Genomic, biochemical and expressional properties reveal strong conservation of the *CLCA2* gene in birds and mammals

Florian Bartenschlager[1], Nikolai Klymiuk[2,3], Achim D. Gruber[1] and Lars Mundhenk[1]

[1] Department of Veterinary Pathology, Faculty of Veterinary Medicine, Freie Universität Berlin, Berlin, Germany
[2] Large Animal Models in Cardiovascular Research, Internal Medical Department I, Technische Universität München, Munich, Germany
[3] Center for Innovative Medical Models, Ludwig-Maximilians-Universität München, Munich, Germany

ABSTRACT

Recent studies have revealed the dynamic and complex evolution of *CLCA1* gene homologues in and between mammals and birds with a particularly high diversity in mammals. In contrast, *CLCA2* has only been found as a single copy gene in mammals, to date. Furthermore, *CLCA2* has only been investigated in few mammalian species but not in birds. Here, we established core genomic, protein biochemical and expressional properties of *CLCA2* in several bird species and compared them with mammalian *CLCA2*. Chicken, turkey, quail and ostrich *CLCA2* were compared to their mammalian orthologues using *in silico*, biochemical and expressional analyses. *CLCA2* was found highly conserved not only at the level of genomic and exon architecture but also in terms of the canonical CLCA2 protein domain organization. The putatively prototypical galline *CLCA2 (gCLCA2)* was cloned and immunoblotting as well as immunofluorescence analyses of heterologously expressed *gCLCA2* revealed protein cleavage, glycosylation patterns and anchoring in the plasma membrane similar to those of most mammalian CLCA2 orthologues. Immunohistochemistry found highly conserved CLCA2 expression in epidermal keratinocytes in all birds and mammals investigated. Our results suggest a highly conserved and likely evolutionarily indispensable role of CLCA2 in keratinocyte function. Its high degree of conservation on the genomic, biochemical and expressional levels stands in contrast to the dynamic structural complexities and proposed functional diversifications between mammalian and avian *CLCA1* homologues, insinuating a significant degree of negative selection of *CLCA2* orthologues among birds and mammals. Finally, and again in contrast to *CLCA1*, the high conservation of *CLCA2* makes it a strong candidate for studying basic properties of the functionally still widely unresolved *CLCA* gene family.

Subjects Biochemistry, Evolutionary Studies, Genetics, Genomics, Molecular Biology
Keywords Evolution , CLCA, Avian, Mammal, Chicken, Ostrich, Turkey, Quail, Keratinocyte, Skin

INTRODUCTION

Chloride channel regulators, calcium activated (*CLCA*) constitute a family of genes that has been correlated to various disease conditions, including chronic inflammatory airway

diseases (*Hoshino et al., 2002*; *Nakanishi et al., 2001*; *Patel, Brett & Holtzman, 2009*; *Range, Mundhenk & Gruber, 2007*), cystic fibrosis (*Hauber et al., 2003*; *Ritzka et al., 2004*; *Young et al., 2007*) and cancer (*Chen et al., 2019*; *Hou et al., 2017*; *Walia et al., 2009*; *Walia et al., 2012*; *Yu, Walia & Elble, 2013*) and shows striking evolutionary dynamics (*Bartenschlager et al., 2022*; *Mundhenk et al., 2018*). In general, CLCAs comprise a prototypical protein domain architecture of a CLCA N-terminal (N-CLCA), a von Willebrand factor type A (vWA), a beta sheet rich (bsr) as well as a carboxy (C)-terminal fibronectin type III domain (fn3) that is separated from N-CLCA, vWA and bsr by proteolytical cleavage (*Patel, Brett & Holtzman, 2009*). Some CLCAs contain a transmembrane (TM) domain that anchors the C-terminal cleavage product in the plasma membrane (*Bartenschlager et al., 2022*; *Braun et al., 2010*; *Elble et al., 2006*; *Patel, Brett & Holtzman, 2009*; *Plog et al., 2012a*). CLCAs lacking the TM appear completely secreted (*Gibson et al., 2005*; *Mundhenk et al., 2006*; *Patel, Brett & Holtzman, 2009*; *Plog et al., 2009*).

On the genomic level, all *CLCA* genes in a given species are located within in a single locus, which is consistently flanked by the *Outer dense fibre of sperm tails 2-like* (*ODF2L*) and *SH3-domain GRB2-like endophilin B1* (*SH3GLB1*) genes (*Bartenschlager et al., 2022*; *Mundhenk et al., 2018*; *Plog et al., 2009*). Although *CLCA* genes have been identified in numerous avian, reptile, amphibian or fish species in the recent past, the main research has been focused on *CLCA* homologues in mammalian species so far (*Cunningham et al., 2021*). From the phylogenetic perspective, mammalian *CLCA* genes have been assigned to four clearly distinct clusters (*Mundhenk et al., 2018*; *Plog et al., 2009*). While clusters 3 and 4 exhibit a complex arrangement of multiple genes that seemingly arose from several independent duplication and inactivation events during the emergence of mammalian species (*Bartenschlager et al., 2022*; *Mundhenk et al., 2018*), both *CLCA* clusters 1 and 2 comprise only one intact single gene in each mammalian species. Recently, we described a *CLCA* locus in the genome of chicken (*Gallus gallus*) which is similarly flanked by the galline *ODF2L* and the *SH3GLB1* genes (*Bartenschlager et al., 2022*; *Mundhenk et al., 2018*). In contrast to the four mammalian *CLCA* (ma*CLCA*) clusters, however, chickens possess only two *CLCA* homologous genes, the galline CLCA1 (*gCLCA1*) and *gCLCA2*. From an evolutionary perspective, *gCLCA1* groups with mammalian clusters *CLCA1/3/4* and might therefore be the avian representative of the monophyletic ancestor of this group (*Bartenschlager et al., 2022*). In contrast to the genetically and functionally divergent *CLCA1/3/4* cluster, g*CLCA2* seems to be much more conserved with g*CLCA2* having a close genetic relationship to ma*CLCA2* (*Bartenschlager et al., 2022*; *Mundhenk et al., 2018*).

Concordantly with the high degree of conservation, maCLCA2 is consistently expressed in keratinocytes of stratified epithelia such as those of skin or esophagus and in certain glandular epithelia in all mammals investigated so far (*Braun et al., 2010*; *Connon et al., 2006*; *Connon et al., 2004*; *Erickson, Gruber & Mundhenk, 2020*; *Hamalainen et al., 2021*; *Plog et al., 2012b*; *Seltmann et al., 2018*; *Walia et al., 2012*). It has been proposed that the CLCA2 protein is involved in epithelial differentiation, growth arrest and maturation of keratinocytes (*Connon et al., 2004*; *Koegel & Alzheimer, 2001*; *Ramena et al., 2016*). Some species-specific differences regarding the expression, however, have been revealed in the mammalian respiratory tract, where human, porcine, and murine, but no feline CLCA2

has been detected in airway submucosal glands (*Dietert et al., 2015*; *Erickson et al., 2018*). Further, additional protein expression has been found in a specific subset of respiratory epithelial cells of the bronchial bifurcation in mice (*Dietert et al., 2015*).

Aiming at clarifying the relevance of non-mammalian *CLCA2*, we characterized the genomic organization and tissue und cellular expression patterns of *CLCA2* in four bird species, including chicken (*Gallus gallus*), turkey (*Meleagris gallopavo*), quail (*Coturnix species (sp.)*), and ostrich (*Struthio camelus*). The genomic organization as well as the protein domain architecture are described for the galline, quail and ostrich orthologues. We further describe the biochemical properties of CLCA2 in chicken as a presumed avian prototype. By comparison with its porcine, feline and murine CLCA2 homologues, conserved architectural elements, specific traits of biochemical processing and tissue expression patterns were identified, which will serve as the basis for functional investigations and structure-function-correlations in the future.

MATERIAL AND METHODS

In silico sequence analysis of gCLCA2 and generation of antibodies

Detailed gene positions, sizes, gene and amino acid (aa) sequences from chicken, quail, ostrich, pig, cat, and mouse *CLCA* loci were extracted from the NCBI (https://www.ncbi.nlm.nih.gov/) and Ensembl (http://www.ensembl.org/index.html) databases as described by *Plog et al. (2009)*. NCBI or Ensembl identifiers for CLCA2 sequences used are listed in Fig. S1. Exon-intron boundaries were established using WebScipio (*Hatje, Hammesfahr & Kollmar, 2013*) and aligned by GenePainter (*Hammesfahr et al., 2013*). Predicted protein domains were identified by the NCBI Conserved Domain Database (*Lu et al., 2020*), EMBL-EBI HMMER web server (*Potter et al., 2018*), Phobius webserver (*Käll, Krogh & Sonnhammer, 2007*), SOSUI (*Hirokawa, Boon-Chieng & Mitaku, 1998*), SignalP 3.0 (*Bendtsen et al., 2004*) algorithms and manual alignments. Asparagine (N)-linked glycosylation sites were predicted using the NetNGlyc webserver 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/). Turkey CLCA2 was not incorporated in the *in silico* analysis due to the low quality of the full-length gene and aa sequences stored in the NCBI and Ensemble databases (XP_031410715.1, XM_031554855.1, ENSMGAT00000009704.2). Phylogenetic relationship based on protein sequences of galline, quail, ostrich, feline, porcine and murine CLCA2 sequences was inferred by using the Maximum Likelihood method and JTT matrix-based model conducted in the MEGA X software package with 100 bootstrap replicates (*Tamura, Stecher & Kumar, 2021*) (S6).

Anti-gCLCA2 antibodies were generated similar to anti-porcine CLCA1 antibodies (*Plog et al., 2009*). In brief, an oligopeptide corresponding to aa 875 to 888 (WTAPGDDFD-KGQAA) in the C-terminal region of gCLCA2 was synthesized and conjugated with *Limulus polyphemus* hemocyanin (LPH). The LPH- conjugated peptide was used for immunization of two rabbits. Specific IgG-antibodies were isolated from the antisera using a cyanogen bromide immunization-peptide coupled sepharose column and named gC2.

Animals and tissues

In accordance with the 3R principle for the reduction of animal experiments, all tissues used in this study were obtained from the veterinary diagnostic pathology tissue archive of the Department of Veterinary Pathology, Freie Universität Berlin, Germany (VetPathFU) and originated either from the veterinary clinical diagnostic service unit or previous experimental studies. No animal was bred, raised, kept or euthanized specifically for this study. For chickens, 45 freshly frozen or formalin fixed and paraffin embedded (FFPE) tissues (Table S2) from ten-week old female individuals (*Gallus gallus domesticus*, Hampshire x White Leghorn, $n = 3$) and the gonads of age-matched male chickens (Hampshire x White Leghorn, $n = 3$) were used from (*Bartenschlager et al., 2022*). In brief, the tissues were by-products from slaughtered animals intended for human consumption. The animals had been bred, housed, and slaughtered in the Albrecht Daniel Thaer-Institute of Agricultural and Horticultural Sciences of the Humboldt-Universität zu Berlin, Germany under the permission of the State Office of Health and Social Affairs (approval number IC 114-ZH70). Weight at harvest was 1–1.2 kg (females) and 1.3–1.5 kg (males). The animals were raised in groups of 25, with infrared heat lamps offered until week five. They were fed with fledgling rearing feed until week eight and young hen feed afterwards. Miscanthus litter was used for housing enrichment. Harvesting was conducted according to national guidelines, which includes anesthesia by head blow and rapid exsanguination *via* jugular veins and carotid arteries. Females were harvested in the morning and males in the morning of the following day. FFPE tissues for immunohistochemical analyses, including esophagus and skin from shanks, abdomen and foot from one female and two male ostriches (*Struthio camelus*), three male quails (*Coturnix sp.*), and skin from three turkeys (*Meleagris gallopavo*) and skin of cats (*Felis catus*) were provided by the veterinary clinical postmortem diagnostic service unit of VetPathFU with no association to animal experiments. The tissues were obtained from the routine diagnostic spectrum to determine the cause of death of animals kept by private owners and free of histopathological changes. The permission to further use these tissues for research purposes was given by signature on the necropsy submission form by the owners. Additionally, FFPE skin samples from each of three mice and pigs from previous experimental studies (*Braun et al., 2010*; *Plog et al., 2012a* approval numbers T 0104/06 and G 0323/06, respectively) were obtained from the archive of VetPathFU and used for immunohistochemical analyses. In brief, 10-weeks-old female C57BL/6J mice were kept in cages enriched with nesting material. All animals had unlimited access to standard pelleted food and tap water. The room temperature was at $22 \pm 2 \, °C$ and the relative humidity at 45–65%. A 12-h light/dark cycle was maintained. For experimental procedures, all mice were sacrificed by cervical dislocation in accordance with the national guidelines. Furthermore, the piglets (Euroc × Piétrain) were 18 days old, male, castrated and kept for four weeks in flatdeck compartments in groups of six piglets enriched with playthings. All animals had unlimited access to mash food and tap water. The room temperature was 28 °C at stabling and was gradually decreased to 22 °C within 10 days with air humidity at approx. 65%. The light programme consisted of a 16 h light and 8 h dark phase. With 45 days of age, all piglets were anesthesized with ketamine hydrochloride (Ursotamin®, 10%; Serumwerk Bernburg AG, Germany) and azaperone

(Stresnil®, Jansen-Cilag, Neuss, Germany) and euthanized using tetracaine hydrochloride, mebezonium iodide and embutramide (T61®, Intervet, Germany). All efforts were made to minimize animal discomfort and suffering.

Molecular cloning and sequencing of *gCLCA2*

The *gCLCA2* open reading frame (ORF) was cloned as described with minor modifications (*Bartenschlager et al., 2022*). In brief, the *gCLCA2* ORF was amplified from a batch of tissues including pharynx, crop, proctodeum and footpad from animal #2 (Table S2). The gCLCA2 ORF was tagged with the enhanced yellow fluorescent protein (*EYFP*) at the C-terminus by cloning it into the p*EYFP*-N1 vector (Clontech, Mountain View, California, USA). The resulting plasmid (*gCLCA2#2*) was sequenced using the primer walking method (Data S5). Three plasmids from independent experiments yielded identical results.

RT-qPCR tissue localization of *gCLCA2* mRNA

mRNA expression was analyzed using RT-qPCR as described (*Bartenschlager et al., 2022*). In brief, total RNA was isolated from galline tissues (Table S2), reverse transcribed, and the cDNA diluted to a final concentration of 1 ng/µl. Specific exon 13/14-boundaries spanning primers (upstream: 5′-CCAGGCTAACAGGACTACC-3′; downstream: 5′-GAAACCTCCTCTTCTGACCTGAAC-3′) were used to detect *gCLCA2* or the reference gene phosphoglycerate kinase (*PGK1*, upstream: 5′-AAAGTTCAGGATAAGATCCAGCTG-3′; downstream 5′-GCCATCAGGTCCTTGACAAT-3′; *Olias et al., 2014*) using a SYBR green qPCR assay (Thermo Fisher Scientific, Waltham, MA, USA). The *gCLCA2*-PCR product corresponds to the gCLCA2 protein region from aa760 to aa809 (QANRTTVPQTAMPWSHAMYIPGYVENGKLKMNPSRPPAI-ENNVQVRRGGF). *gCLCA2* mRNA was considered to be expressed when C_t-values of 35 or less were detected in at least two out of three tested animals.

Transient transfection of HEK293 cells

HEK293 cells (ATCC, Manassas, Virginia, USA) were transiently transfected as described with minor modifications (*Bartenschlager et al., 2022*). In brief, cells were grown in six-well plates in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% heat-inactivated fetal calf serum (FCS), 1% HEPES, and 1% penicillin/streptomycin. When reaching 80–90% confluence, the cells were transfected with 2 µg of a plasmid containing gCLCA2#2 or EYFP alone (mock) using 8 µl polyethylenimine (PEI) per well. 12 h post transfection, the cells were washed with phosphate buffered saline (PBS) and serum-free DMEM was added. 48 h after transfection, the cells of each well were lyzed using 500 µl radioimmunoprecipitation assay (RIPA) buffer supplemented with a protease inhibitor cocktail (complete Mini, EDTA-free, Roche Diagnostics, Rotkreuz, Switzerland). The protein concentration of supernatants and cell lysates were quantified using the bicinchoninic acid (Thermo Fisher Scientific, Waltham, Massachusetts, USA) method prior to freezing at −20 °C.

Endoglycosidase treatment

For glycosylation analysis, lysates from *gCLCA2*-transfected cells were deglycosylated by incubation with 25 U/ml endo H, 50 U/ml PNGase F or left untreated at 37 °C over night

according to the manufacturer's protocols (New England Biolabs, Ipswich, Massachusetts, USA).

Immunoblotting

Cell lysates and supernatants of *gCLCA2* transfected cells were analyzed using immunoblotting as described with minor modifications (*Bartenschlager et al., 2022*). In brief, samples of cell lysates or concentrated cell culture supernatant were reduced in 1,4-dithiothreitol (DTT) and separated using a 10% SDS-polyacrylamide gel electrophoresis. Proteins were transferred to a (PVDF)-membrane (https://www.linguee.de/englisch-deutsch/uebersetzung/polyvinylidene+fluoride.html) and blocked with 5% non-fat milk. Membranes were probed with antibody gC2 in a three-fold dilution series from 5 µg/ml to 0.05 µg/ml, or mouse monoclonal anti-YFP (cat. G163; ABM, Vancouver, Canada) diluted at 1:500, or mouse monoclonal anti-beta-actin (A5441, Sigma-Aldrich, St. Louis, Missouri, USA) diluted at 1:1,000. Membranes were incubated with horseradish peroxidase-conjugated goat anti-rabbit (115-035-068, Jackson Immuno Research Laboratories, Inc., West Grove, Pennsylvania, USA) or goat anti-mouse (111-035-144, Jackson Immuno Research Laboratories, Inc.) secondary antibodies and developed using enhanced chemiluminescence (Supersignal West Pico Plus, Thermo Fisher Scientific, Waltham, MA, USA). The gCLCA2 protein was only detected by this technique when using the anti-YFP antibody; however, it was undetectable when the gC2 antibodies were used.

Immunocytochemistry of transfected HEK293 cells

Immunocytochemistry was performed as described with minor modifications (*Bartenschlager et al., 2022*). In brief, HEK293 (ATCC) cells were grown on 8-well tissue chamber slides and transfected with *gCLCA2#2* or *EYFP-* mock plasmids. 48 h after transfection, the cells were briefly fixed in ice-cold methanol followed by a 4% paraformaldehyde fixation for 10 min. After permeabilization with 0.1% Triton X-100 in PBS and blocking with 10% goat normal serum (GS) and 0.05% Tween 20 in PBS, cells were probed with untreated or pre-absorbed antibody gC2 each used at 2 µg/ml or irrelevant affinity-purified rabbit polyclonal anti-porcine CFTR antibody (*Plog et al., 2010*) (S3). Alexa fluor 568 conjugated goat anti-rabbit (AB_143157, Invitrogen, Carlsbad, California, USA) were used as secondary antibodies followed by 4′, 6-diamidino-2-phenylindole (DAPI) nuclear counterstain. All *in vitro* experiments were repeated three times.

Tissue and cellular localization of gCLCA2 protein using immunohistochemistry and immunofluorescence

All galline tissues in which g*CLCA2* mRNA was detected at C_t-values below 35 were analyzed *via* immunofluorescence to identify gCLCA2 expressing cell types. Furthermore, skin and esophagus from chicken, ostriches, quails, as well as skin of turkeys, mice, pigs, and cats were analyzed *via* immunohistochemistry. Immunofluorescence and immunohistochemistry were performed as described with minor modifications (*Bartenschlager et al., 2022*). In brief, FFPE-tissues were cut, mounted on adhesive glass slides, and dewaxed. For immunohistochemistry, endogenous peroxidase was blocked by adding 0.5% H_2O_2 in methanol. For immunofluorescence analysis, tissue sections of chickens were permeabilized

with 0.1% Triton X-100 in PBS. Antigen was retrieved using 1 mg/ml recombinant protease from *Streptomyces griseus*. Slides were blocked with 10% Roti-ImmunoBlock and 20% GS in PBS for immunohistochemistry and 10% GS and 0.05% Tween 20 in PBS for immunofluorescence, both for 30 min. The slides were probed with the immunopurified gC2 or irrelevant affinity-purified rabbit polyclonal (anti-porcine CFTR, *Plog et al., 2010*) antibodies at 2 μg/ml. Additionally, mouse monoclonal anti-cytokeratin (AE1/AE3, M3515, Agilent Dako, Santa Clara, California, USA) antibodies were used at 1:400. For immunofluorescence, Alexa fluor 568-conjugated goat anti-rabbit (AB_143157; Invitrogen, Waltham, MA, USA) secondary antibodies were used diluted at 1:200, followed by DAPI nuclear counterstain. For immunohistochemistry with AE1/AE3 primary antibodies, 3,3'-diaminobenzidine (DAB) was added after incubation with goat anti-mouse biotinylated secondary antibodies (BA-9200; Vector Laboratories, Burlingame, California, USA) diluted at 1: 200 and an avidin-biotin complex. For immunohistochemistry with the gC2 primary antibody, DAB was added after incubation with the Histofine Simple Stain Mouse MAX PO anti-rabbit polymer kit (414341F; Nichirei Biosciences Inc., Tokyo, Japan). Potential cross reactivity of the gC2 antibody with pig, cat, mouse, turkey, quail and ostrich CLCA2 orthologues was tested by epitope sequence alignment using the NCBI Protein Blast (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins, Table S4).

RESULTS

Avian and mammalian *CLCA2* genes and their overall protein structures are conserved

Similar to mammals (*Bartenschlager et al., 2022*; *Mundhenk et al., 2018*), avian *CLCA2* (a*CLCA2*) are single-copy genes located directly adjacent to the *ODF2L* gene in all species analyzed here (Fig. 1A). While in mammals the region between *CLCA2* and *SH3GLB1* comprise the complex and divergent *CLCA1/3/4* locus, birds contain only one single *CLCA1* gene (Fig. 1A). The presence of only two *CLCA* homologues but also the shorter intronic and intergenic regions that correspond to the more compact organization of avian genomes (*Ellegren, 2005*) make the consensus avian *CLCA* gene locus much smaller than that in mammals (Fig. 1A, *Bartenschlager et al., 2022*). The a*CLCA2* genes comprise 14 exons that encode for a single, putatively functional protein in all bird species analyzed here (Figs. 1B, 2), identical to all mammals investigated to date (*Patel, Brett & Holtzman, 2009*). The nucleotide numbers vary in only four exons by three to nine nucleotides between avian and mammalian *CLCA2* genes (Fig. 1B), causing only slightly distinct protein lengths (Fig. 2). Overall, our findings establish a high degree of evolutionary conservation of the genomic *CLCA2* architecture across birds and mammals (Fig. 1B). For all *CLCA2* orthologous genes investigated here, the predicted proteins share the canonical CLCA protein domain architecture as described (*Patel, Brett & Holtzman, 2009*). *In silico* prediction suggests a signal peptide within the first 22 to 43 aa, indicating entry into the secretory pathway which is a highly conserved trait in all avian and mammalian CLCA2 sequences (Fig. 2). An N-CLCA domain (PFAM identifier: pfam08434) containing a proteolytic HExxH-motif and a cysteine-rich domain is prepended to a vWA domain

(PFAM identifier: pfam13519) and a bsr domain (Fig. 2, S1, *Patel, Brett & Holtzman, 2009*). In contrast to other CLCA proteins but consistent with maCLCA2, the vWA domain of CLCA2 of birds does not contain an intact metal ion dependent adhesion site (MIDAS, DxSxS-T4-D5). In concordance to maCLCA2, an fn3 (PFAM identifier: PF00041) and a TM domain are predicted in the C-terminal cleavage product for aCLCA2 (Fig. 2, Fig. S1). The presence of designated beta4-integrin binding motif (IBM, consensus sequence F(S/N)R(I/L/V)(S/T)S, *Abdel-Ghany et al., 2003*) appears less consistent. Human and porcine CLCA2 show such an IBM within the vWA domain (human aa480-485: FSRISS (*Patel, Brett & Holtzman, 2009*), pig aa480-485: FSRISS, Fig. S1) while it is absent from chicken, quail, ostrich, feline and murine CLCA2 (Fig. S1). Another IBM motif is lacking in the C-terminal cleavage product of birds and pigs whereas such a motif was found in human, feline and murine CLCA2 (human aa480-485: FSRISS (*Patel, Brett & Holtzman, 2009*), cat aa741-746: FSRVSS, Fig. S1, mouse aa740-745: FSRVSS (*Patel, Brett & Holtzman, 2009*), Fig. S1).

Noteworthy, aCLCA2 amino acid (aa) sequences are 921 to 930 aa long and therefore shorter than all of their mammalian homologues with 942 to 944 aa (Fig. 2). In addition, aCLCA2 do not contain a predicted glycosylation site after the predicted transmembrane domain (Fig. 2).

A phylogeny based on galline, quail, ostrich, feline, porcine and murine CLCA2 protein sequences revealed a monophyletic aCLCA2 group separate from maCLCA2 (Fig. S6).

The gCLCA2 protein shares many biochemical properties with mammalian CLCA2 proteins
Posttranslational cleavage
The cleavage of a precursor protein into a larger N- and a shorter C-terminal subunit belongs to the conserved properties of all maCLCA proteins (*Patel, Brett & Holtzman, 2009*). It is thought to be mediated by the zinc-binding HExxH motif in the N-CLCA domain, cleaving the protein at a canonical cleavage site (*Bothe et al., 2012*; *Lenart et al., 2013*; *Pawłowski et al., 2006*; *Yurtsever et al., 2012*). Consistently, canonical HExxH motifs and putative proteolytic cleavage sites were identified in aCLCA2 proteins (Fig. 2). The predicted cleavage was verified *in vitro* for gCLCA2 as a putative avian prototype by immunoblot analysis of lysates from heterologously transfected HEK293 cells. A band consistent with a precursor protein at approx. 145 kilodalton (kDa) and a band consistent with the C-terminal cleavage product of approx. 57 kDa were detected (Fig. 3), suggesting that posttranslational cleavage of CLCA2 also occurs in chicken.

N-glycosylation and cleavage in the medial Golgi
maCLCA2 proteins are multiple N-linked glycosylated (*Braun et al., 2010*; *Elble et al., 2006*; *Gruber et al., 1999*; *Plog et al., 2012b*). Consistently, our *in silico* analyses predicted five N-linked glycosylation sites for aCLCA2 (Fig. 2). To verify this prediction experimentally, lysates from *gCLCA2*- transfected HEK293 cells were treated with endoglycosidases endo H and PNGase F and immunoblotted to identify the kind and extent of glycosylation. The approx. 145 kDa precursor protein was sensitive to endo H and PNGase F treatments (Fig. 4), resulting in a size shift from approx. 145 kDa to approx. 130 kDa, suggestive
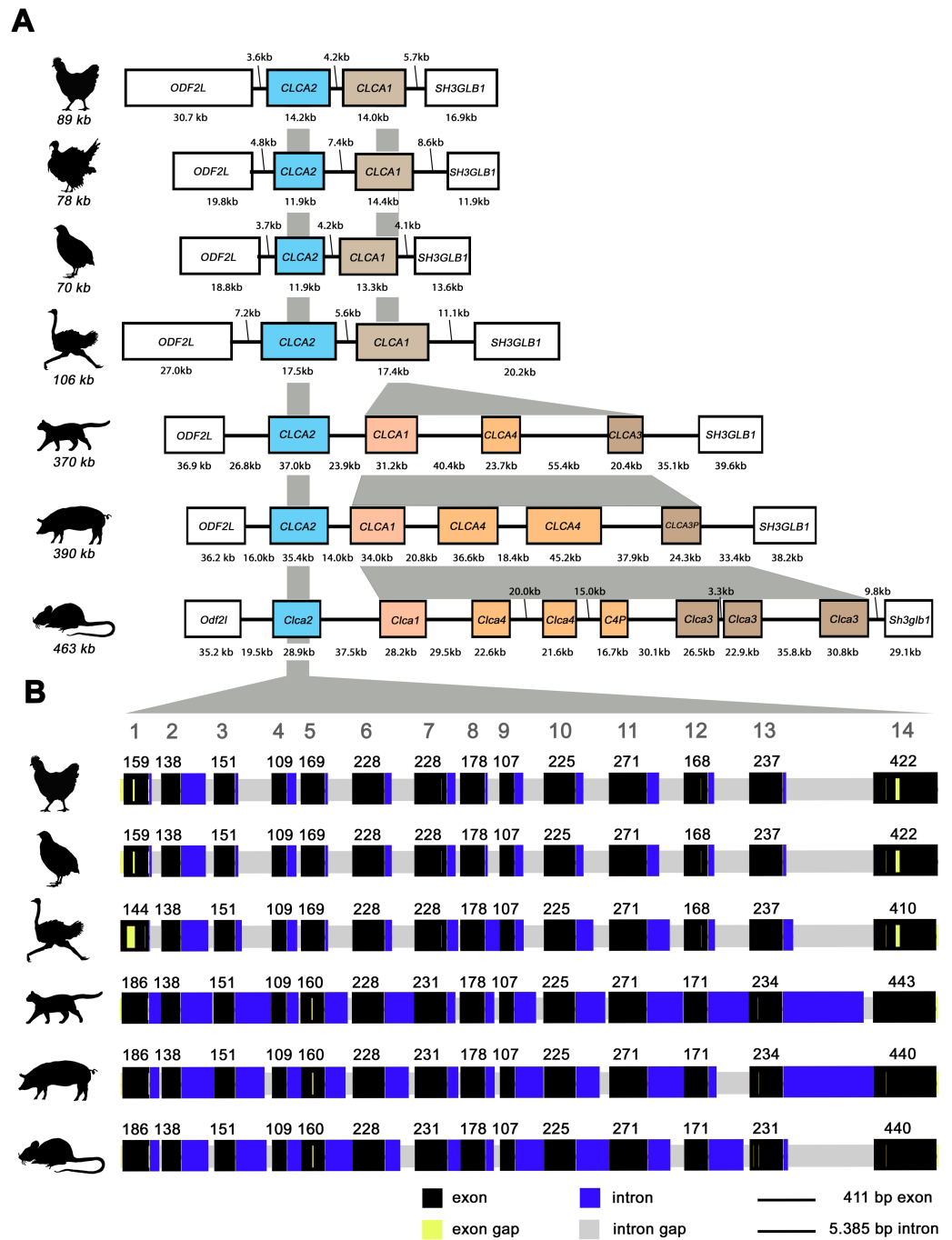
Figure 1 **The genomic architecture of avian and mammalian *CLCA2* genes is highly conserved while *CLCA1* homologues are diverse only in mammals but not in birds.** (A) Boxes indicate genes, black lines illustrate intergenic regions. Box colors highlight *CLCA* orthologues with white boxes indicating neighboring genes. Gray areas indicate regions of high sequence homology. The avian *CLCA* loci are scaled 2-fold larger for illustrative purposes. kb, kilobases.

Full-size ⊡ DOI: 10.7717/peerj.14202/fig-1

Figure 1 (...continued)
(B) The exon architecture of avian and mammalian *CLCA2* is highly conserved. Identical to mammals, a*CLCA2* genes are encoded by 14 exons with concurring exon lengths in exons 2, 3, 4, 6, 8, 9, 10 and 11. Black boxes represent exons and blue boxes depict introns. Gaps in the alignment are highlighted in yellow for exons and in grey for introns. Large grey digits at the top indicate exon numbers, small black digits indicate base pairs per exon. bp: base pairs.
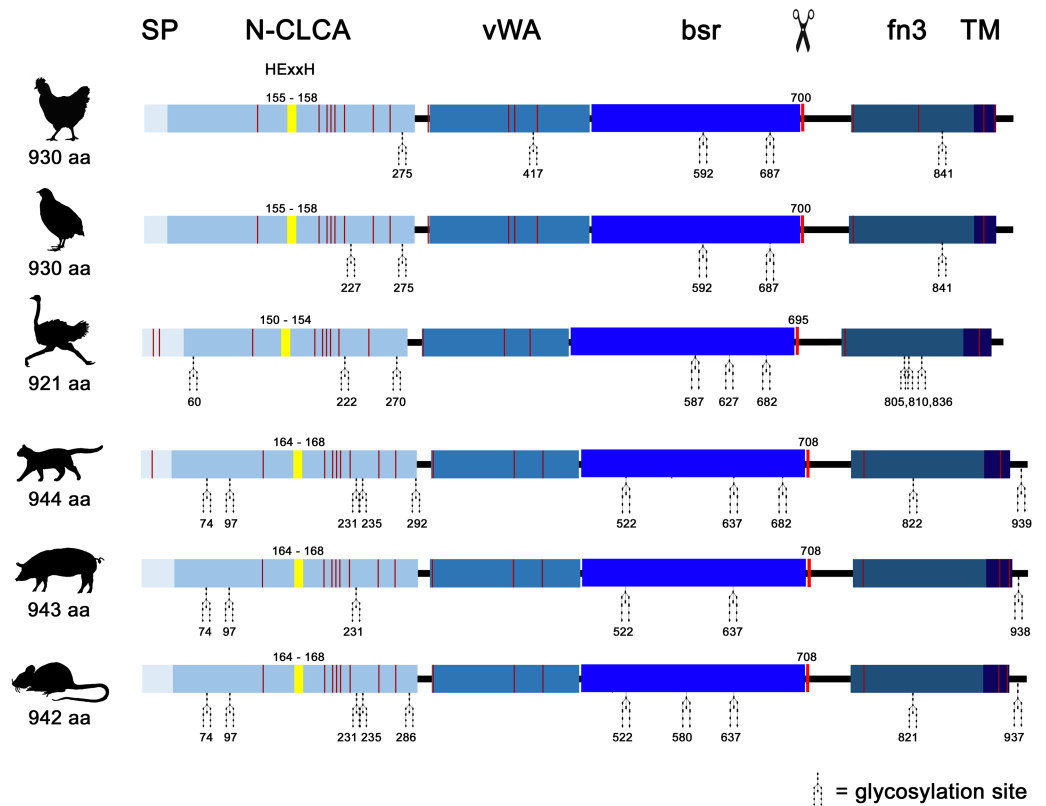
Figure 2 Avian and mammalian CLCA2 proteins share the canonical CLCA2 predicted protein architecture. Graphical alignment of CLCA2 orthologues. Protein domains are indicated by boxes in shades of blue and entitled by: SP, signal peptide; N-CLCA, N-CLCA domain; vWA, von Willebrand factor type A domain; bsr, beta sheet rich domain; fn3, fibronectin type III domain; TM, transmembrane domain. Yellow vertical boxes illustrate HExxH-motifs, red vertical boxes and the scissor sketch depict putative cleavage sites, vertical red lines constitute cysteines. Numbers indicate aa positions or lengths, respectively.
Full-size ☑ DOI: 10.7717/peerj.14202/fig-2

of an immature high mannose-type glycosylation pattern. In contrast, the C-terminal cleavage product was resistant to endo H treatment but sensitive to PNGase F treatment, as suggested by a reduction in size from approx. 57 kDa to approx. 54 kDa (Fig. 4). Our results indicate that virtually all predicted sites might be glycosylated, presuming a molecular weight of approx. 3 kDa per glycosylation (*Pult et al., 2011*). Furthermore, the complex high mannose-rich glycosylation pattern of the C-terminal cleavage products, in contrast to the immature glycosylated precursor protein, suggests cleavage of gCLCA2 early in the medial Golgi, similar to what has been observed for the murine CLCA2 (*Braun*
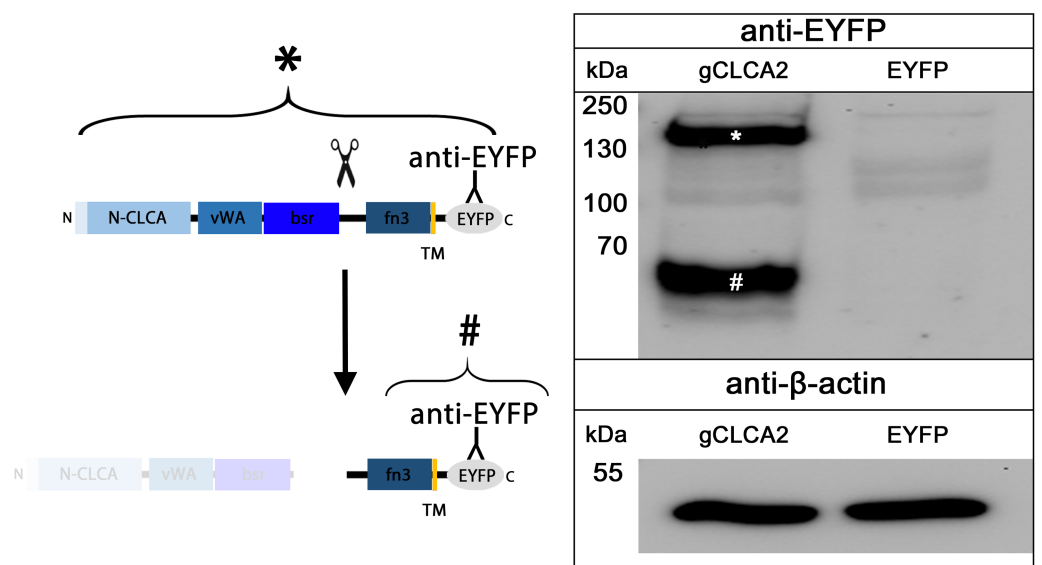
Figure 3 Posttranslational cleavage of the gCLCA2 precursor protein. Left panel: illustration of protein cleavage: N-CLCA: N-CLCA domain, vWA: von Willebrand factor type A domain, bsr, beta sheet rich domain; fn3, fibronectin type III domain; TM, transmembrane domain; EYFP, enhanced yellow fluorescent protein tag; scissor sketch, putative cleavage site; Y, antibody binding site. Right panel: immunoblot analysis of lysates from *gCLCA2* or *EYFP*-mock plasmid transfected HEK293 cells using an anti-EYFP antibody. An asterisk (*) indicates uncleaved precursor protein, a hash (#) indicates C-terminal cleavage product. To control for equal total protein loading, the samples were identically immunoblotted with primary anti-beta-actin antibodies (bottom panel). Representative image of three independent experiments are shown.
Full-size ⬛ DOI: 10.7717/peerj.14202/fig-3

et al., 2010). Thus, glycosylation and cleavage in the medial Golgi also appear as conserved traits.

Anchoring in the plasma membrane

Identically to its maCLCA2 orthologues, a TM domain in the C-terminal subunit was predicted for aCLCA2, which anchors the protein to the plasma membrane (Fig. 2; *Braun et al., 2010*; *Elble et al., 2006*; *Gruber et al., 1999*; *Plog et al., 2012b*). Consistently, a prominent green autofluorescent EYFP signal was detected along the plasma membrane of HEK293 cells transfected with the *gCLCA2* construct that contained an EYFP tag downstream to the TM domain (Figs. 5A and Figs. 5B1). A coinciding signal (Fig. 5B3) was found by immunocytochemistry using the anti-gC2 antibody directed against the fn3 domain of the protein, which is located upstream to the C-terminal transmembrane domain (Fig. 5B2). This is in contrast to the diffuse green signal of cytosolic EYFP protein (Fig. 5C). Therefore, anchorage of the CLCA2 protein in the cell membrane *via* a TM domain seems also conserved with mammals.

aCLCA2 is expressed in stratified squamous epithelia of skin and mucous membranes

The tissue and cellular expression patterns of gCLCA2 were examined at the mRNA- and protein-levels. *gCLCA2* mRNA was detected by RT-qPCR in all tested locations of the skin
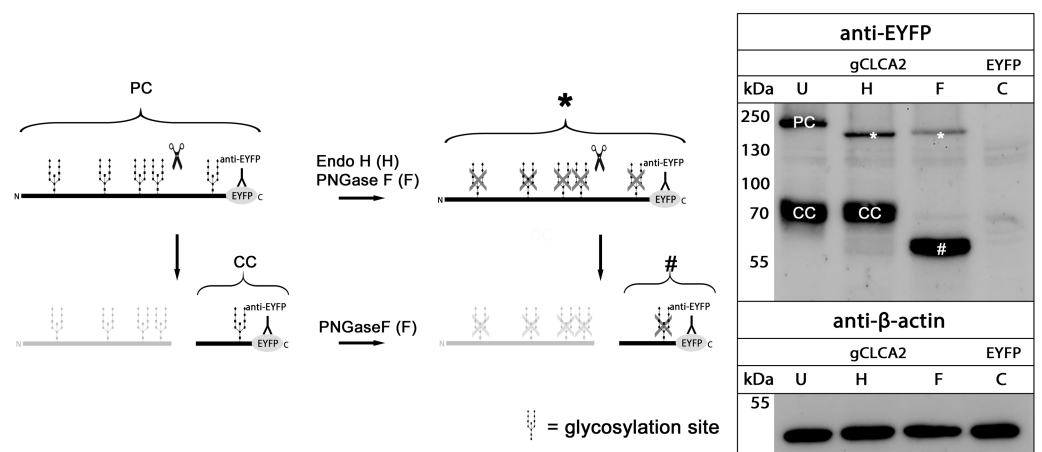
Bartenschlager et al. (2022), *PeerJ*, DOI 10.7717/peerj.14202   11/25

Figure 4 **gCLCA2 is a multiple N-linked glycosylated protein and cleaved early in the medial golgi.** Left panel: illustration of protein cleavage and deglycosylation: scissor sketch, cleavage; Y, antibody binding site; PC, precursor protein; CC, C-terminal cleavage product; asterisk, deglycosylated, uncleaved precursor protein; hash, deglycosylated C-terminal cleavage product. Right panel: immunoblot analysis of lysates from HEK293 cells transfected with *gCLCA2* or the *EYFP*-mock plasmids (C), using an anti-EYFP antibody. Lysates were treated with endoglycosidases endo H (H), PNGase F (F) or left untreated (U). To control for equal total protein loading, the samples were immunoblotted with anti-beta-actin antibodies (bottom panel). Representative image of three independent experiments.

Full-size ⬛ DOI: 10.7717/peerj.14202/fig-4

(back, foot, wattle, ball of the foot, proctodeum) and skin appendages (feather follicle, beak, uropygial gland) as well as in organs with keratinizing mucosal membranes, such as the nose, pharynx, esophagus, crop, and proctodeum (Fig. 6, Table S2). Additionally, *gCLCA2* mRNA was found in the trachea, cecum, kidney, bursa of Fabricius, thyroid gland, sciatic nerve, eye and in the liver (Fig. 6, Table S2). gCLCA2 protein was exclusively detected immunohistochemically in keratinocytes of the skin, skin appendages and keratinizing mucosal membranes of the larynx, esophagus and crop (Fig. 7A, Fig. S3). gCLCA2 was localized in all layers of the stratified epithelium similar to the epithelial cell marker cytokeratins (Fig. 7B). At the subcellular level, signals consistently appeared as multiple, evenly distributed dots within the cytosol, with no specific signal enrichment detected at the plasma membrane. However, we failed to detect the CLCA2 protein in other tissues with notable *gCLCA2* mRNA presence (Fig. 6). Abundant expression in epidermal keratinocytes appears as a consistent hallmark of CLCA2 proteins, as verified in chicken, turkeys, quails, ostriches, cats, pigs and mice (Fig. 8).

DISCUSSION

Previous investigations on the *CLCA* gene family have revealed unusually complex evolutionary developments in some of its members (*Bartenschlager et al., 2022*; *Mundhenk et al., 2018*). The divergent cluster of mammalian *CLCA1/3/4* genes is characterized by multiple and independent duplication and inactivation events. This suggests flexible adaptation to environmental conditions and separation and specification of gene functions (*Mundhenk et al., 2018*; *Patel, Brett & Holtzman, 2009*; *Plog et al., 2015*). In sharp contrast,
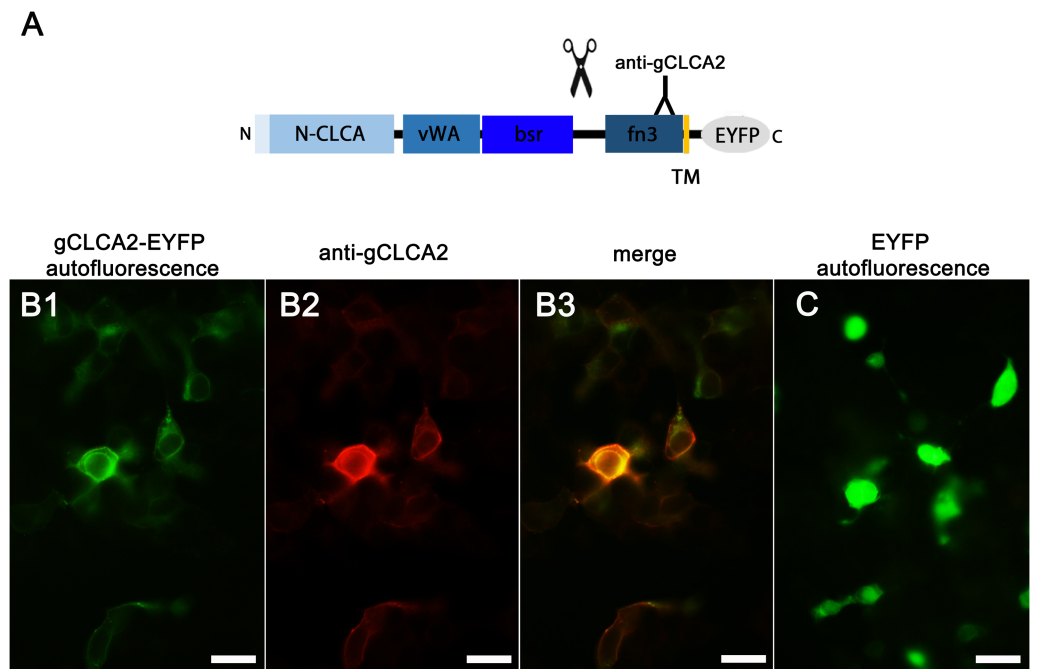
A

Figure 5 Heterologously expressed gCLCA2 is associated with the plasma membrane. (A) illustration of the anti-gCLCA2 binding site: N-CLCA, N-CLCA domain; vWA, von Willebrand factor type A domain; bsr, beta sheet rich domain; scissor sketch, putative cleavage site; fn3, fibronectin type III domain; TM, transmembrane domain; EYFP, enhanced yellow fluorescent protein tag; Y, antibody binding site. (B) auto- and immunofluorescent localization of the C-terminal cleavage product of gCLCA2 using antibody gC2 in HEK293 cells transfected with the *gCLCA2* plasmid. B1, green signal; EYFP, autofluorescence; B2, red signal anti-gCLCA2 immunofluorescence; B3, orange; merged signals of EYFP autofluorescence and anti-gCLCA2 immunofluorescence. Alexa fluor 568-conjugated secondary antibodies. (C) autofluorescence (green signals) of HEK293 cells transfected with the *EYFP* mock plasmid. Bars indicate 20 μm. Representative images of three independent experiments are shown.

Full-size 🖼 DOI: 10.7717/peerj.14202/fig-5

the *CLCA2* gene appears as a consistent single intact gene in mammalian and avian species examined here (Fig. 1A). It thus appears that *CLCA2* is the most conserved gene of the family, providing an opportunity to gain a consistent and detailed insight into basic properties of *CLCA* genes. Here, our systematic comparisons of the *CLCA2* genomic and protein structures, biochemical properties and tissue as well as cellular expression levels confirmed the high degree of conservation within and between birds and mammals.

Similar to all previously studied *CLCA* genes, avian *CLCA2* (a*CLCA2*) genes are encoded by 14 exons with small in-frame deletions, compared to ma*CLCA2* (Fig. 1B), leading to slightly shorter protein sequences. Furthermore, aCLCA2 match the canonical protein architecture of CLCA that had been described for mammalian species (*Braun et al., 2010*; *Elble et al., 2006*; *Evans, Thoreson & Beck, 2004*; *Patel, Brett & Holtzman, 2009*; *Plog et al., 2012b*), namely the consecutive sequence of a signal peptide followed by the domains N-CLCA, vWA, bsr, fn3. These domains as well as the consistent proteolytic cleavage motif located between bsr and fn3 might constitute a not yet defined common functional property of intact CLCA gene family members. Obviously, however, the evolutionary turnover
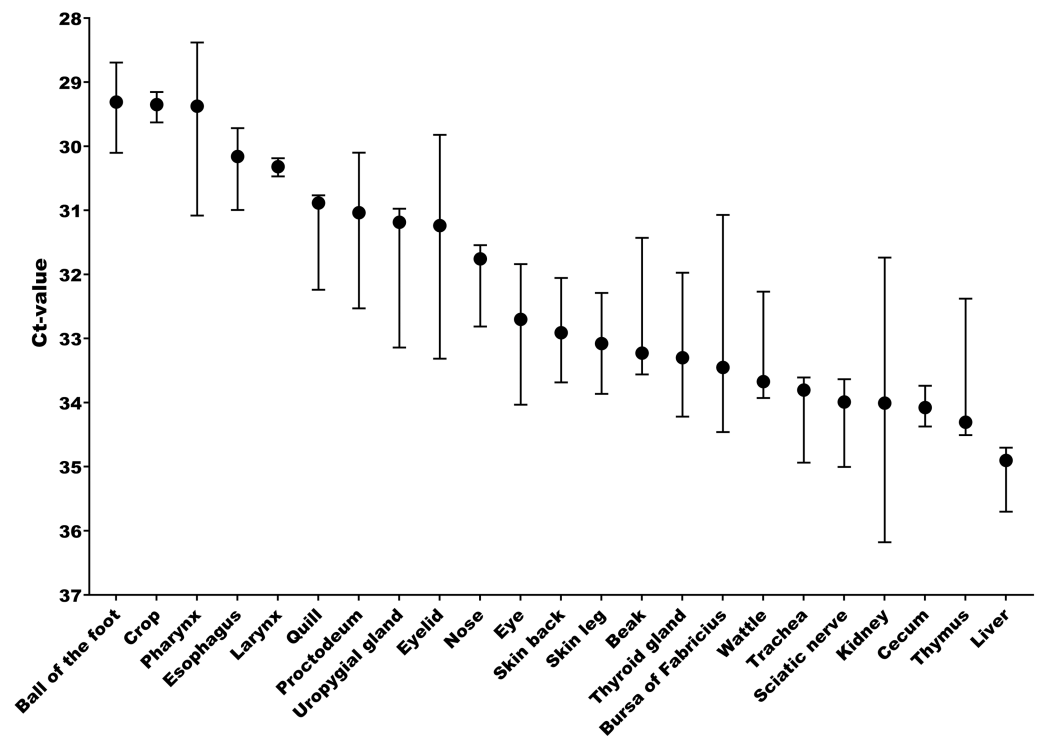
Figure 6 **Abundant g*CLCA2* mRNA was detected by quantitative RT-PCR in tissues containing keratinizing epithelium.** Black dots represent median Ct values, error bars indicate the range. $n = 3$ animals per group. Ct, cycle threshold.

Full-size 🖼 DOI: 10.7717/peerj.14202/fig-6

has modified functional properties, as suggested by the lack of an otherwise consistent TM domain in maCLCA1 (*Bartenschlager et al., 2022*). Similarly, an intact MIDAS site that is present in the mammalian CLCA1/3/4 cluster and its avian CLCA1 orthologue (*Bartenschlager et al., 2022*) was found to be interrupted in CLCA2 proteins in mammals and birds. This conserved feature of CLCA2 might point to a similar sub-functionalization in mammals and birds.

Cleavage of gCLCA2 into a larger N-terminal part and a smaller membrane-anchored C-terminal tail is presumably caused by proteolytic cleavage between the bsr and fn3 domain in the medial Golgi. This is similar to what has been found in its murine orthologue (*Braun et al., 2010*) whereas the human orthologue is cleaved extracellularly after its insertion into the plasma membrane (*Elble et al., 2006*). Thus, cleavage in the medial Golgi appears to be the evolutionarily older, prototypical process, which might have been changed in certain lineages, including humans.

The CLCA2 protein was consistently found to be expressed in epidermal keratinocytes of various cutaneous and mucosal body coverings of chicken, turkey, quail and ostrich. This highly select cellular protein expression is similar to that of human, murine, feline, and porcine orthologues (*Braun et al., 2010*; *Connon et al., 2006*; *Connon et al., 2004*; *Erickson, Gruber & Mundhenk, 2020*; *Hamalainen et al., 2021*; *Plog et al., 2012b*; *Seltmann et al.,*
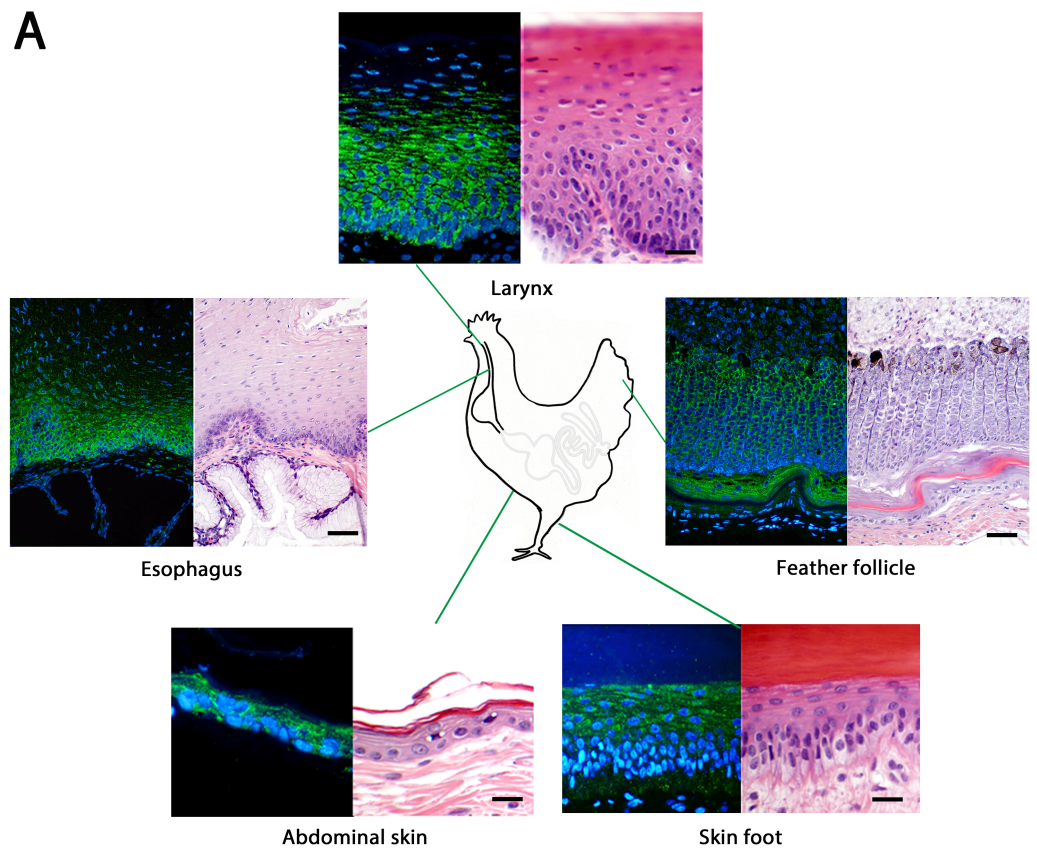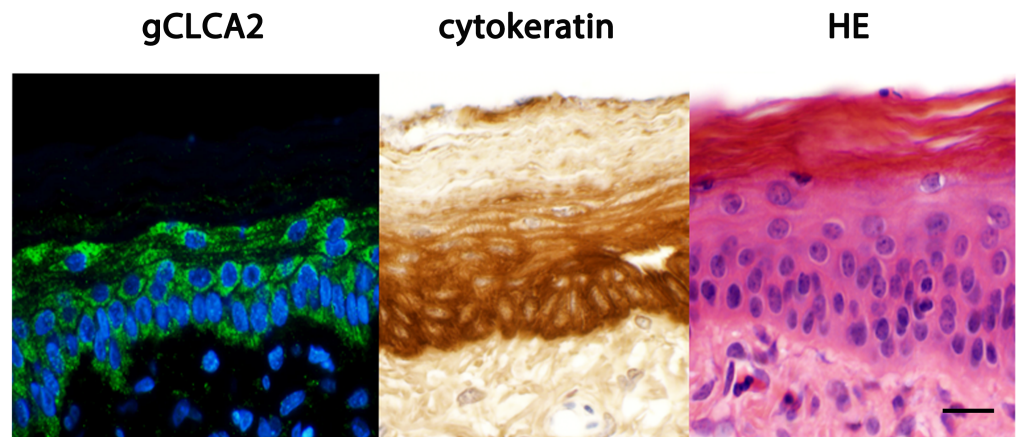
Figure 7 Immunofluorescence localization of the gCLCA2 protein in the cells of stratified/cornified epithelia of the skin, mucus membranes and feather follicles. (A) Immunofluorescence on galline tissues using the gC2 antibody (left panels), compared to hematoxylin and eosin (HE) stained tissue sections of FFPE-samples (right panels) for structural comparisons. Alexa fluor 488-conjugated secondary antibodies were used for visualization (green) and DAPI as nuclear
Full-size ⊡ DOI: 10.7717/peerj.14202/fig-7

Figure 7 (...continued)
counterstain (blue). Bars: 10 μm. (B) the gCLCA2 protein is located in stratified epithelial cells charac-
terized by the expression of cytokeratins. Left panel: Immunofluorescent localization of the gCLCA2 pro-
tein in stratified epithelium of wattle skin using antibody gC2 and Alexa fluor 488-conjugated secondary
antibodies with DAPI nuclear counterstain (blue). Central panel: immunohistochemical localization of
acidic and basic cytokeratins in a serial tissue section using the AE1/AE3 antibody cocktail, biotinylated
anti-mouse IgG secondary antibodies conjugated with horseradish peroxidase and DAB as chromogen
(brown) with hematoxylin counterstain (blue). Right panel: tissue morphology visualized using hema-
toxylin (blue) and eosin (pink) stain on a consecutive tissue section. Bars: 10 μm. Representative images
of sections from three animals.

2018). Despite the obvious differences of avian and mammalian skin anatomy (*Akat et
al., 2022*), CLCA2 expression in keratinizing epithelial cells seems to be highly conserved,
supporting the notion that skin and skin appendages including hair in mammals and
feathers in birds share a common ancestry (*Di-Poï & Milinkovitch, 2016*). The idea of the
symplesiomorphic nature of CLCA2 protein abundance in hair (*Plog et al., 2012b*) and
feather follicles might be extended to other sites as well. It had been reported that corneal
epithelial cells of galline embryos express CLCA2 protein (*Connon et al., 2006*). Like in
skin, hair and feathers, expression in corneal epithelial cells is plausible as they derive from
a common ectodermal origin with keratinocytes. Accordingly, *gCLCA2* specific mRNA was
found in the ten-week old chicken eye in our study, while no protein was detected using
immunohistochemistry. It will be interesting to explore whether impaired sensitivity for
protein detection was the reason for this discrepancy or if CLCA2 expression in corneal
epithelial cells depends on the developmental status. Similarly, the detection of *gCLCA2*
mRNA in several other tissues, albeit at lower mRNA expression levels, with concomitant
lack of detectable CLCA2 protein is in line with similar findings on murine CLCA2 (*Braun
et al., 2010*). On the one hand, the *gCLCA2* specific amplicon detected by RT-qPCR is
slightly (129 nucleotides) upstream to the sequence that encodes the antibody-binding site
for the anti-gCLCA2 antibody used in this study. Therefore, it cannot be fully excluded that
these tissues, which apparently express *gCLCA2*-mRNA but lack expression of gCLCA2
protein, may express truncated or otherwise modified variants of *gCLCA2*. On the other
hand, this discrepancy underscores established difficulties in exploring gene products
that appear at low expression levels or are restricted to small niches of expression sites.
Theoretically, non-translated transcripts of *CLCA2* might also have regulatory properties,
but to our knowledge, no such mechanisms have ever been proposed for any member of
the *CLCA* gene family. For this reason, we propose the most relevant role of CLCA2 in
keratinocytes of skin, hair and feather follicles in mammals and birds, respectively. The
highly conserved expression in the epidermis may suggest an indispensable role for CLCA2
in keratinocyte function during the course of evolution. It remains to be established,
however, to which functional aspect this may pertain, including skin barrier function,
osmolar homeostasis, cell signaling or local immunity.

Despite the fact that most of the assessed traits appear conserved between aCLCA2 and
maCLCA2, two inconsistencies were found: first, while we failed to detect the gCLCA2
protein in the airways of chicken, it is abundantly expressed in epithelial cells of the
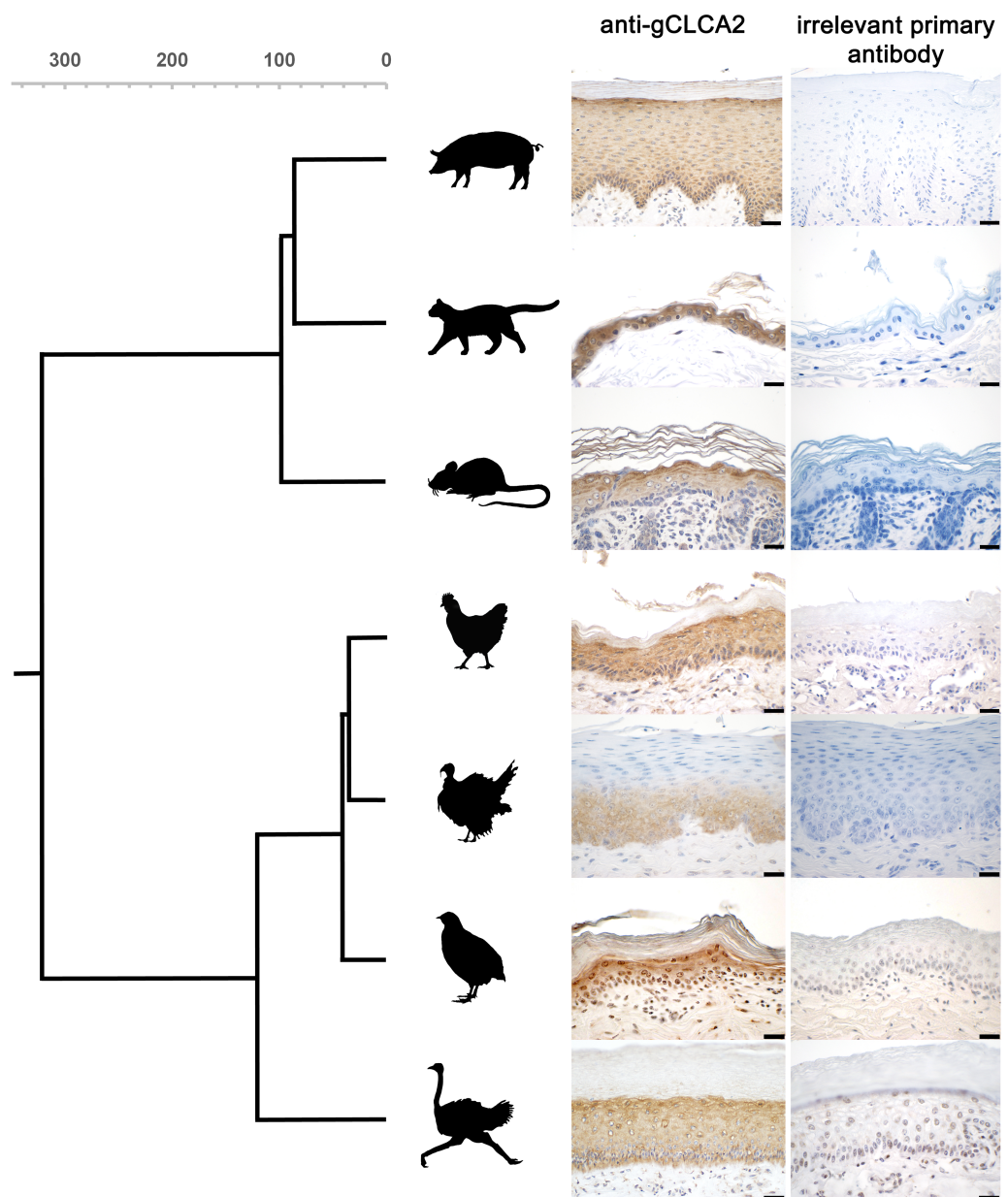respiratory tract of some albeit not all mammals (*Dietert et al., 2015*). Given the established

Figure 8 **CLCA2 expression is conserved in viable epidermal keratinocytes across avian and mammalian species.** Immunohistochemistry using the gC2 antibody and an irrelevant control antibody (anti-porcine CFTR). Biotinylated anti-mouse secondary antibodies and DAB as chromogen (brown) and hematoxylin counterstain (blue). Gray scalebar: species divergence in million years. Black bars: 20 μm. Representative images of sections from three animals per species.

Full-size ⊡ DOI: 10.7717/peerj.14202/fig-8

evolutionary relationships, this may propose a potential adaptation of CLCA2 in the respiratory tract during mammalian evolution and its subsequent loss in cats or other carnivores. Theoretically, a more ancient function of CLCA2 in the respiratory tract and an independent loss in avian and select mammalian species may be conceived, possibly to be confirmed in a common ancestor of avian and mammals. Even more

complicated to interpret is the presence of IBM motifs in the vWA and fn3 domains of CLCA2 that are thought to mediate cell–cell adhesion via interaction between CLCA2 and beta4-integrin (*Abdel-Ghany et al., 2003*; *Abdel-Ghany et al., 2001*). However, the inconsistency of functional data on this motif and its relatively loose consensus sequence F(S/N)R(I/L/V)(S/T)S raise doubt on whether this motif has any physiological function in CLCA2.

To date, knowledge about functional properties of CLCA2 are still limited, as is the case for CLCA proteins in general. CLCA2 seems to lack significant associations with human diseases, (OMIM Database entry #604003), which is usually the prime driver for generating animal models to investigate more complex functions *in vivo*. This is in contrast to ma *CLCA1*, for which its proposed role as modifier in cystic fibrosis (*Hauber et al., 2003*; *Ritzka et al., 2004*; *Young et al., 2007*) has stimulated the generation of several knockout models in mice (*Erickson et al., 2018*; *Erickson et al., 2015*; *Long et al., 2006*; *Mundhenk et al., 2012*; *Nyström et al., 2018*; *Patel et al., 2006*; *Robichaud et al., 2005*). The lack of a relevant phenotype in any of these models raises general reservations regarding the suitability of deleting a gene for exploring its function, but in the case of *CLCA*, the interpretation was complicated even further. The similar protein architectures of CLCAs and, at least in mice, the overlapping tissue expression patterns of CLCA1 and CLCA2 (*Long et al., 2006*; *Mundhenk et al., 2012*; *Patel et al., 2006*; *Robichaud et al., 2005*) immediately suggested a functional cross-compensation between different CLCA members in the respiratory tract. Given the tissue expression pattern of distinct CLCA members in the epidermis of mice (*Seltmann et al., 2018*), a mutually overlapping and possibly redundant function of certain CLCAs would also question the suitability of the mouse as model organism for studying *CLCA2* function keratinocytes. In chicken, however, the overall architecture of the *CLCA* locus is much less complex than that in mammals (*Bartenschlager et al., 2022*; *Mundhenk et al., 2018*). Moreover, the non-overlapping expression of *gCLCA1* in enterocytes (*Bartenschlager et al., 2022*) and *gCLCA2* in keratinocytes largely excludes the possibility of mutual compensation in chickens. In combination with the high degree of conservation among avian and mammalian *CLCA2* and the recent progress in genome editing techniques in birds (*Chojnacka-Puchta & Sawicka, 2020*; *Morin, Véron & Marcelle, 2017*), the simple structure of the galline *CLCA* gene locus together with its distinct expression pattern might provide a suitable setting for *in vivo* explorations of *CLCA2* functions using genetically edited chickens.

CONCLUSIONS

Our data provide strong evidence for a high conservation of *CLCA2* in mammalian and avian species during evolution. This is in stark contrast to the dynamics and proposed complex functional adaptations in the *CLCA1/3/4* cluster. The slow evolutionary dynamics of *CLCA2* genes insinuates a significant degree of negative selection and a strong functional conservation of *CLCA2* orthologues among birds and mammals, particularly in epidermal keratinocytes. This proposes gCLCA2 as a suitable object for studying basic functional properties of *CLCA* 2. Furthermore, the simple structure of the *CLCA* locus in birds and

distinct and rather simple expression patterns in chicken may serve as an ideal frame to experimentally address overall CLCA functions in the latter, rather than in mammals.

ACKNOWLEDGEMENTS

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

Grant Disclosures

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Florian Bartenschlager conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored and reviewed drafts of the article, was responsible for data curation, methodology, investigation, project administration and approved the final draft.
- Nikolai Klymiuk conceived and designed the experiments, authored or reviewed drafts of the article, and was responsible for supervision, investigation and approved the final draft.
- Achim D. Gruber conceived and designed the experiments, authored or reviewed drafts of the article, and was responsible for supervision, resources, project administration and approved the final draft.
- Lars Mundhenk conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the article, and was responsible for supervision, project administration, methodology and approved the final draft.

Animal Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

The State Office of Health and Social Affairs Berlin approved the studies (IC 114-ZH70; G 0323/06; T 0104/06).

Data Availability

The following information was supplied regarding data availability:

The raw data, the individual Ct values of gCLCA2 and PGK, a loading control, and the raw, uncropped blot images are available in the Supplementary Files.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.14202#supplemental-information.

REFERENCES

Abdel-Ghany M, Cheng H-C, Elble RC, Lin H, Di Biasio J, Pauli BU. 2003. The interacting binding domains of the $\beta4$ integrin and calcium-activated chloride channels (CLCAs) in metastasis. *Journal of Biological Chemistry* **278**:49406–49416 DOI 10.1074/jbc.M309086200.

Abdel-Ghany M, Cheng H-C, Elble RC, Pauli BU. 2001. The breast cancer $\beta4$ integrin and endothelial human CLCA2 mediate lung metastasis. *Journal of Biological Chemistry* **276**:25438–25446 DOI 10.1074/jbc.M100478200.

Akat E, Yenmiş M, Pombal MA, Molist P, Megías M, Arman S, Veselỳ M, Anderson R, Ayaz D. 2022. Comparison of vertebrate skin structure at class level: a review. *The Anatomical Record* Epub ahead of print 27 February 2022 DOI 10.1002/ar.24908.

Bartenschlager F, Klymiuk N, Weise C, Kuropka B, Gruber AD, Mundhenk L. 2022. Evolutionarily conserved properties of CLCA proteins 1, 3 and 4, as revealed by phylogenetic and biochemical studies in avian homologues. *PLOS ONE* **17**:e0266937 DOI 10.1371/journal.pone.0266937.

Bendtsen JD, Nielsen H, Von Heijne G, Brunak S. 2004. Improved prediction of signal peptides: signalP 3.0. *Journal of Molecular Biology* **340**:783–795 DOI 10.1016/j.jmb.2004.05.028.

Bothe MK, Mundhenk L, Beck CL, Kaup M, Gruber AD. 2012. Impaired autoproteolytic cleavage of mCLCA6, a murine integral membrane protein expressed in enterocytes, leads to cleavage at the plasma membrane instead of the endoplasmic reticulum. *Molecules and Cells* **33**:251–257 DOI 10.1007/s10059-012-2217-1.

Braun J, Bothe MK, Mundhenk L, Beck CL, Gruber AD. 2010. Murine mCLCA5 is expressed in granular layer keratinocytes of stratified epithelia. *Histochemistry and Cell Biology* **133**:285–299 DOI 10.1007/s00418-009-0667-0.

Chen H, Liu Y, Jiang C-J, Chen Y-M, Li H, Liu Q-A. 2019. Calcium-activated chloride channel A4 (CLCA4) plays inhibitory roles in invasion and migration through suppressing epithelial-mesenchymal transition via PI3K/AKT signaling in colorectal cancer. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research* **25**:4176 DOI 10.12659/MSM.914195.

Chojnacka-Puchta L, Sawicka D. 2020. CRISPR/Cas9 gene editing in a chicken model: current approaches and applications. *Journal of Applied Genetics* **61**:221–229 DOI 10.1007/s13353-020-00537-9.

Connon CJ, Kawasaki S, Liles M, Koizumi N, Yamasaki K, Nakamura T, Quantock AJ, Kinoshita S. 2006. Gene expression and immunolocalisation of a calcium-activated chloride channel during the stratification of cultivated and developing corneal epithelium. *Cell and Tissue Research* **323**:177–182 DOI 10.1007/s00441-005-0059-2.

Connon CJ, Yamasaki K, Kawasaki S, Quantock AJ, Koizumi N, Kinoshita S. 2004. Calcium-activated chloride channel-2 in human epithelia. *Journal of Histochemistry & Cytochemistry* **52**:415–418 DOI 10.1177/002215540405200313.

Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean Irina M, Austine-Orimoloye O, Azov Andrey G, Barnes I, Bennett R, Berry A, Bhai J, Bignell A, Billis K, Boddu S, Brooks L, Charkhchi M, Cummins C, Da Rin Fioretto L, Davidson C, Dodiya K, Donaldson S, El Houdaigui B, El Naboulsi T, Fatima R, Giron CG, Genez T, Martinez Jose G, Guijarro-Clarke C, Gymer A, Hardy M, Hollis Z, Hourlier T, Hunt T, Juettemann T, Kaikala V, Kay M, Lavidas I, Le T, Lemos D, Marugán JC, Mohanan S, Mushtaq A, Naven M, Ogeh Denye N, Parker A, Parton A, Perry M, Piližota I, Prosovetskaia I, Sakthivel Manoj P, Salam Ahamed Imran A, Schmitt Bianca M, Schuilenburg H, Sheppard D, Pérez-Silva José G, Stark W, Steed E, Sutinen K, Sukumaran R, Sumathipala D, Suner M-M, Szpak M, Thormann A, Tricomi FF, Urbina-Gómez D, Veidenberg A, Walsh Thomas A, Walts B, Willhoft N, Winterbottom A, Wass E, Chakiachvili M, Flint B, Frankish A, Giorgetti S, Haggerty L, Hunt Sarah E, Isley Garth RI, Loveland Jane E, Martin Fergal J, Moore B, Mudge Jonathan M, Muffato M, Perry E, Ruffier M, Tate J, Thybert D, Trevanion Stephen J, Dyer S, Harrison Peter W, Howe Kevin L, Yates Andrew D, Zerbino Daniel R, Flicek P. 2021. Ensembl 2022. *Nucleic Acids Research* **50**:D988–D995 DOI 10.1093/nar/gkab1049.

Di-Poï N, Milinkovitch MC. 2016. The anatomical placode in reptile scale morphogenesis indicates shared ancestry among skin appendages in amniotes. *Science Advances* **2**:e1600708 DOI 10.1126/sciadv.1600708.

Dietert K, Mundhenk L, Erickson NA, Reppe K, Hocke AC, Kummer W, Witzenrath M, Gruber AD. 2015. Murine CLCA5 is uniquely expressed in distinct niches of airway epithelial cells. *Histochemistry and Cell Biology* **143**:277–287 DOI 10.1007/s00418-014-1279-x.

Elble RC, Walia V, H-c Cheng, Connon CJ, Mundhenk L, Gruber AD, Pauli BU. 2006. The putative chloride channel hCLCA2 has a single C-terminal transmembrane segment. *Journal of Biological Chemistry* **281**:29448–29454 DOI 10.1074/jbc.M605919200.

Ellegren H. 2005. The avian genome uncovered. *Trends in Ecology & Evolution* **20**:180–186 DOI 10.1016/j.tree.2005.01.015.

Erickson NA, Dietert K, Enders J, Glauben R, Nouailles G, Gruber AD, Mundhenk L. 2018. Soluble mucus component CLCA1 modulates expression of leukotactic cytokines and BPIFA1 in murine alveolar macrophages but not in bone marrow-derived macrophages. *Histochemistry and Cell Biology* **149**:619–633 DOI 10.1007/s00418-018-1664-y.

Erickson N, Gruber A, Mundhenk L. 2020. The family of chloride channel regulator, calcium-activated proteins in the feline respiratory tract: a comparative perspective on airway Diseases in man and animal models. *Journal of Comparative Pathology* **174**:39–53 DOI 10.1016/j.jcpa.2019.10.193.

Erickson NA, Nyström EE, Mundhenk L, Arike L, Glauben R, Heimesaat MM, Fischer A, Bereswill S, Birchenough GM, Gruber AD. 2015. The goblet cell protein Clca1 (alias mClca3 or Gob-5) is not required for intestinal mucus synthesis, structure and barrier function in naive or DSS-challenged mice. *PLOS ONE* **10**:e0131991 DOI 10.1371/journal.pone.0131991.

Evans SR, Thoreson WB, Beck CL. 2004. Molecular and functional analyses of two new calcium-activated chloride channel family members from mouse eye and intestine. *Journal of Biological Chemistry* **279**:41792–41800 DOI 10.1074/jbc.M408354200.

Gibson A, Lewis AP, Affleck K, Aitken AJ, Meldrum E, Thompson N. 2005. hCLCA1 and mCLCA3 are secreted non-integral membrane proteins and therefore are not ion channels. *Journal of Biological Chemistry* **280**:27205–27212 DOI 10.1074/jbc.M504654200.

Gruber AD, Schreur KD, Ji H-L, Fuller CM, Pauli BU. 1999. Molecular cloning and transmembrane structure of hCLCA2 from human lung, trachea, and mammary gland. *American Journal of Physiology-Cell Physiology* **276**:C1261–C1270 DOI 10.1152/ajpcell.1999.276.6.C1261.

Hamalainen L, Bart G, Takabe P, Rauhala L, Deen A, Pasonen-Seppanen S, Karkkainen E, Karna R, Kumlin T, Tammi MI, Tammi RH. 2021. The calcium-activated chloride channel-associated protein rCLCA2 is expressed throughout rat epidermis, facilitates apoptosis and is downmodulated by UVB. *Histochemistry and Cell Biology* **155**:605–615 DOI 10.1007/s00418-021-01962-5.

Hammesfahr B, Odronitz F, Mühlhausen S, Waack S, Kollmar M. 2013. GenePainter: a fast tool for aligning gene structures of eukaryotic protein families, visualizing the alignments and mapping gene structures onto protein structures. *BMC Bioinformatics* **14**:1–11.

Hatje K, Hammesfahr B, Kollmar M. 2013. WebScipio: reconstructing alternative splice variants of eukaryotic proteins. *Nucleic Acids Research* **41**:W504–W509 DOI 10.1093/nar/gkt398.

Hauber HP, Manoukian JJ, Nguyen LH, Sobol SE, Levitt RC, Holroyd KJ, McElvaney NG, Griffin S, Hamid Q. 2003. Increased expression of interleukin-9, interleukin-9 receptor, and the calcium-activated chloride channel hCLCA1 in the upper airways of patients with cystic fibrosis. *The Laryngoscope* **113**:1037–1042 DOI 10.1097/00005537-200306000-00022.

Hirokawa T, Boon-Chieng S, Mitaku S. 1998. SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* **14**:378–379 DOI 10.1093/bioinformatics/14.4.378.

Hoshino M, Morita S, Iwashita H, Sagiya Y, Nagi T, Nakanishi A, Ashida Y, Nishimura O, Fujisawa Y, Fujino M. 2002. Increased expression of the human Ca2+-activated

Cl − channel 1 (CaCC1) gene in the asthmatic airway. *American Journal of Respiratory and Critical Care Medicine* **165**:1132–1136 DOI 10.1164/ajrccm.165.8.2107068.

Hou T, Zhou L, Wang L, Kazobinka G, Zhang X, Chen Z. 2017. CLCA4 inhibits bladder cancer cell proliferation, migration, and invasion by suppressing the PI3K/AKT pathway. *Oncotarget* **8**:93001 DOI 10.18632/oncotarget.21724.

Käll L, Krogh A, Sonnhammer EL. 2007. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. *Nucleic Acids Research* **35**:W429–W432 DOI 10.1093/nar/gkm256.

Koegel H, Alzheimer C. 2001. Expression and biological significance of Ca2+-activated ion channels in human keratinocytes. *The FASEB Journal* **15**:145–154 DOI 10.1096/fj.00-0055com.

Lenart A, Dudkiewicz M, Grynberg M, Pawłowski K. 2013. CLCAs-a family of metalloproteases of intriguing phylogenetic distribution and with cases of substituted catalytic sites. *PLOS ONE* **8**:e62272 DOI 10.1371/journal.pone.0062272.

Long AJ, Sypek JP, Askew R, Fish SC, Mason LE, Williams CM, Goldman SJ. 2006. Gob-5 contributes to goblet cell hyperplasia and modulates pulmonary tissue inflammation. *American Journal of Respiratory Cell and Molecular Biology* **35**:357–365 DOI 10.1165/rcmb.2005-0451OC.

Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Marchler GH, Song JS. 2020. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Research* **48**:D265–D268 DOI 10.1093/nar/gkz991.

Morin V, Véron N, Marcelle C. 2017. CRISPR/Cas9 in the chicken embryo. In: Sheng G, ed. *Avian and reptilian developmental biology. Methods in molecular biology*, vol. 1650. New York, NY: Humana Press DOI 10.1007/978-1-4939-7216-6_7.

Mundhenk L, Alfalah M, Elble RC, Pauli BU, Naim HY, Gruber AD. 2006. Both cleavage products of the mCLCA3 protein are secreted soluble proteins. *Journal of Biological Chemistry* **281**:30072–30080 DOI 10.1074/jbc.M606489200.

Mundhenk L, Erickson NA, Klymiuk N, Gruber AD. 2018. Interspecies diversity of chloride channel regulators, calcium-activated 3 genes. *PLOS ONE* **13**:e0191512 DOI 10.1371/journal.pone.0191512.

Mundhenk L, Johannesson B, Anagnostopoulou P, Braun J, Bothe MK, Schultz C, Mall MA, Gruber AD. 2012. mCLCA3 does not contribute to calcium-activated chloride conductance in murine airways. *American Journal of Respiratory Cell and Molecular Biology* **47**:87–93 DOI 10.1165/rcmb.2010-0508OC.

Nakanishi A, Morita S, Iwashita H, Sagiya Y, Ashida Y, Shirafuji H, Fujisawa Y, Nishimura O, Fujino M. 2001. Role of gob-5 in mucus overproduction and airway hyperresponsiveness in asthma. *Proceedings of the National Academy of Sciences of the United States of America* **98**:5175–5180 DOI 10.1073/pnas.081510898.

Nyström EE, Birchenough GM, van der Post S, Arike L, Gruber AD, Hansson GC, Johansson ME. 2018. Calcium-activated chloride channel regulator 1 (CLCA1) controls mucus expansion in colon by proteolytic activity. *EBioMedicine* **33**:134–143 DOI 10.1016/j.ebiom.2018.05.031.

Olias P, Adam I, Meyer A, Scharff C, Gruber AD. 2014. Reference genes for quantitative gene expression studies in multiple avian species. *PLOS ONE* **9**:e99678 DOI 10.1371/journal.pone.0099678.

Patel AC, Brett TJ, Holtzman MJ. 2009. The role of CLCA proteins in inflammatory airway disease. *Annual Review of Physiology* **71**:425–449 DOI 10.1146/annurev.physiol.010908.163253.

Patel AC, Morton JD, Kim EY, Alevy Y, Swanson S, Tucker J, Huang G, Agapov E, Phillips TE, Fuentes ME. 2006. Genetic segregation of airway disease traits despite redundancy of calcium-activated chloride channel family members. *Physiological Genomics* **25**:502–513 DOI 10.1152/physiolgenomics.00321.2005.

Pawłowski K, Lepistö M, Meinander N, Sivars U, Varga M, Wieslander E. 2006. Novel conserved hydrolase domain in the CLCA family of alleged calcium-activated chloride channels. *Proteins: Structure, Function, and Bioinformatics* **63**:424–439 DOI 10.1002/prot.20887.

Plog S, Grötzsch T, Klymiuk N, Kobalz U, Gruber AD, Mundhenk L. 2012a. The porcine chloride channel calcium-activated family member pCLCA4a mirrors lung expression of the human hCLCA4. *Journal of Histochemistry & Cytochemistry* **60**:45–56 DOI 10.1369/0022155411426455.

Plog S, Klymiuk N, Binder S, Van Hook MJ, Thoreson WB, Gruber AD, Mundhenk L. 2015. Naturally occurring deletion mutants of the pig-specific, intestinal crypt epithelial cell protein CLCA4b without apparent phenotype. *PLOS ONE* **10**:e0140050 DOI 10.1371/journal.pone.0140050.

Plog S, Mundhenk L, Bothe MK, Klymiuk N, Gruber AD. 2010. Tissue and cellular expression patterns of porcine CFTR: similarities to and differences from human CFTR. *Journal of Histochemistry & Cytochemistry* **58**:785–797 DOI 10.1369/jhc.2010.955377.

Plog S, Mundhenk L, Klymiuk N, Gruber AD. 2009. Genomic, tissue expression, and protein characterization of pCLCA1, a putative modulator of cystic fibrosis in the pig. *Journal of Histochemistry & Cytochemistry* **57**:1169–1181 DOI 10.1369/jhc.2009.954594.

Plog S, Mundhenk L, Langbein L, Gruber AD. 2012b. Synthesis of porcine pCLCA2 protein during late differentiation of keratinocytes of epidermis and hair follicle inner root sheath. *Cell and Tissue Research* **350**:445–453 DOI 10.1007/s00441-012-1482-9.

Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. 2018. HMMER web server: 2018 update. *Nucleic Acids Research* **46**:W200–W204 DOI 10.1093/nar/gky448.

Pult F, Fallah G, Braam U, Detro-Dassen S, Niculescu C, Laube B, Schmalzing G. 2011. Robust post-translocational N-glycosylation at the extreme C-terminus of membrane and secreted proteins in Xenopus laevis oocytes and HEK293 cells. *Glycobiology* **21**:1147–1160 DOI 10.1093/glycob/cwr013.

Ramena G, Yin Y, Yu Y, Walia V, Elble RC. 2016. CLCA2 interactor EVA1 is required for mammary epithelial cell differentiation. *PLOS ONE* **11**:e0147489 DOI 10.1371/journal.pone.0147489.

Range F, Mundhenk L, Gruber A. 2007. A soluble secreted glycoprotein (eCLCA1) is overexpressed due to goblet cell hyperplasia and metaplasia in horses with recurrent airway obstruction. *Veterinary Pathology* **44**:901–911 DOI 10.1354/vp.44-6-901.

Ritzka M, Stanke F, Jansen S, Gruber AD, Pusch L, Woelfl S, Veeze HJ, Halley DJ, Tümmler B. 2004. The CLCA gene locus as a modulator of the gastrointestinal basic defect in cystic fibrosis. *Human Genetics* **115**:483–491 DOI 10.1007/s00439-004-1190-y.

Robichaud A, Tuck SA, Kargman S, Tam J, Wong E, Abramovitz M, Mortimer J, Burston HE, Masson P, Hirota J. 2005. Gob-5 is not essential for mucus overproduction in preclinical murine models of allergic asthma. *American Journal of Respiratory Cell and Molecular Biology* **33**:303–314 DOI 10.1165/rcmb.2004-0372OC.

Seltmann K, Meyer M, Sulcova J, Kockmann T, Wehkamp U, Weidinger S, Keller Uaufdem, Werner S. 2018. Humidity-regulated CLCA2 protects the epidermis from hyperosmotic stress. *Science Translational Medicine* **10**:eaao4650 DOI 10.1126/scitranslmed.aao4650.

Tamura K, Stecher G, Kumar S. 2021. MEGA11: molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution* **38**:3022–3027 DOI 10.1093/molbev/msab120.

Walia V, Ding M, Kumar S, Nie D, Premkumar LS, Elble RC. 2009. hCLCA2 is a p53-inducible inhibitor of breast cancer cell proliferation. *Cancer Research* **69**:6624–6632 DOI 10.1158/0008-5472.CAN-08-4101.

Walia V, Yu Y, Cao D, Sun M, McLean JR, Hollier BG, Cheng J, Mani SA, Rao K, Premkumar L. 2012. Loss of breast epithelial marker hCLCA2 promotes epithelial-to-mesenchymal transition and indicates higher risk of metastasis. *Oncogene* **31**:2237–2246 DOI 10.1038/onc.2011.392.

Young FD, Newbigging S, Choi C, Keet M, Kent G, Rozmahel RF. 2007. Amelioration of cystic fibrosis intestinal mucous disease in mice by restoration of mCLCA3. *Gastroenterology* **133**:1928–1937.e1923 DOI 10.1053/j.gastro.2007.10.007.

Yu Y, Walia V, Elble RC. 2013. Loss of CLCA4 promotes epithelial-to-mesenchymal transition in breast cancer cells. *PLOS ONE* **8**:e83943 DOI 10.1371/journal.pone.0083943.

Yurtsever Z, Sala-Rabanal M, Randolph DT, Scheaffer SM, Roswit WT, Alevy YG, Patel AC, Heier RF, Romero AG, Nichols CG. 2012. Self-cleavage of human CLCA1 protein by a novel internal metalloprotease domain controls calcium-activated chloride channel activation. *Journal of Biological Chemistry* **287**:42138–42149 DOI 10.1074/jbc.M112.410282.