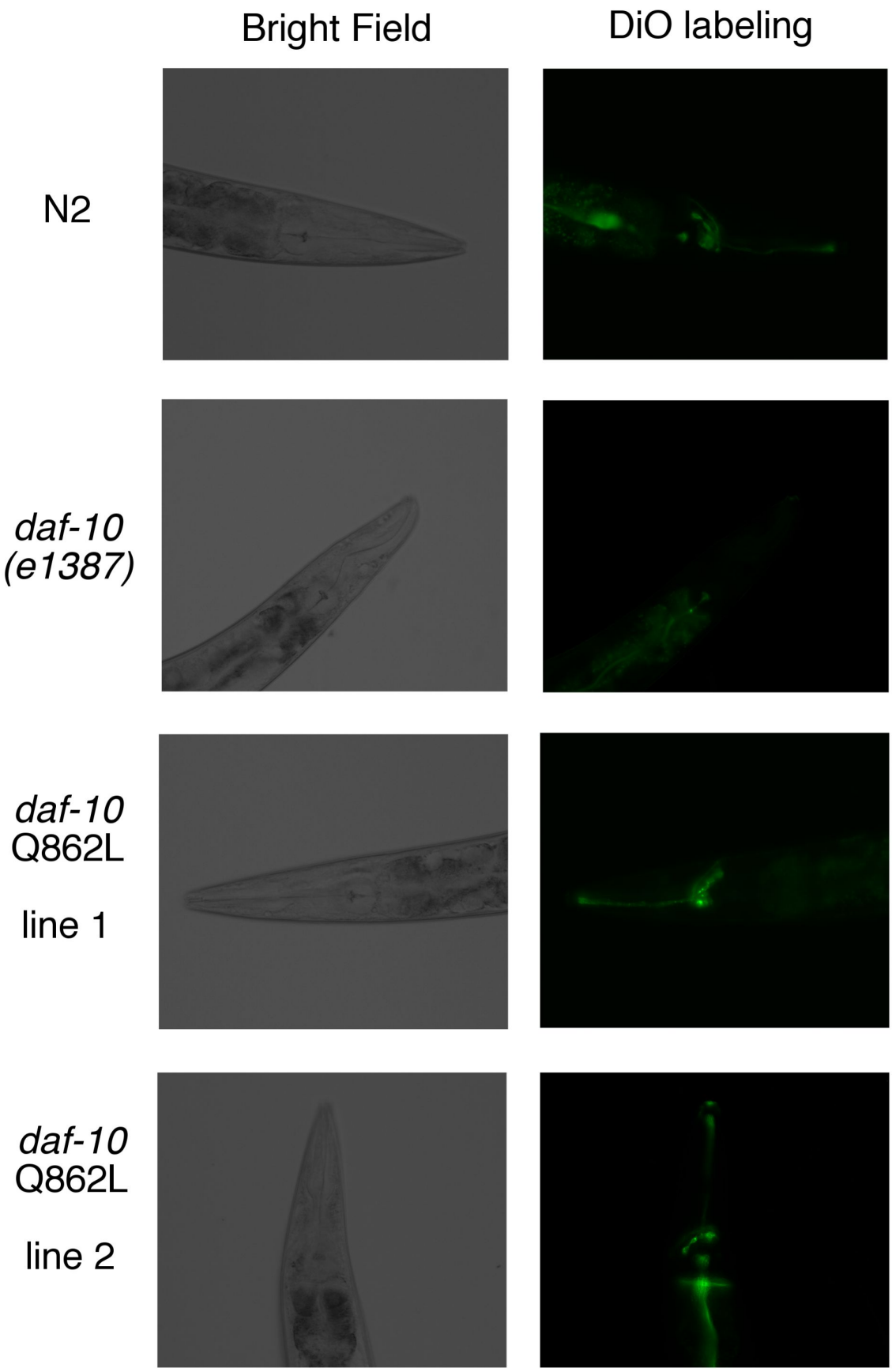

Fig. S7. Ciliated neurons of *daf-10* Q862L are able to uptake DiO dye.
The ciliated neurons of N2 control, *daf-10(e1387)*, and knock-in *daf-10* Q862L worms were stained with DiO fluorescent dye. As described previously in the literature, *daf-10(e1387)* worms do not uptake DiO, which suggest that their cilia are strongly compromised. In contrast, *daf-10* Q862L worms uptake DiO and neither the staining pattern nor the strength was qualitatively different from N2.

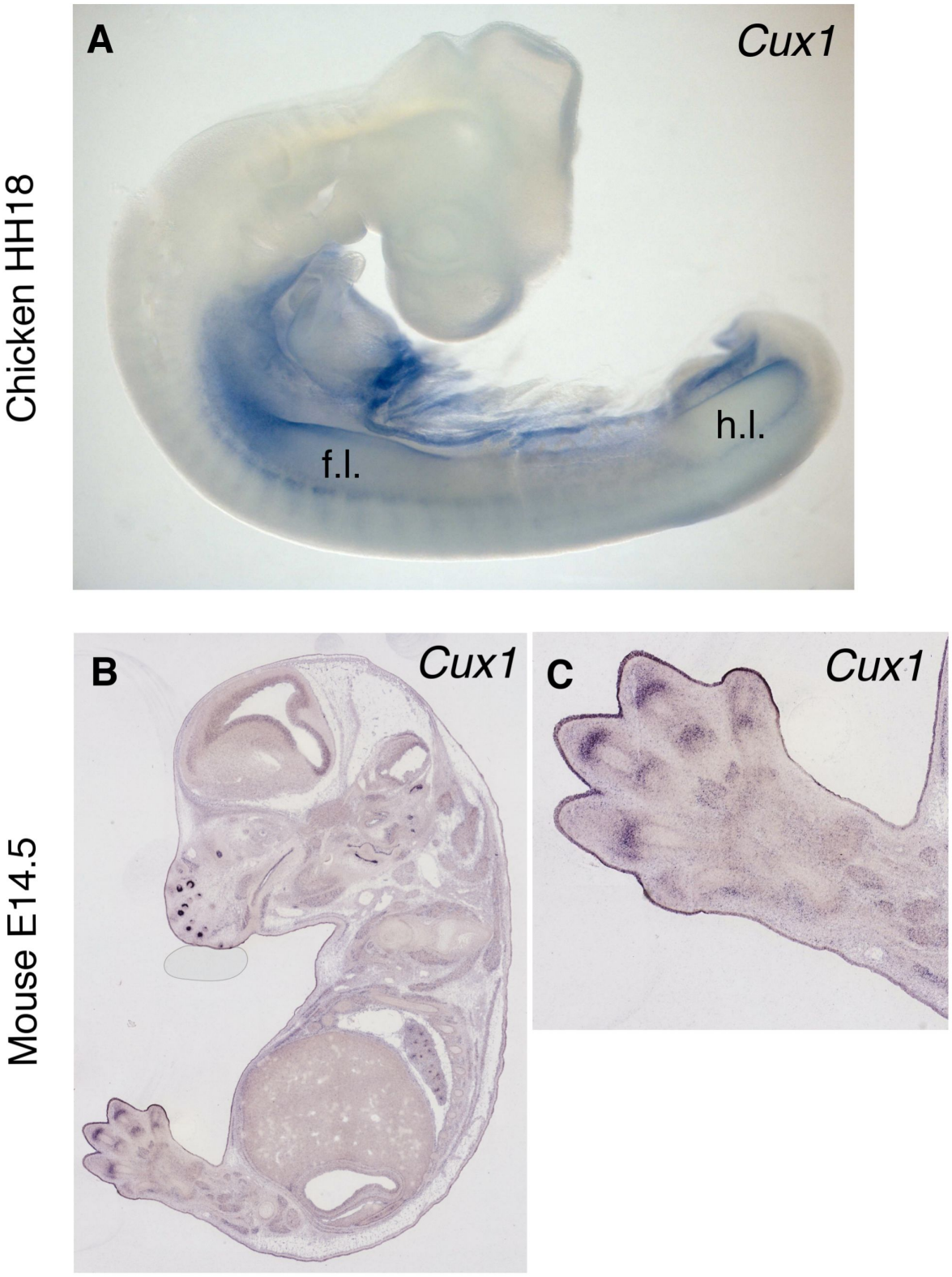

Fig. S8. Expression of Cux1 during chicken and mouse embryonic development.
(A) mRNA in situ hybridization of Cux1 in HH18 stage chicken generated and made available by the GEISHA database. Reproduced with permission (85). Cux1 is expressed in forelimb (f.l.) and hindlimb ectoderm (h.l.). **(B)** mRNA in situ hybridization of Cux1 in E14.5 mice generated and made available by the Eurexpress digital transcriptome atlas (86). Cux1 expression is strongest in developing vibrissae and limbs. (**C)** Magnification of the mouse hindlimb. Cux1 is observed in chondrocytes that will later become joints. A similar expression pattern was also observed in the chicken limb by Lizagarra and colleagues (87).

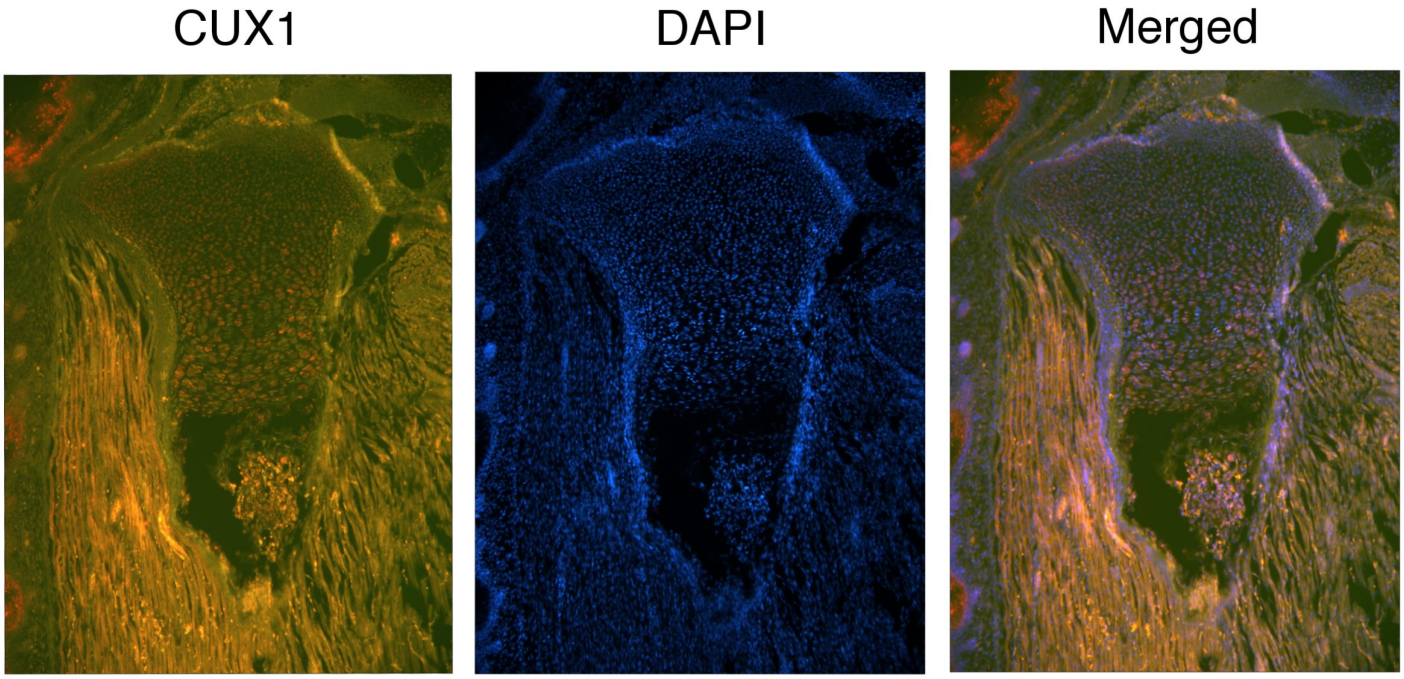

Fig. S9. Immunolocalization of CUX1 in developing bone.
Expression of CUX1 in the chondrocytes of the developing tibia of a mouse embryo (E18.5) (left). Nuclear staining (DAPI) (middle). Merged imaged (right). A rabbit polyclonal anti-mouse M-222 CUX1 antibody was used for the immunohistochemistry (Santa Cruz Biotechnology). CUX1 expression is observed in developing chondrocytes including those in the growth plate.

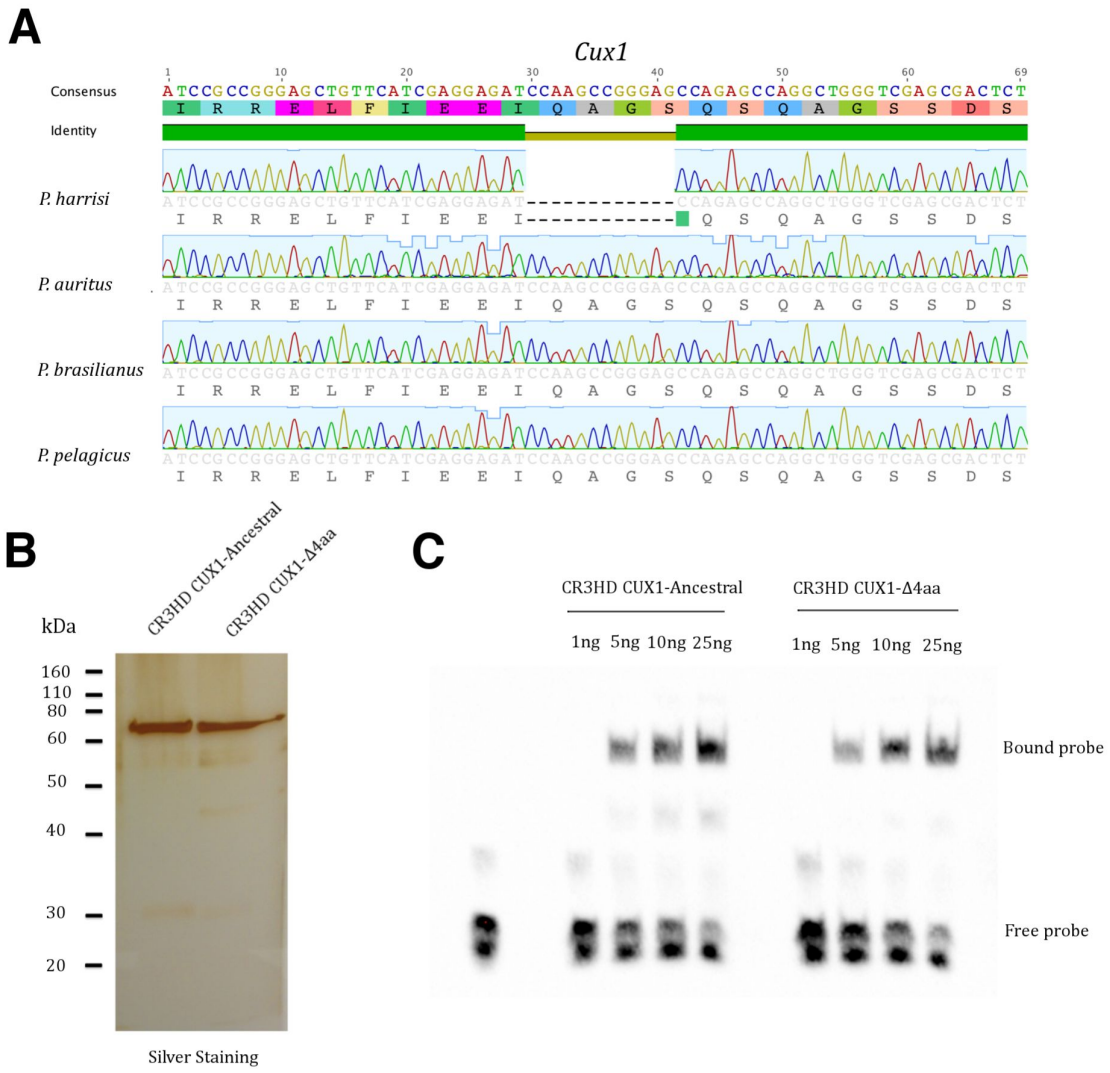

Fig. S10. *Cux1* deletion variant confirmation and CUX1 DNA binding experiments.
(A) We Sanger sequenced the area surrounding the predicted Cux1 deletion in P. harrisi, P. auritus, P. brasilianus and P. pelagicus. Alignment of the resulting sequences confirmed the 12bp deletion in P. harrisi. The deletion was not detected in the other cormorants. **(B)** Silver staining of the purified CR3HD CUX1-Ancestral and CUX1-Δ4aa variants. The proteins matched the expected sizes for MBP-CR3HD CUX1-5xHisTag protein fusions. **(C)** Electrophoretic Mobility Shift Assay (EMSA). Purified CR3HD CUX1-Ancestral and CUX1-Δ4aa variants were equally able to shift a labeled probed containing a CUX1 CR3HD binding motif.