

## RESEARCH ARTICLE

# Evolutionarily conserved properties of CLCA proteins 1, 3 and 4, as revealed by phylogenetic and biochemical studies in avian homologues

Florian Bartenschlager<sup>1</sup>, Nikolai Klymiuk<sup>2,3</sup>, Christoph Weise<sup>4</sup>, Benno Kuropka<sup>4</sup>, Achim D. Gruber<sup>1</sup>, Lars Mundhenk<sup>1\*</sup>

**1** Faculty of Veterinary Medicine, Department of Veterinary Pathology, Freie Universität Berlin, Berlin, Germany, **2** Large Animal Models in Cardiovascular Research, Internal Medical Department I, Technical University of Munich, Munich, Germany, **3** Center for Innovative Medical Models, Ludwig-Maximilians University Munich, Munich, Germany, **4** Institute of Chemistry and Biochemistry, Core Facility BioSupraMol, Freie Universität Berlin, Berlin, Germany

\* [Lars.mundhenk@fu-berlin.de](mailto:Lars.mundhenk@fu-berlin.de)



## OPEN ACCESS

**Citation:** 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(4): e0266937. <https://doi.org/10.1371/journal.pone.0266937>

**Editor:** Cyril Charles, Ecole Normale Supérieure de Lyon, FRANCE

**Received:** May 20, 2021

**Accepted:** March 30, 2022

**Published:** April 13, 2022

**Copyright:** © 2022 Bartenschlager et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its [Supporting information files](#).

**Funding:** The author(s) received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

Species-specific diversities are particular features of mammalian chloride channel regulator, calcium activated (*CLCA*) genes. In contrast to four complex gene clusters in mammals, only two *CLCA* genes appear to exist in chickens. *CLCA2* is conserved in both, while only the galline *CLCA1* (*gCLCA1*) displays close genetic distance to mammalian clusters 1, 3 and 4. In this study, sequence analyses and biochemical characterizations revealed that *gCLCA1* as a putative avian prototype shares common protein domains and processing features with all mammalian *CLCA* homologues. It has a transmembrane (TM) domain in the carboxy terminal region and its mRNA and protein were detected in the alimentary canal, where the protein was localized in the apical membrane of enterocytes, similar to *CLCA4*. Both mammals and birds seem to have at least one TM domain containing *CLCA* protein with complex glycosylation in the apical membrane of enterocytes. However, some characteristic features of mammalian *CLCA1* and 3 including entire protein secretion and expression in cell types other than enterocytes seem to be dispensable for chicken. Phylogenetic analyses including twelve bird species revealed that avian *CLCA1* and mammalian *CLCA3* form clades separate from a major branch containing mammalian *CLCA1* and 4. Overall, our data suggest that *gCLCA1* and mammalian *CLCA* clusters 1, 3 and 4 stem from a common ancestor which underwent complex gene diversification in mammals but not in birds.

## Introduction

The *chloride channel regulator, calcium activated* (*CLCA*) gene family has been implicated in the pathophysiology of inflammatory airway diseases, such as asthma [1–3] or cystic fibrosis [4–6], tumor progression and metastasis [7–11]. In this context, different *CLCA* proteins have

been linked to cellular functions including mucus processing [12,13], activation of airway macrophages [14–16], the modulation of the anion channel TMEM16A [17,18], cellular differentiation such as mucus cell metaplasia [1,19–21], cellular senescence and apoptosis [7,22–27], or as adhesion molecules for metastatic tumor cells [28,29]. Due to their proposed biomedical relevance, CLCA molecules have been studied in healthy and diseased humans [2,4,5,26,30–37] and mammalian model organisms including mouse [38–44], pig [45–49], horse [3,50], and cat [51,52].

The mammalian genome contains multiple *CLCA* gene copies, which can be classified into four distinct clusters based on genetic distances [51]. All *CLCA* genes are located in a single, conserved gene locus flanked by the two genes “*outer dense fibre of sperm tails 2-like*” (*ODF2L*) and “*SH3-domain GRB2-like endophilin B1*” (*SH3GLB1*). Clusters 1 and 2 are conserved in mammals, and each species carries only one functional *CLCA1* and *CLCA2* homologue. In contrast, clusters 3 and 4 are diverse, characterized by inactivation and duplication events in several mammalian species. In pigs, sheep, and dry-nose primates including humans, the *CLCA3* gene underwent independent mutational inactivation including several frameshift and nonsense mutations as well as the introduction of an additional exon [51]. In contrast to this pseudogenization, the gene is duplicated in cattle and mice leading to two or three functional orthologues, respectively. In other species such as cats, cluster 3 consists of only a single functional orthologue. In cluster 4, the *CLCA4* gene is duplicated in pigs and mice with two or three representatives, respectively. In humans, cats, and cattle only one functional *CLCA4* gene exists in this cluster. Taken together, the *CLCA* gene family is characterized by conserved clusters, such as 1 and 2, and by diverse clusters, such as 3 and 4 in mammals [51].

This genetic heterogeneity is also reflected in the cellular expression patterns of the proteins. *CLCA1* was detected in mucin-producing cells of different mammal species [3,30,49,50,52,53] and the *CLCA2* proteins are commonly expressed in keratinocytes of stratified epithelia [27,47,52,54–56]. In contrast, the expression patterns of the cluster 3 proteins are diverse in mammalian species. A single intact *CLCA3* has so far only been found and characterized to date in cats. It is expressed in mucus-producing submucosal cells, ciliated epithelial cells, and esophageal keratinocytes [51]. In species with duplicated and intact *CLCA3* genes such as mouse and cattle (*Bos taurus*), their gene products were localized to keratinocytes, secretory epithelial cells, smooth muscle cells, and vascular endothelial cells [38,40,42,57–61]. In mammals, the *CLCA4* proteins are predominantly expressed in enterocytes [46,48,62]. The duplication products of *CLCA4* may lead to different cellular niches. In pigs, for example, one was localized to enterocytes of the villus tips of the small intestine while the other duplicate was found in crypt epithelial cells of small and large intestine [46,48]. In humans and pigs and in contrast to mice, *CLCA4* members are also expressed in the airways [33,46,63]. In cluster 3 and 4, the duplication events seem to broaden the tissue and cellular expression pattern of the *CLCA* proteins with a species-specific evolution. This duplication may be associated with a sub- or neofunctionalization, which is unknown to date.

Despite this diversity at the genomic level and in the expression pattern of the particular *CLCA* clusters in mammals, a common feature of their functional protein products is the entry into the secretory pathway and an autocatalytic posttranslational cleavage of the precursor protein into a larger amino (N)-terminal and a smaller carboxy (C)-terminal subunit [64]. The N-terminal subunit encompasses a cysteine-rich CLCA domain with a HExxH motif (PFAM identifier: pfam08434), a von Willebrand factor type A (vWA) domain (PFAM identifier: pfam13519) and a  $\beta$ -sheet rich domain [64]. The C-terminal cleavage fragment has a fibronectin type III domain (PFAM identifier: PF00041) [64]. Additionally, based on biochemical properties, the mammalian *CLCA* clusters can be divided into two groups: the proteins of cluster 1/3 are soluble proteins, which are secreted in their entirety by the cells [51,52,64–66],

whereas the proteins of cluster 2/4 are anchored in the plasma membrane via a C-terminal transmembrane domain [47,62,64,67]. Only one study on the phylogenetic distribution of the zinc-binding amino-acid (aa) motif HExxH of the CLCA domain has been published so far [68]. Furthermore, no CLCA homologues of other vertebrates have been investigated in detail to date, which limits assertions about novel acquired and conserved traits of CLCA proteins in the mammal lineage. Recently, it was reported that the *CLCA* genes in chicken are flanked by *ODF2L* and *SH3GLB1* identically to mammals; however, only two homologues were found in this gene locus [51]. One galline homologue clustered with the conserved mammalian *CLCA2* genes. In contrast, the second *CLCA* homologue, *gCLCA1*, clustered not only with the mammalian cluster 1, but also with the diverse clusters 3 and 4.

Here, we test the hypothesis that *gCLCA1* combines expressional and protein-biochemical characteristics of mammalian *CLCA1*, 3, and 4, which suggests a common ancestor with an independent expansion in the early mammalian lineage. We analyzed the genomic organization and protein structure of *gCLCA1* *in silico*, characterized its protein-biochemical properties and expression patterns in comparison to mammalian *CLCA1*, 3, and 4. The data may help to elucidate the properties and putative functions of early ancestral members of this gene family and infer novel properties of the diverse *CLCA* clusters in mammals.

## Materials and methods

### Characterization of the *gCLCA* gene locus and phylogenetic analyses

In addition to the architecture of the *gCLCA* locus [51] detailed gene positions and sizes and exon-intron boundaries were extracted from the NCBI (<https://www.ncbi.nlm.nih.gov/>) database and Plog et al. [49]. In brief, *CLCA* protein sequences were obtained from the NCBI and ensembl (<http://www.ensembl.org/index.html>) databases and aligned. Phylogenetic trees were generated using the phylip package (Phylogeny Inference Package 3.6. J. Felsenstein. Department of Genome Sciences, University of Washington, Seattle, WA, USA). First, three independent trees were generated by the maximum likelihood (ml), most parsimony (mp), and 100 neighbor joining (nj) methods. The final tree was based on the ml tree, with branch nodes occurring also in mp. Mammalian and avian *CLCA2* were used as an outgroup. For better identification of distinct evolutionary speed in different protein domains, separate trees were calculated for the N-*CLCA* domain which includes a catalytic domain (cat, corresponding to aa45-200 in chicken *CLCA1*) and a cysteine-rich domain (cys, aa201-289), the von Willebrand factor A domain (vWA, aa297-483), a  $\beta$ -sheet rich domain (bsr, aa484-698), and fibronectin type III domain (fn3, aa754-874).

### *In silico* amino-acid sequence analysis of *gCLCA1* and generation of antibodies

The aa sequence of the *gCLCA1* was analyzed via NCBI Conserved Domain Database [69], EMBL-EBI HMMER web server [70], SUPERFAMILY 2 database [71,72], Phobius webserver [73], SOSUI [74], CCTOP [75], and SignalP 5.0 [76] algorithms to identify putative protein domains. Asparagine (N)-linked glycosylation sites were predicted using NetNGlyc 1.0 (<http://www.cbs.dtu.dk/services/NetNGlyc/>). The putative proteolytic cleavage site was identified by interspecies comparison with aa sequences of mammalian *CLCA* proteins. Anti-*gCLCA1* antibodies were generated as described [49]. In brief, according to immunogenicity predictions, two oligopeptides were synthesized corresponding to aa92 to 105 (KKNSTYSRLKTESY, *gCN1*) and aa821 to 834 (ASVPSDDEGNTSDG, *gCC1*). *Limulus polyphemus* hemocyanin-conjugated peptides were used for immunization of two rabbits each. Specific IgG-antibodies

were purified from the antisera using a cyanogen bromide immunization peptide-coupled sepharose column. The immunopurified polyclonal antibodies generated against the N-terminal and C-terminal part of the gCLCA1 protein were named gC1-N1 and gC1-C1, respectively.

### Animals and tissue processing

44 tissues each (S1 File) from ten-week old female chickens (Hampshire x White Leghorn, n = 3) and the gonads of male chickens (Hampshire x White Leghorn, n = 3) of the same age were snap-frozen on dry ice subsequent to a brief immersion in 2-methylbutane. For immunohistology or immunofluorescence, tissues were fixed in 4% neutral-buffered formaldehyde for 48 h and embedded in paraffin (FFPE-tissues). All 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 (State Office of Health and Social Affairs Berlin, 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 on miscanthus litter in groups of 25, with infrared heat lamps offered until week five. The animals were fed with fledgling rearing feed until week eight and pullet feed afterwards. Females were harvested in the morning and males in the morning of the following day.

### Molecular cloning of gCLCA1 and generation of gCLCA1 mutants

*gCLCA1* was cloned and mutants were generated as described in detail in S2 File. In brief, the *gCLCA1* open reading frame (ORF) was amplified from reverse transcribed RNA extracted from the proctodeum and tagged with the enhanced yellow fluorescent protein (*EYFP*) at the C-terminus by cloning into the *pEYFP-N1* vector (*gCLCA1WT*). For analysis of an EQ mutation in the zinc-binding aa motif in accordance with Pawlowski et al. [77] and Bothe et al. [78], the wild type motif was replaced by a synthesized gene fragment containing a substitution of glutamic acid (E) with glutamine (Q) at position 164 (*gCLCA1E164Q*). As none of the specific anti-gCLCA1 antibodies were applicable for immunoblotting (see Immunoblotting section), a *gCLCA1*-fragment was substituted by a synthesized gene fragment containing a recognition site for murine N-terminal CLCA1 antibodies ( $\alpha$ -p3b2, [53]) in order to enable for the detection of the N-terminal cleavage product of the murinized wild type (*gCLCA1Nmabc1*) or murinized EQ mutant (*gCLCA1Nmabc1E164Q*) *in vitro*.

### Tissue expression pattern of gCLCA1 mRNA

The *gCLCA1* mRNA organ-specific expression pattern was analyzed by RT-qPCR as described in detail in S2 File. In brief, total RNA was isolated from tissues (S1 File), reverse transcribed, and the synthesized cDNA was diluted to a final concentration of 1 ng/ $\mu$ l. Specific, exon-boundaries spanning primers were used to detect *gCLCA1* or the reference gene phosphoglycerate kinase 1 (*PGK1*) [79] using a SYBR green qPCR assay. *gCLCA1*-mRNA was considered to be expressed, when  $C_t$ -values of 35 or below were detected in at least two of three samples tested.

### Transient transfection of HEK293 cells

HEK293 cells were transfected as described in detail in S2 File. In brief, the 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 the

cells were approximately 80–90% confluent, they were transfected using 2 µg of a plasmid (*EYFP*-mock, pcDNA3.1<sup>+</sup>-mock, *gCLCA1*WT, *gCLCA1E164Q*, *gCLCA1Nmbc1*, *gCLCA1Nmbc1E164Q*, *mCLCA1* [78], *mCLCA1E157Q* [78], *mCLCA4a* [80], or *mCLCA4aE157Q* [80]) and 8 µl polyethylenimine (PEI, 1 mg/ml) per well. 12 h after transfection, the cells were washed with phosphate buffered saline (PBS) and serum-free DMEM was added to the cells. 72 h after transfection, the cell culture medium was removed and centrifuged at 14000 x g for one h at 4°C. The supernatant was concentrated using Vivaspin 2 protein concentrator spin columns (Sartorius, Göttingen, Germany) with a 10 kDa molecular weight cutoff. The cells of each well were lysed using 500 µl radioimmunoprecipitation assay (RIPA) buffer supplemented with a protease inhibitor cocktail. The protein concentration of supernatants and cell lysates were quantified using the bicinchoninic acid (BCA) method prior to freezing at –20°C.

### Endoglycosidase treatment

For glycosylation analysis, lysates from *gCLCA1Nmbc1*-transfected cells were deglycosylated by incubation with 25 U/ml endo H, 50 U/ml PNGase F, or left untreated at 37°C overnight according to the manufacturer's protocol (New England Biolabs, Ipswich, Massachusetts, USA) and consequently immunoblotted as described in the Immunoblotting section.

### Immunoblotting

Cell lysates and supernatants of transfected cells were analyzed by immunoblotting as described in detail in [S2 File](#). In brief, samples of cell lysates or concentrated cell culture supernatant were reduced in 1,4-dithiothreitol (DTT) and separated in a 10% acrylamide gel. The proteins were transferred onto a polyvinylidene fluoride (PVDF)-membrane and blocked with 5% non-fat milk. The membranes were probed with gC1-N1 and gC1-C1 both at a three-fold dilution series from 5 µg/ml to 0.05 µg/ml, or mouse monoclonal anti-YFP (cat. G163, abm, Vancouver, Canada) at 1:500, or rabbit polyclonal anti-mCLCA1 [53] at 1:500, or rabbit polyclonal anti-mCLCA4a [62], or mouse monoclonal anti beta-actin (A5441, MilliporeSigma, Saint Louis, Missouri, USA) antibodies at 1:1000. The 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. The immunoblotted gCLCA1 protein was not detectable using antibodies gC1-N1 or gC1-C1 at any dilution used ([S3 File](#)).

### Immunocytochemistry of transfected HEK293 cells

Procedures are described in detail in [S2 File](#). In brief, cells were grown on 8-well tissue chamber slides and transfected with *gCLCA1*WT or *EYFP*-mock plasmids. 72 h after transfection, the cells were briefly fixed in ice-cold methanol (100%) for five min followed by a fixation in 4% paraformaldehyde for ten min. The cells in each well were permeabilized with 0.1% Triton X-100 in PBS and blocked with 10% goat normal serum (GS) and 0.05% Tween 20 in PBS. After blocking, the cells were probed with untreated or pre-absorbed gC1-N1, or gC1-C1, or irrelevant affinity-purified rabbit polyclonal (anti-pig CFTR [81]) antibodies (each used at 1 µg/ml). 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 for each experiment. No specific signals were detected using the gC1-N1 antibody ([S3 File](#)).

## Tissue expression pattern of gCLCA1 protein using immunofluorescence and immunohistochemistry

Methods are described in detail in [S2 File](#). In brief, all tissues from the three animals that showed a *gCLCA1*-specific Ct-value below 35 were analyzed using immunofluorescence. FFPE-tissues were cut at 3  $\mu\text{m}$  thickness, mounted on adhesive glass slides, and dewaxed. For immunohistochemistry, endogenous peroxidase was blocked by 0.5%  $\text{H}_2\text{O}_2$ . Antigens were retrieved using 1 mg/ml recombinant protease from *Streptomyces griseus* or microwave heating (600 W) for 15 min in 10 mM citric acid, pH 6.0, containing 0.05% Triton X-100. Slides were blocked with 10% Roti-ImmunoBlock and 20% GS in PBS and probed with gC1-C1 at 1  $\mu\text{g}/\text{ml}$  or rabbit monoclonal anti-villin (ab130751, Abcam, Cambridge, United Kingdom) at 1:400 or irrelevant affinity-purified rabbit polyclonal (anti-pig CFTR [81]) at 1  $\mu\text{g}/\text{ml}$  antibodies. 3,3'-Diaminobenzidine (DAB) stain was added after incubation with goat anti-rabbit biotinylated secondary antibodies (BA-1000, Vector Laboratories, Burlingame, California, USA) at 1:200 and an avidin-biotin complex solution. Slides were counterstained with 1% alcian blue in HCl at pH 1.0 and 0.1% nuclear fast red-aluminum sulfate.

For immunofluorescence analysis, tissue sections were permeabilized with 0.1% Triton X-100 in PBS and blocked with 10% GS and 0.05% Tween 20 in PBS. The slides were incubated with the gC1-C1 antibody at 1  $\mu\text{g}/\text{ml}$  or an irrelevant affinity-purified rabbit polyclonal (anti-pig CFTR) antibody [81] using the same concentration. Alexa fluor 568-conjugated goat anti-rabbit (AB\_143157, Invitrogen) secondary antibodies at 1:200 and a DAPI nuclear counterstain were used for immunofluorescence.

Unless otherwise noted, all *in vitro* studies were performed in a minimum of three independent cell batches where all transfections were conducted in parallel.

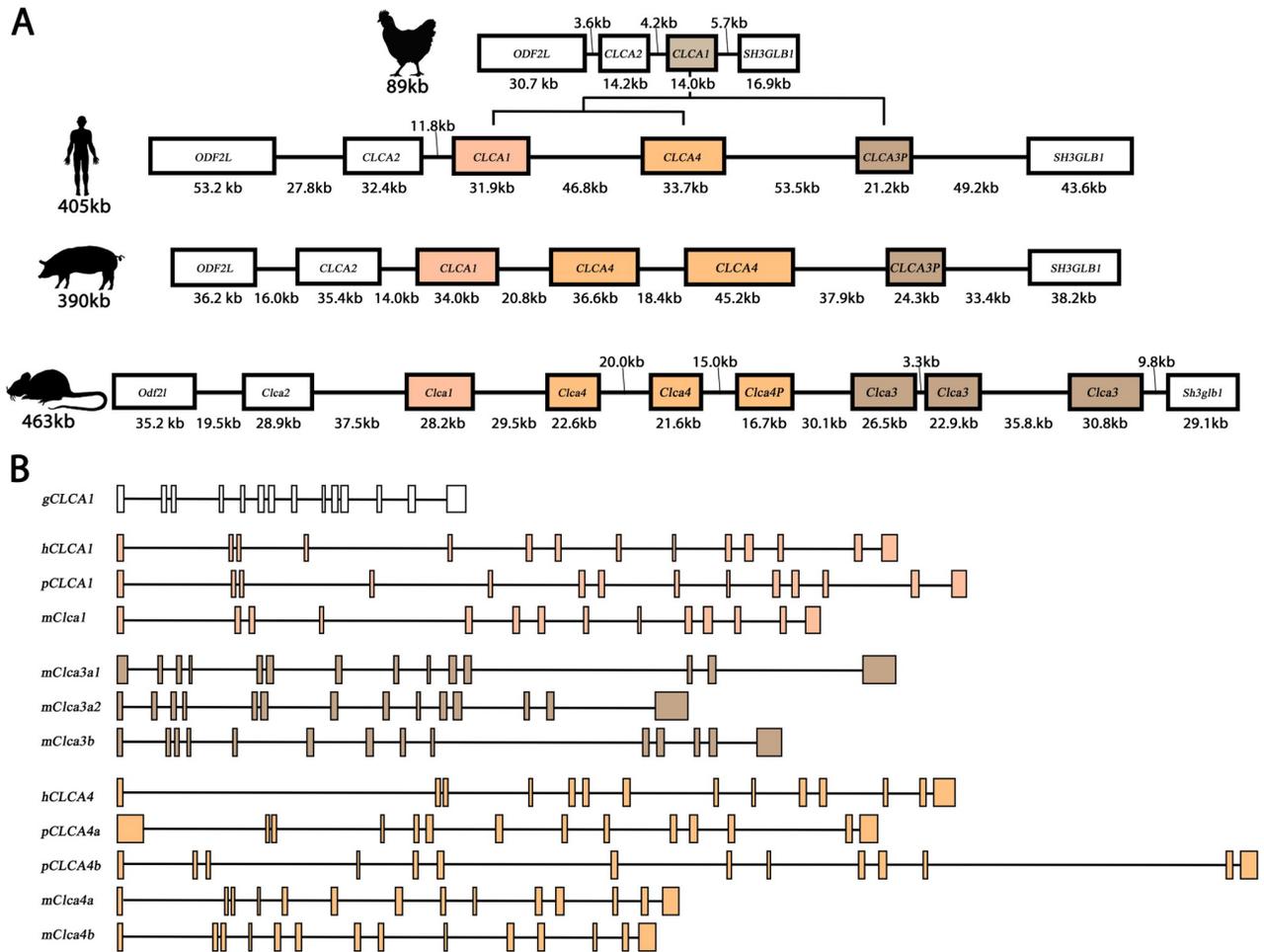
## Results

### The galline *gCLCA1* gene shares characteristics with its mammalian orthologues and encodes a putatively functional protein

The galline *CLCA* locus is located on chromosome 8 flanked by the genes *ODF2L* and *SH3GLB1* and it is shorter than mammalian *CLCA* loci ([49], [Fig 1A](#)). Compared to the human, porcine, and murine genes of cluster 1 ([49], [Fig 1A and 1B](#)), the *gCLCA1* gene is shorter, however, it is composed of 14 exons just as all intact mammalian *CLCA* genes ([Fig 1B](#)). The *gCLCA1* ORF consists of 2808 base pairs (bp, 936 aa), which is longer than those of mammalian homologues, such as human *CLCA1* (2742 bp, 914 aa) or murine *CLCA4b* (2775 bp, 925 aa) [64,82]. The predicted galline *CLCA1* gene does not contain premature stop codons or frameshift mutations as it had been identified for mammalian *CLCA3* [51] and therefore may encode a functional protein. Compared to the reference sequence of the NCBI Genbank (XM\_422360.6, *Gallus gallus*), our *gCLCA1* ORF contained four synonymous and four non-synonymous single nucleotide polymorphisms (SNPs, [S4 File](#)).

### The gCLCA1 protein contains canonical CLCA domains

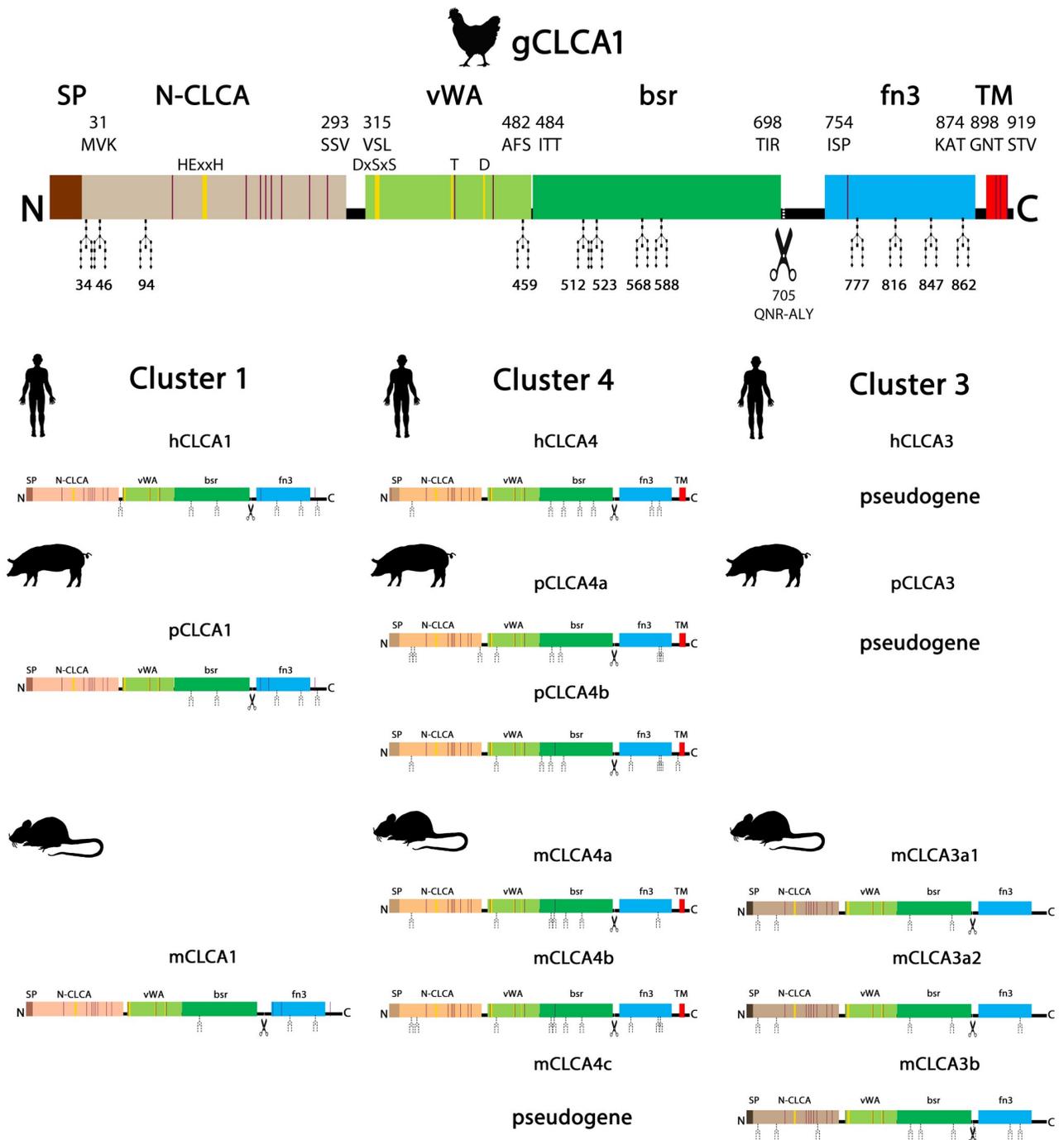
All known mammalian CLCA proteins enter a secretory pathway and contain an N-CLCA (PFAM identifier: pfam08434) and a vWA (PFAM identifier: pfam13519) domain. For *gCLCA1*, the SignalP 5.0 algorithm also predicted a signal peptide between aa1 and 29. Furthermore, the NCBI Conserved Domain Database (CDD), EMBL-EBI HMMER and SUPERFAMILY 2 databases identified an N-CLCA (aa31 to 293) and a vWA domain (aa315 to 482, [Fig 2](#)). A cysteine-rich domain within the N-terminal cleavage product, which has been described for mammalian CLCA proteins [77], was also identified in *gCLCA1* ([Fig 2](#)).



**Fig 1. *gCLCA1* gene locus and structure as compared with mammalian homologues.** (A) Comparisons of *CLCA* gene loci of the chicken and of the three mammalian representatives human, pig, and mouse. The galline *CLCA* locus consists of two *CLCA* genes, *gCLCA1* and *gCLCA2*, and is shorter than that of mammals. Genes are depicted as boxes and noncoding, intergenic segments are represented as black lines. The chicken locus is scaled 2-fold larger for illustrative purposes. P = pseudogene, kb = kilobases. (B) Comparative gene structure of *gCLCA1* and functional mammalian *CLCA1*, 3 and 4 genes. Like the mammalian *CLCA* genes of clusters 1, 3, and 4, *gCLCA1* is encoded by 14 exons (vertical boxes). Due to short intronic segments (horizontal black lines), the *gCLCA1* gene is shorter than its orthologues of the mammal cluster 1, 3, and 4.

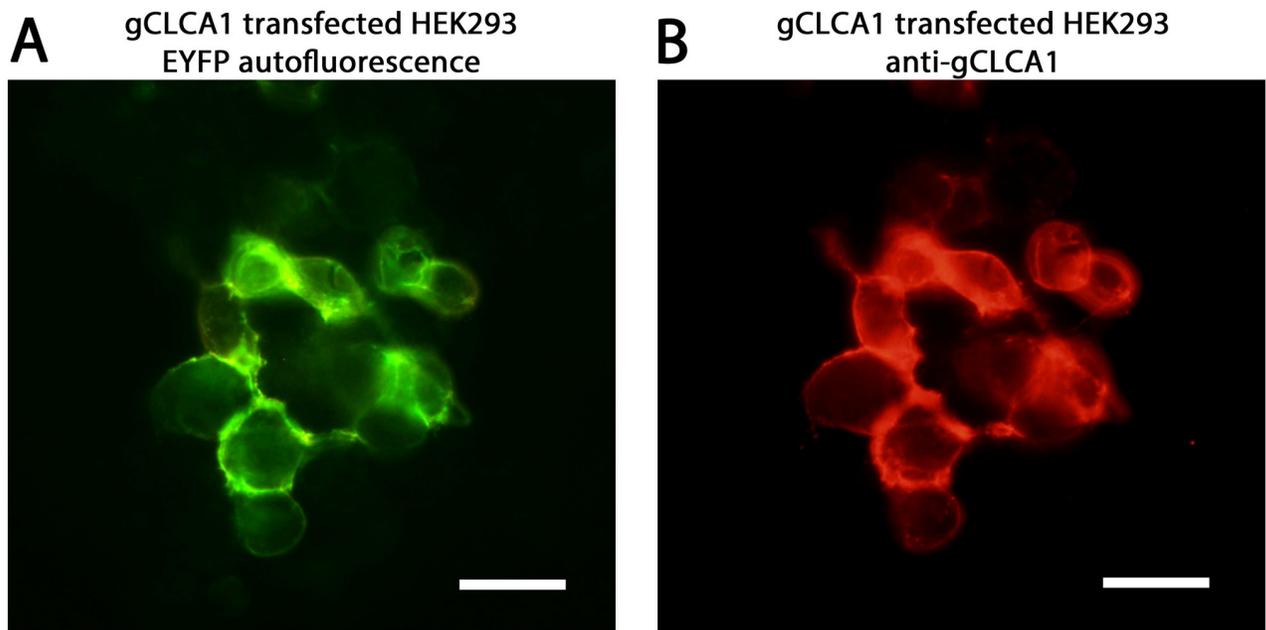
<https://doi.org/10.1371/journal.pone.0266937.g001>

Similar to the transmembrane proteins of the mammalian cluster 4, the *in silico* analysis consistently predicted a single hydrophobic transmembrane domain between aa897 (SOSUI, Phobius) / 898 (CCTOP) and aa919 (SOSUI) / 922 (CCTOP, Phobius) (Fig 2) for *gCLCA1*. To verify the presence of the predicted transmembrane domain, the cellular location of the C-terminal cleavage product of the *gCLCA1* protein was experimentally analyzed. A green signal of the C-terminal EYFP-tag was detected along the plasma membrane of cells transfected with the *gCLCA1* WT plasmid (Fig 3A). This is in contrast to the diffuse green, cytoplasmic autofluorescent EYFP signal as detected in cells transfected with the EYFP-mock plasmid (S6 File). pcDNA3.1+ transfected HEK293 cells failed to show a specific green autofluorescent signal (S6 File). A red signal along the plasma membrane was found by immunostaining of the cells with the gC1-C1 antibody which binds to the C-terminal subunit of *gCLCA1* (Fig 3B, S3 File). These results suggest the presence of a C-terminal transmembrane domain in *gCLCA1* similar to mammalian *CLCA4* proteins.



**Fig 2. The basic functional domains of mammalian CLCA proteins are conserved in gCLCA1.** Schematic depiction of the gCLCA1 protein and of CLCA proteins of clusters 1, 3, and 4 in three mammalian representatives (human, pig and mouse). A cleavable N-terminal signal peptide (dark brown box), an N-CLCA domain (N-CLCA, light brown box), a vWA-domain (vWA, light green box), a  $\beta$ -sheet rich domain (bsr, dark green box) and a fibronectin type III domain (fn3, light blue box) were predicted by *in silico* analyses and manual sequence alignments. The N-CLCA domain comprises an intact zinc-dependent metalloprotease motif (HExxH, vertical yellow bar). The vWA domain contains a MIDAS site (DxSxS, T, D, vertical yellow bar). Vertical dark red bars indicate cysteine residues, predominantly in the N-CLCA domain. The gCLCA1 protein is putatively cleaved after the amino acids QNR at aa position 705 (scissor). gCLCA1 has a putative C-terminal transmembrane domain (TM, light red box) similarly to the CLCA representatives of cluster 4. The gCLCA1 protein is an N-linked glycoprotein with 12 predicted glycosylation sites. The number of glycosylation sites is higher than in the CLCA proteins of cluster 4, followed by the proteins of cluster 3 and 1. Putative glycosylation sites are indicated by U-shaped icons. Numbers display the absolute position, for protein domains further specified by the first or last three amino acids of each domain, respectively.

<https://doi.org/10.1371/journal.pone.0266937.g002>



**Fig 3. Localization of the C-terminal cleavage product of gCLCA1 in the plasma membrane.** (A) The C-terminal cleavage product of the gCLCA1EYFP fusion protein was detected at the plasma membrane of transiently transfected HEK293 cells via autofluorescence of the EYFP-tag (green). (B) A corresponding red signal was reproducible by immunofluorescence using the gC1-C1 (anti-gCLCA1 C-terminal) antibody. Alexa fluor 568-conjugated secondary antibodies. Bars indicate 20  $\mu$ m. Representative images of three independent experiments.

<https://doi.org/10.1371/journal.pone.0266937.g003>

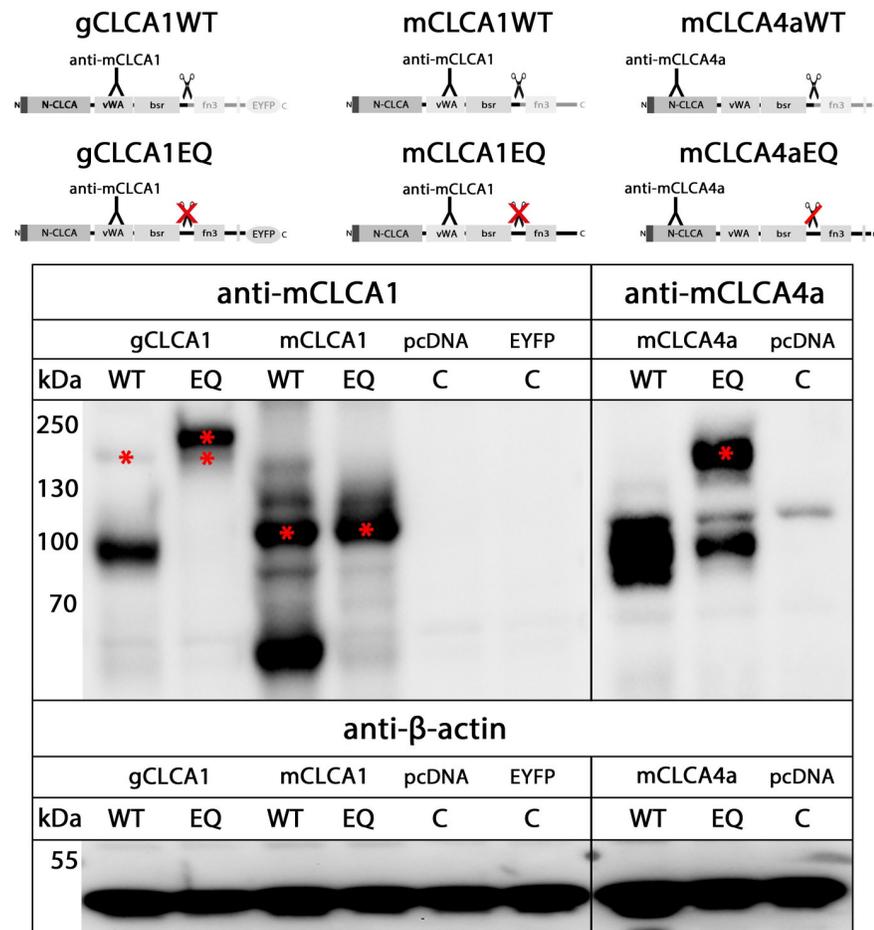
### The gCLCA1 protein has biochemical properties similar to those of mammalian CLCAs Posttranslational Cleavage

The zinc-binding aa motif HExxH of the CLCA domain, which is necessary for autocatalytic cleavage of a precursor protein into a larger N- and a smaller C-terminal cleavage product [64], is a common feature of all known mammalian CLCA proteins [77,83]. An intact HExxH motif (HEWAH, aa163–167) and a putative proteolytic cleavage site (QNR/ALY, after R705) were also identified in the gCLCA1 protein (Fig 2). It has been reported that mutagenic disruption of the mammalian HExxH motif abrogates the cleavage of mammalian CLCA1 or only impairs that of CLCA4 proteins [78,80]. To evaluate a putative HExxH-dependent cleavage of gCLCA1, lysates from cells transfected with the gCLCA1 wild type and the gCLCA1E164Q mutant were immunoblotted with anti-EYFP antibodies. For cells expressing the wild-type protein, a weak band of the precursor protein at ~154 kilodalton (kDa) and a strong band of the C-terminal cleavage product at ~64 kDa were detected (Fig 4). In contrast, no cleavage product was visible in lysates from cells transfected with the gCLCA1E164Q mutant. A single strong band of the uncleaved, mature glycosylated protein at ~166 kDa and the weak band of the immature glycosylated precursor protein at ~154 kDa (Fig 4) were detected. These findings are consistent with an autocatalytic cleavage of gCLCA1 as previously described for mammalian CLCAs. Additionally, the EQ mutation of gCLCA1 leads to a total abrogation of the cleavage event similar to mammalian CLCA1 (Fig 5).

### Secretion of the N-terminal cleavage product

It has been shown that the N-terminal subunit of mammalian CLCA4 is secreted [46,62]. To test for secretion of the N-terminal cleavage product of the gCLCA1, conditioned cell culture media and lysates from cells transfected with gCLCA1 were analyzed. The ~102 kDa N-





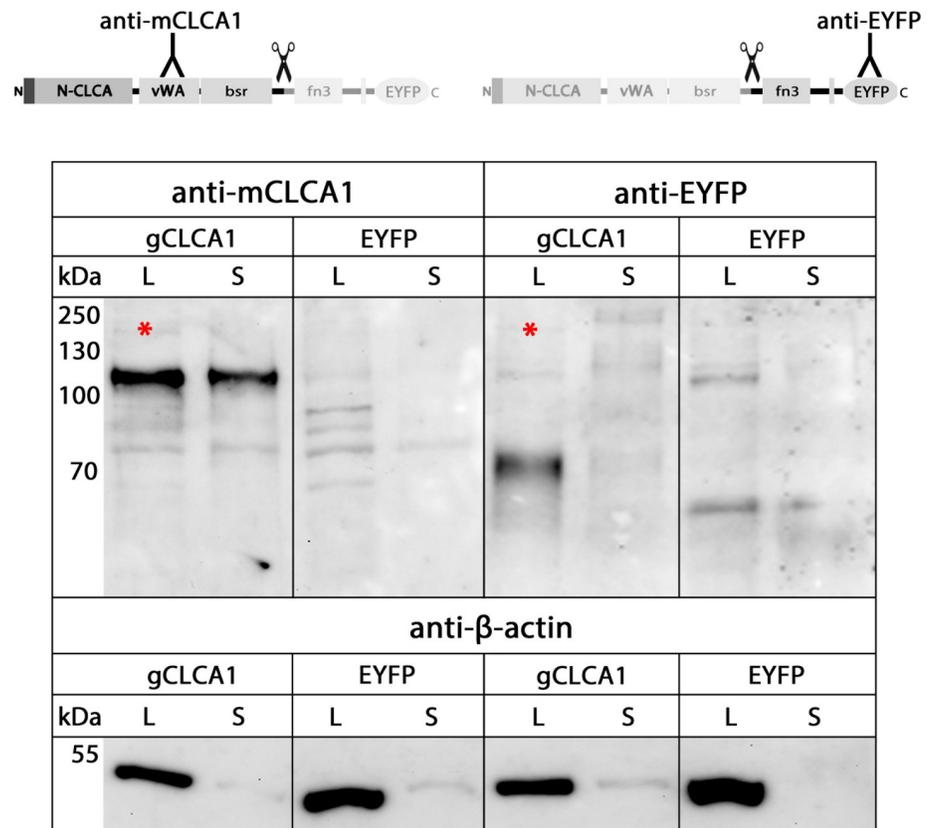
**Fig 5. The gCLCA1EQ mutant eliminated autocatalytic cleavage similar to CLCA proteins of mammalian cluster 1.** Immunoblotting of HEK293 cell lysates. The left panel illustrates immunoblots from murinized galline CLCA1, the central panel from murine CLCA1 and the right panel from murine CLCA4a constructs. The proteins of the three homologues differ in their molecular weights. The mutation of the HExxH motif of gCLCA1 eliminated the cleavage of the ~154 kDa precursor protein (\*) similarly to the murine CLCA1 protein. In contrast, the cleavage of the murine CLCA4aEQ mutant was only impaired, but not totally absent. Asterisks (\*) indicate the uncleaved protein of the respective CLCA homolog. To control for equal total protein loading, the samples were identically immunoblotted with primary anti-beta-actin antibodies. WT = HExxH wild type motif, EQ = EQ mutation in the HExxH motif, C = mock-transfected control. Representative images of three independent experiments.

<https://doi.org/10.1371/journal.pone.0266937.g005>

terminal cleavage product was detected in the supernatant and the cell lysate while the ~64 kDa C-terminal cleavage product was exclusively found in the cell lysate (Fig 6), which further supports the presence of a C-terminal transmembrane domain. Consistently, only tryptic peptides from the N-terminal cleavage product were found in the supernatant from cells transfected with the gCLCA1 using LC-ESI-MS/MS (S5 File). Thus, similar to mammalian CLCA4 proteins, the N-terminal cleavage product of gCLCA1 was secreted.

### N-glycosylation and cleavage in the medial Golgi

Mammalian CLCA proteins of clusters 1, 3, and 4 are N-linked glycoproteins, with different number of glycosylation sites [46,49,51,52,62,66]. Consistent with these results, the NetNGlyc 1.0 algorithms predicted 12 potential N-linked glycosylation sites for gCLCA1, located at aa

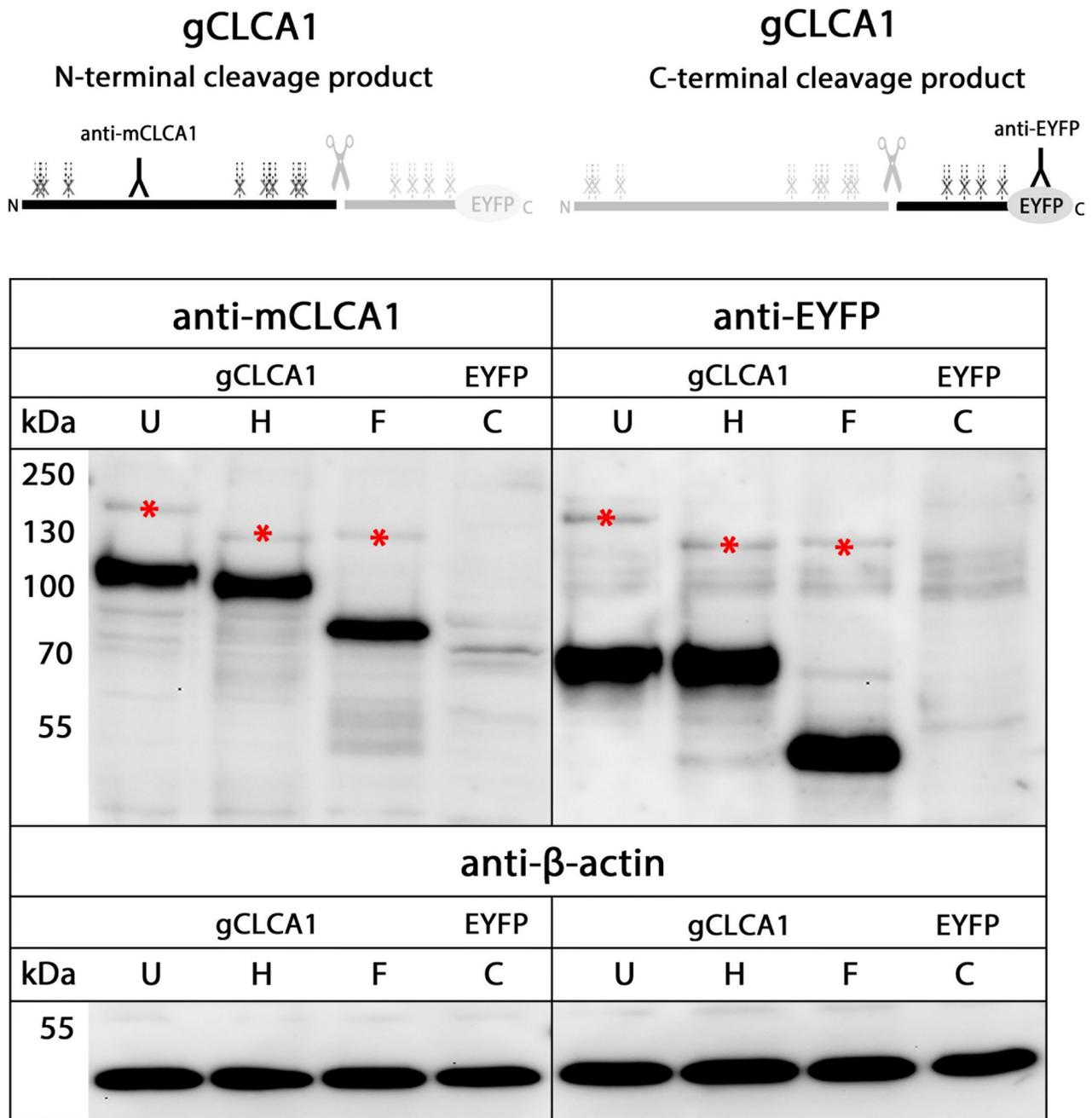


**Fig 6. Secretion of only the N-terminal cleavage product of gCLCA1 into the cell culture supernatant.**

Immunoblotting of lysates (L) and cell culture supernatants (S) from HEK293 cells transfected with the *gCLCA1Nmbc1* (gCLCA1) and *EYFP*-mock plasmid (EYFP). The N-terminal cleavage product was detected at ~102 kDa in cell lysate and supernatant. In contrast, the C-terminal, transmembrane domain containing cleavage product was exclusively detected at ~64 kDa in the cell lysate, but not in the supernatant. A weak band of the immature glycosylated gCLCA1 precursor protein was detectable in the cell lysate at ~154 kDa (\*). The same amounts of proteins for the respective lysates and supernatants from transfected cells were analyzed in these experiments. To control for equal total protein loading of cell lysates and for lack of contamination of the supernatants by cell debris, the samples were identically immunoblotted with primary anti-beta-actin antibodies. Representative images of three independent experiments.

<https://doi.org/10.1371/journal.pone.0266937.g006>

positions 34, 46, 94, 459, 512, 523, 568, 588, 777, 816, 847, and 862 (Fig 2). To corroborate these predictions, cell lysate from *gCLCA1Nmbc1*-transfected cells was treated with endoglycosidases endo H and PNGase F for identification of the kind and extent of the glycosylation. The ~154 kDa precursor protein was sensitive to endo H and PNGase F (Fig 7) resulting in a size shift from ~154 kDa to ~130 kDa which shows an immature high mannose-type glycosylation pattern. In contrast, the N- and C-terminal cleavage products were resistant to endo H but sensitive to PNGase F treatment, shown by a reduction in size from ~102 kDa to ~78 kDa (Fig 7) and ~64 kDa to ~52 kDa (Fig 7), respectively. This indicates that the majority of the predicted eight consensus glycosylation sites in the N-terminal subunit and four glycosylation sites in the C-terminal cleavage product may be used for glycosylation, when a molecular weight of ~3 kDa per site is assumed [84]. Furthermore, the complex, high mannose-rich glycosylation pattern of the N- and C-terminal cleavage products, in contrast to the precursor protein, suggests a cleavage of gCLCA1 in the medial Golgi.

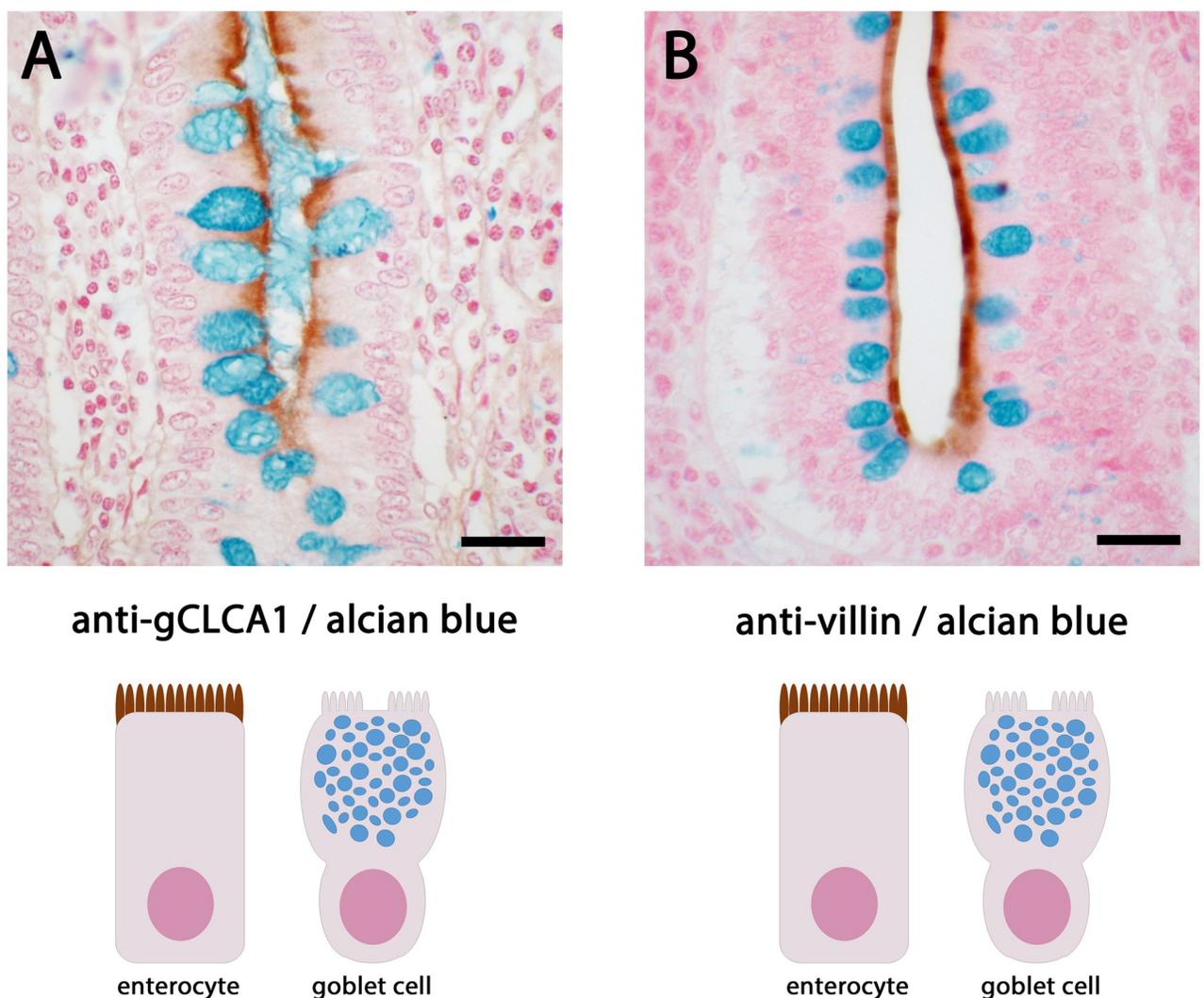


**Fig 7. Glycosylated gCLCA1 is cleaved in the medial Golgi.** Immunoblot of cell lysate of HEK293 cells transiently transfected with the *gCLCA1Nmabc1* plasmid (gCLCA1) and a control *EYFP*-mock plasmid (EYFP). The gCLCA1 precursor (\*) was mannose-rich glycosylated, reduced in size by endo H and PNGase F treatment. In contrast, both cleavage products were solely sensitive to PNGase F as shown by a reduction in size of the N-terminal and C-terminal cleavage product to ~78 kDa or ~52 kDa, respectively. The cleavage event seems to occur in the medial Golgi. U = Untreated, C = untreated mock-control, H = endo H treated, F = PNGase F treated cell lysate. To control for equal total protein loading, the samples were identically immunoblotted with primary anti-beta-actin antibodies. Representative images of three independent experiments.

<https://doi.org/10.1371/journal.pone.0266937.g007>

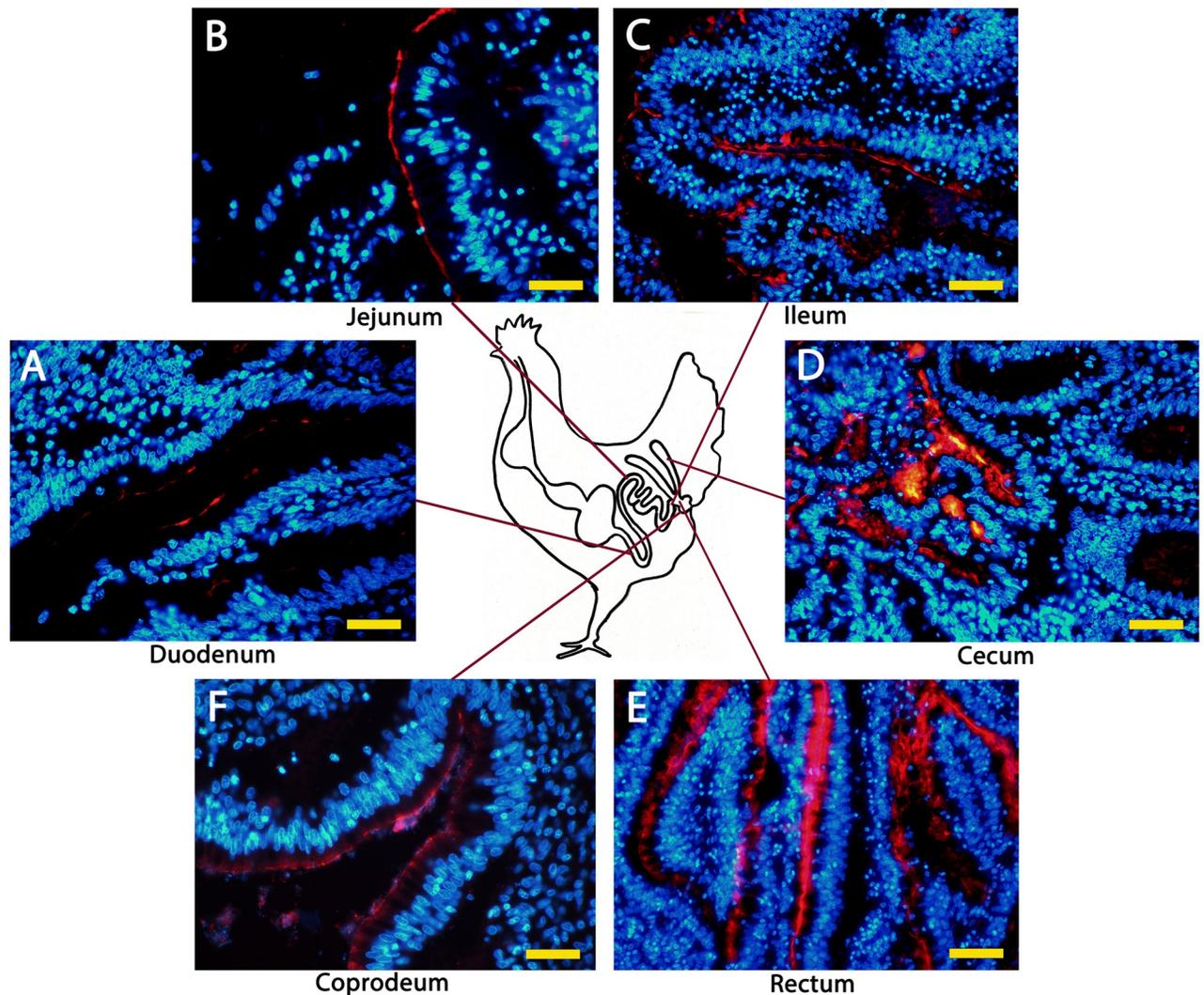
### The gCLCA1 protein is expressed in enterocytes throughout the intestine

The expression pattern of gCLCA1 was identified on the mRNA and protein levels. *gCLCA1* mRNA was detected in all segments of the alimentary tract (pharynx, esophagus, duodenum, jejunum, ileum, cecum, rectum, coprodeum), the Bursa of Fabricius, the eye, and the lung (S1 File). These tissues were further analyzed via immunofluorescence and immunohistochemistry to identify the cell types that express gCLCA1. The protein was exclusively found at the brush border of intestinal epithelial cells (Fig 8A), which also express villin, a structural marker of enterocytes (Fig 8B, [85]). However, mucin-producing intestinal goblet cells lacked any gCLCA1 expression (Fig 8A). In all segments of the intestinal tract, gCLCA1 protein was detected in enterocytes (Fig 9). Additionally, it was detected along the bursal epithelium (S7 File). No specific gCLCA1 protein signals were detected in any other tissue analyzed.



**Fig 8. gCLCA1 protein is expressed in the brush border membrane of enterocytes, but not in goblet cells.** (A) The gCLCA1 protein (brown) was detected at the apical surface of enterocytes of colonic crypts by immunohistochemistry using the gC1-C1 (anti-gCLCA1 C-terminal) antibody. A counterstain with alcian blue highlights goblet cells (dark blue) that did not show any gCLCA1 protein expression. (B) The cellular expression pattern was similar to villin, a marker for the brush border of enterocytes (brown). Color was developed using DAB as substrate (brown). Bars indicate 20  $\mu$ m.

<https://doi.org/10.1371/journal.pone.0266937.g008>



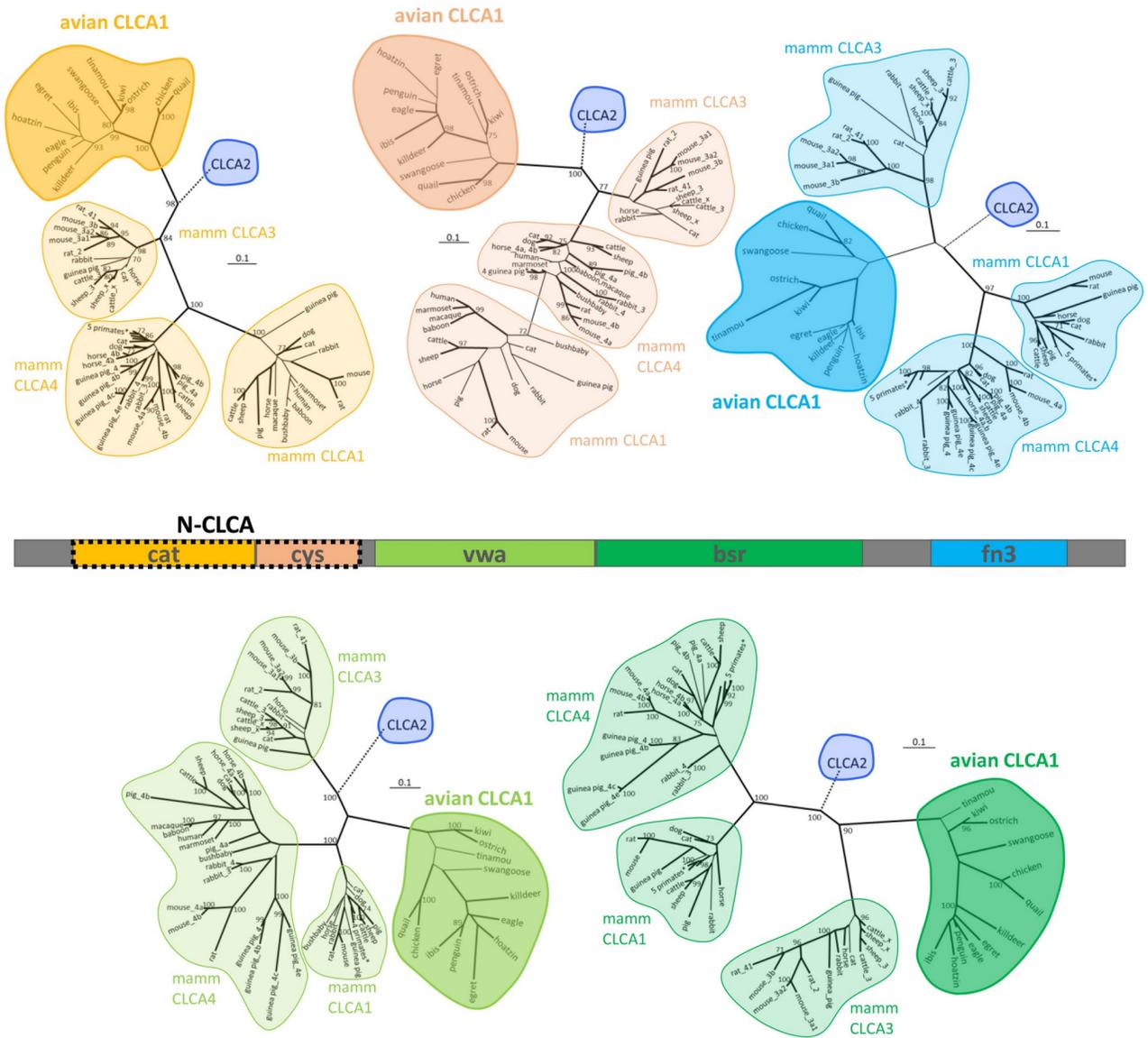
**Fig 9. gCLCA1 protein is expressed along the chicken intestine.** Anti-gCLCA1 C-terminal antibody was used to detect gCLCA1 along the epithelium in duodenum (A), jejunum (B), ileum (C), cecum (D), rectum (E) and coprodeum (F). Alexa fluor 568-conjugated secondary antibodies and DAPI counterstain (blue). Bars indicate 40 / 100  $\mu$ m.

<https://doi.org/10.1371/journal.pone.0266937.g009>

Compared to the expression pattern of CLCA proteins in mammals, the expression of *gCLCA1* showed congruence with mammalian CLCA4 proteins [46,48,62].

### The phylogenetic signal of full-length gCLCA1 scatters on the domain level

Phylogenetic analyses of five protein domains reveal a clustering of the CLCA proteins in these domains (Fig 10). Designated mammalian and avian CLCA2 form a outgroup from the avian CLCA1 and mammalian CLCA1, 3, 4. Any of the sequences cluster with the same group throughout the protein. The hierarchy within the mammalian clusters is not fully consistent, due to the known evolutionary dynamics of the rapid formation of major mammalian branches such as primates, rodents, ungulates, or carnivores within a short period of time and the independent expansion at different paces later on. Similar processes seem to have taken place in avian diversification [86] and accordingly, the relationship between the analyzed avian species remains elusive. Only close relationships such as between chicken and quail and, with



**Fig 10. Phylogenetic relationship of CLCA proteins.** Maximum likelihood (ml), most parsimony (mp) and 100 neighbor joining (nj) trees were generated for the five distinct CLCA domains and merged. Tree was based on the ml tree, with branches in thick, when occurring in the mp tree and the branch nodes indicated when occurring in more than 70 nj trees. CLCA2 orthologues were reduced to an outgroup. Other clusters were highlighted. To avoid confusion due to limited space, the five primate species analyzed (human, macaque, baboon, marmoset, bushbaby) are occasionally not separately defined (\*). In the case of CLCA duplicates in mammalian species, the sequences are designated by the abbreviation given in GenBank (i.e. “mouse\_3a2” for murine Clca3a2) or by “x” if the available sequence has been designated as “CLCA-like” or “unknown”.

<https://doi.org/10.1371/journal.pone.0266937.g010>

some restrictions, kiwi, ostrich and tinamou, are consistently reflected throughout the CLCA protein. Adding to the separation of the analyzed clusters from the CLCA2 family, mammalian CLCA1 and CLCA4 form a major branch, independent from avian CLCA1 and mammalian CLCA3. The analysis of the five domains shows variable genetic distances among avian CLCA1 and mammalian CLCA3 and CLCA1/4. The relationship between CLCA2, avian CLCA1, mammalian CLCA3 and CLCA1/4 is therefore inconsistent throughout the full-length protein.

## Discussion

In contrast to the mammalian *CLCA* gene family, which consists of up to eight members, there are only two *CLCA* homologues in chicken [51], suggesting that the mammalian and the avian *CLCA* loci have been subjected to different evolutionary pressures. *CLCA2* is conserved in mammalian and chicken loci [51], whereas the galline *CLCA1* displays low genetic distance to all other three mammalian *CLCA* proteins. Here, we characterized gCLCA1 as a putative prototype of the avian class in terms of protein structure, biochemical properties, and expression pattern and compared the results with the nature of the mammalian *CLCA* 1, 3 and 4 proteins. The gCLCA1 protein shares features of all *CLCA* family members such as a signal peptide, a *CLCA* domain with a HExxH motif, a vWA domain, and an autocatalytic cleavage site (Fig 2, Table 1). The phylogenetic analysis revealed separated *CLCA* clusters, an evident separation of the avian and mammalian *CLCA2* from the other clusters and a consistent relationship between the mammalian *CLCA1* and *CLCA4* clusters. The position of avian *CLCA1* in the phylogenetic tree was difficult to interpret, confirming the previously postulated idea of high dynamics of mammalian *CLCA* clusters 1, 3, and 4 [51]. It appears that avian *CLCA1* and its mammalian relatives may have developed independently from a common ancestor, the avian *CLCA1* thus sharing select properties of mammalian *CLCA* 1, 3, and 4.

**Table 1. Comparison of gCLCA1 with the mammalian *CLCA* 1, 3, and 4.**

Trait	gCLCA1	Mammalian <i>CLCA</i> 1	Mammalian <i>CLCA</i> 3	Mammalian <i>CLCA</i> 4
Number of exons	14	14	14	14
Duplication of genes	-	-	+	+
Pseudogenization	-	-	+	(+)
Signal peptide	+	+	+	+
N- <i>CLCA</i> domain	+	+	+	+
HExxH motif	+	+	+	+
vWA domain	+	+	+	+
Canonical cleavage site	+	+	+	+
TM domain	+	-	-	+
N-terminal subunit secreted in vitro	+	+	+	+
C-terminal subunit secreted in vitro	-	+	+	-
Posttranslational cleavage	+	+	+	+
EQ mutation abrogates cleavage	+	+	n.t.	- (*)
N-glycosylation sites	+++	+	++	+++
Protein expressing cell type	Enterocyte	Mucin-producing cells of intestinal, airway and female reproductive tract	<ul style="list-style-type: none"> <li>• Ciliated epithelial airway cell</li> <li>• mucus producing airway submucosal cell</li> <li>• keratinocyte</li> <li>• endothelial cell</li> <li>• smooth muscle cell</li> </ul>	<ul style="list-style-type: none"> <li>• Enterocyte</li> <li>• ciliated epithelial airway cell</li> </ul>

\* only partially inhibited.

n.t.: Not tested.

N: Amino terminal.

C: Carboxy terminal.

vWA: Von Willebrand factor type A.

TM: Transmembrane.

<https://doi.org/10.1371/journal.pone.0266937.t001>

Regarding the cellular expression pattern, gCLCA1 shares more properties with the CLCA4 proteins than with CLCA proteins 3 and 1 (Table 1). gCLCA1 protein was detected at the apical surface of enterocytes along the alimentary canal, and this expression pattern mirrors that of mammalian CLCA4 proteins [46,48,62]. In contrast, mucin-producing goblet cells, which abundantly express CLCA1 in all mammalian species investigated [2,3,49,50,53], did not appear to express gCLCA1. Recently, it has been shown that murine CLCA1 has mucus-processing properties [13] and controls mucus expansion in the colon [12]; the presence of a predicted equivalent protease motif in the secreted N-terminal cleavage product of gCLCA1 suggests that this protein may participate in mucus homeostasis in the chicken intestine. In addition, we found gCLCA1 protein in the epithelial lining of the cloaca and the Bursa of Fabricius, the latter is an immunological organ unique to birds. gCLCA1 mRNA was present in other organs, such as pharynx, esophagus, eye, and lung, but this was not accompanied by detectable gCLCA1 protein levels.

The presence of a transmembrane domain in the C-terminal subunit, and the secretion of only the N-terminal subunit are other traits shared by gCLCA1 and mammalian CLCA4 proteins. In contrast, mammalian CLCA1 and CLCA3 are soluble proteins and secreted in their entirety [3,64–66]. The function of the different gCLCA1 protein domains as well as the relevance of its glycosylation remain to be investigated.

gCLCA1 constructs with a modified HEXxH motif lacked posttranslational autocatalytic cleavage abilities, as reported for mutated human, murine and porcine CLCA1 [77,78]; in contrast, equivalent mutations in the HEXxH domain of CLCA4 impair autocatalysis only partially [80]. Thus, whereas gCLCA1 appears to share tissue distribution and cellular expression patterns with mammalian CLCA4, regarding autocatalytic properties it would be functionally closer to mammalian CLCA proteins of cluster 1.

### Which characteristics might a common ancestor of gCLCA1 and the mammalian CLCA 1, 3 and 4 proteins have had?

The protein and expressional characteristics of gCLCA1 and the comparison with its mammalian relatives allow us to speculate about the nature of their evolutionary ancestor. The commonly identified traits of gCLCA1 and the mammalian CLCA4 proteins, such as “proteolytic cleavage”, “anchoring of the C-terminal subunit in the plasma membrane via a transmembrane (TM) domain”, “secretion of the N-terminal subunit”, glycosylation and enterocyte-exclusive expression (Table 1) might be considered as a molecular sympleiomorphy and thus be shared also by a common ancestor.

Such a hypothesis might gain some support from more distantly related species. *In silico* analysis of the 940 aa *chloride channel accessory 4 gene 1* (CLCA4.1, NP\_001267595.1), a CLCA homologue from *Xenopus laevis*, suggested heavy glycosylation and a strongly hydrophobic C-terminus consistent with a transmembrane domain. RT-qPCR analysis had identified the gut as an expressing tissue, however, the expressing cell type of this CLCA member is still unknown [87]. Our comparison of the aa sequence of *Xenopus* CLCA4.1 (xCLCA4.1) with gCLCA1 and proteins of the mammalian CLCA 1, 3, and 4 identified a proteolytically active HEXxH site as well as a canonical cleavage site at position 696 (RSR-ALY) in the amphibian CLCA member, which suggests a putative posttranslational cleavage of the protein. Thus, a common CLCA ancestor might have been a glycosylated and cleavable transmembrane protein expressed in the intestine.

The three mammalian CLCA 1, 3, and 4, seem to have developed from duplication events that did not happen in chicken. Gene duplication, as a key driver of evolution [88], is documented in nine vertebrate species for at least 37,000 genes [89]. Either, as the most common

case [90], the duplication is completely lost or the duplicated gene becomes inactivated, resulting in pseudogenization. Pseudogenization is known for the *CLCA3* gene in certain mammals such as dry-nose primates, pigs, and sheep [51]. On the other hand, the duplication can be maintained and adapted, resulting in the gain of a new function (neofunctionalization) or the original function is now shared among both, the original and the duplicated gene (subfunctionalization) [90]. Additional species-specific duplication events of clusters 3 and 4 have been described in mammals. Three and two intact *CLCA3* genes are present in the murine or bovine CLCA locus, respectively [51]. Duplication of the *CLCA4* gene has also been identified in the pig and the mouse [51]. The duplication of functional *CLCA* genes in mammals seems to amplify the expression pattern of this gene family. In contrast to the expression of the galline CLCA1 and mammalian CLCA4 proteins in enterocytes, CLCA1 proteins are expressed by mucin-producing cells such as goblet cells of the intestinal and respiratory tract [2,3,49,53]. The expression pattern of the duplicated CLCA3 members is broader than of other clusters [51] and is not restricted to epithelial cells. CLCA3 members are also expressed in cells derived from the mesoderm such as endothelial or smooth muscle cells [42]. In addition to the diversity of the cellular expression pattern of mammalian CLCA proteins, structural variations in terms of the lack of a transmembrane domain are evident in the CLCA proteins 1 and 3 compared to 4 and the gCLCA1. The structural differences of the CLCA proteins and the diverse expression pattern may reflect that the *CLCA* gene duplication events in mammals resulted in a sub- or even neofunctionalization of the genes. The chicken lacks the diversity of mammalian *CLCA* genes and the fully secreted CLCA 1/3 homologues seem to be dispensable for the chicken. In contrast, the expression of a glycosylated, a single TM domain containing CLCA protein like gCLCA1 or mammalian CLCA4 in the intestine suggests a taxon spanning, integral trait of CLCA homologues in this anatomical section.

## Supporting information

**S1 File. RT-qPCR of *gCLCA1* and *PGK1*.**

(DOCX)

**S2 File. In-detail protocols.**

(DOCX)

**S3 File. Testing of anti-gCLCA1 primary antibodies.**

(DOCX)

**S4 File. SNPs of the *gCLCA1* clone used in this study.**

(DOCX)

**S5 File. Detection of the N-terminal cleavage products of gCLCA1 in the cell supernatant using liquid chromatography-tandem mass spectrometry (LC-MS/MS).**

(DOCX)

**S6 File. EYFP- and HEK293 autofluorescence.**

(DOCX)

**S7 File. gCLCA1 protein expression at the apical brush border of bursal surface epithelium (red).**

(DOCX)

**S1 Raw images.**

(PDF)

## Acknowledgments

The excellent technical support of Annett Kannegießer, Michaela Scholz, Charlene Lamprecht, Nicole Huth, Angela Linke, Cornelia Ziegler, Jana Enders, Sophie Merz, Laura A. Eisenhauer, and Mandana J. Poetsch is greatly appreciated. We would like to acknowledge the assistance of the Core Facility BioSupraMol supported by the German Research Foundation. This work is part of the Ph.D. thesis of F.B.

## Author Contributions

**Conceptualization:** Florian Bartenschlager, Nikolai Klymiuk, Achim D. Gruber, Lars Mundhenk.

**Data curation:** Florian Bartenschlager, Nikolai Klymiuk, Christoph Weise, Benno Kuropka, Lars Mundhenk.

**Formal analysis:** Florian Bartenschlager, Nikolai Klymiuk, Christoph Weise, Benno Kuropka, Lars Mundhenk.

**Investigation:** Florian Bartenschlager, Nikolai Klymiuk.

**Methodology:** Florian Bartenschlager, Nikolai Klymiuk, Christoph Weise, Benno Kuropka, Lars Mundhenk.

**Project administration:** Florian Bartenschlager, Achim D. Gruber, Lars Mundhenk.

**Resources:** Achim D. Gruber.

**Supervision:** Achim D. Gruber, Lars Mundhenk.

**Validation:** Florian Bartenschlager.

**Visualization:** Florian Bartenschlager, Nikolai Klymiuk.

**Writing – original draft:** Florian Bartenschlager, Nikolai Klymiuk, Benno Kuropka, Lars Mundhenk.

**Writing – review & editing:** Christoph Weise, Achim D. Gruber.

## References

1. Nakanishi A, Morita S, Iwashita H, Sagiya Y, Ashida Y, Shirafuji H, et al. Role of gob-5 in mucus overproduction and airway hyperresponsiveness in asthma. *Proceedings of the National Academy of Sciences*. 2001; 98(9):5175–80. <https://doi.org/10.1073/pnas.081510898> PMID: 11296262
2. Hoshino M, Morita S, Iwashita H, Sagiya Y, Nagi T, Nakanishi A, et al. Increased expression of the human Ca<sup>2+</sup>-activated Cl<sup>-</sup> channel 1 (CaCC1) gene in the asthmatic airway. *American journal of respiratory and critical care medicine*. 2002; 165(8):1132–6. <https://doi.org/10.1164/ajrcm.165.8.2107068> PMID: 11956057
3. Range F, Mundhenk L, Gruber A. A soluble secreted glycoprotein (eCLCA1) is overexpressed due to goblet cell hyperplasia and metaplasia in horses with recurrent airway obstruction. *Veterinary pathology*. 2007; 44(6):901–11. <https://doi.org/10.1354/vp.44-6-901> PMID: 18039903
4. Ritzka M, Stanke F, Jansen S, Gruber AD, Pusch L, Woelfl S, et al. The CLCA gene locus as a modulator of the gastrointestinal basic defect in cystic fibrosis. *Human genetics*. 2004; 115(6):483–91. <https://doi.org/10.1007/s00439-004-1190-y> PMID: 15490240
5. Hauber HP, Manoukian JJ, Nguyen LH, Sobol SE, Levitt RC, Holroyd KJ, et al. 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*. 2003; 113(6):1037–42. <https://doi.org/10.1097/00005537-200306000-00022> PMID: 12782818
6. Young FD, Newbigging S, Choi C, Keet M, Kent G, Rozmahel RF. Amelioration of cystic fibrosis intestinal mucous disease in mice by restoration of mCLCA3. *Gastroenterology*. 2007; 133(6):1928–37. e3. <https://doi.org/10.1053/j.gastro.2007.10.007> PMID: 18054564

7. Walia V, Ding M, Kumar S, Nie D, Premkumar LS, Elble RC. hCLCA2 is a p53-inducible inhibitor of breast cancer cell proliferation. *Cancer research*. 2009; 69(16):6624–32. <https://doi.org/10.1158/0008-5472.CAN-08-4101> PMID: 19654313
8. Walia V, Yu Y, Cao D, Sun M, McLean JR, Hollier BG, et al. Loss of breast epithelial marker hCLCA2 promotes epithelial-to-mesenchymal transition and indicates higher risk of metastasis. *Oncogene*. 2012; 31(17):2237–46. <https://doi.org/10.1038/onc.2011.392> PMID: 21909135
9. Yu Y, Walia V, Elble RC. Loss of CLCA4 promotes epithelial-to-mesenchymal transition in breast cancer cells. *PloS one*. 2013; 8(12):e83943. <https://doi.org/10.1371/journal.pone.0083943> PMID: 24386311
10. Hou T, Zhou L, Wang L, Kazobinka G, Zhang X, Chen Z. CLCA4 inhibits bladder cancer cell proliferation, migration, and invasion by suppressing the PI3K/AKT pathway. *Oncotarget*. 2017; 8(54):93001. <https://doi.org/10.18632/oncotarget.21724> PMID: 29190973
11. Chen H, Liu Y, Jiang C-J, Chen Y-M, Li H, Liu Q-A. 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 journal of experimental and clinical research*. 2019; 25:4176. <https://doi.org/10.12659/MSM.914195> PMID: 31164625
12. Nyström EE, Birchenough GM, van der Post S, Arike L, Gruber AD, Hansson GC, et al. Calcium-activated chloride channel regulator 1 (CLCA1) controls mucus expansion in colon by proteolytic activity. *EBioMedicine*. 2018; 33:134–43. <https://doi.org/10.1016/j.ebiom.2018.05.031> PMID: 29885864
13. Nyström EE, Arike L, Ehrencrona E, Hansson GC, Johansson ME. Calcium-activated chloride channel regulator 1 (CLCA1) forms non-covalent oligomers in colonic mucus and has mucin 2-processing properties. *Journal of Biological Chemistry*. 2019; 294(45):17075–89. <https://doi.org/10.1074/jbc.RA119.009940> PMID: 31570526
14. Ching JC, Lobanova L, Loewen ME. Secreted hCLCA1 is a signaling molecule that activates airway macrophages. *PloS one*. 2013; 8(12):e83130. <https://doi.org/10.1371/journal.pone.0083130> PMID: 24349445
15. Erickson NA, Dietert K, Enders J, Glauben R, Nouailles G, Gruber AD, et al. 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*. 2018; 149(6):619–33. <https://doi.org/10.1007/s00418-018-1664-y> PMID: 29610986
16. Keith BA, Ching JC, Loewen ME. Von Willebrand Factor Type A domain of hCLCA1 is sufficient for U-937 macrophage activation. *Biochemistry and biophysics reports*. 2019; 18:100630. <https://doi.org/10.1016/j.bbrep.2019.100630> PMID: 30984882
17. Sala-Rabanal M, Yurtsever Z, Nichols CG, Brett TJ. Secreted CLCA1 modulates TMEM16A to activate Ca<sup>2+</sup>-dependent chloride currents in human cells. *elife*. 2015; 4:e05875.
18. Sharma A, Ramena G, Yin Y, Premkumar L, Elble RC. CLCA2 is a positive regulator of store-operated calcium entry and TMEM16A. *PLoS One*. 2018; 13(5):e0196512. <https://doi.org/10.1371/journal.pone.0196512> PMID: 29758025
19. Morton J, Alevy Y, Peltz G, Holtzman M. Functional genomics of asthma: role of CLCA3 in goblet cell metaplasia (GCM) but not airway hyperreactivity (AHR). *Journal of Allergy and Clinical Immunology*. 2004; 113(2):S205.
20. Patel AC, Morton JD, Kim EY, Alevy Y, Swanson S, Tucker J, et al. Genetic segregation of airway disease traits despite redundancy of calcium-activated chloride channel family members. *Physiological genomics*. 2006; 25(3):502–13. <https://doi.org/10.1152/physiolgenomics.00321.2005> PMID: 16569774
21. Patel A, Battaile J, Alevy Y, Patterson G, Swanson S. Homologous mouse Clca3, 5, 6 and human hCLCA1, 2, 4 gene clusters control mucous cell metaplasia. *Am J Respir Crit Care Med*. 2007; 175:A499.
22. Elble RC, Pauli BU. Tumor suppression by a proapoptotic calcium-activated chloride channel in mammary epithelium. *Journal of Biological Chemistry*. 2001; 276(44):40510–7. <https://doi.org/10.1074/jbc.M104821200> PMID: 11483609
23. Beckley JR, Pauli BU, Elble RC. Re-expression of detachment-inducible chloride channel mCLCA5 suppresses growth of metastatic breast cancer cells. *Journal of Biological Chemistry*. 2004; 279(40):41634–41. <https://doi.org/10.1074/jbc.M408334200> PMID: 15292178
24. Tanikawa C, Nakagawa H, Furukawa Y, Nakamura Y, Matsuda K. CLCA2 as a p53-inducible senescence mediator. *Neoplasia*. 2012; 14(2):141–9. <https://doi.org/10.1593/neo.111700> PMID: 22431922
25. Rovillain E, Mansfield L, Caetano C, Alvarez-Fernandez M, Caballero OL, Medema RH, et al. Activation of nuclear factor-kappa B signalling promotes cellular senescence. *Oncogene*. 2011; 30(20):2356–66. <https://doi.org/10.1038/onc.2010.611> PMID: 21242976

26. Seltmann K, Meyer M, Sulcova J, Kockmann T, Wehkamp U, Weidinger S, et al. Humidity-regulated CLCA2 protects the epidermis from hyperosmotic stress. *Science translational medicine*. 2018; 10(440). <https://doi.org/10.1126/scitranslmed.aao4650> PMID: 29743348
27. Hamalainen L, Bart G, Takabe P, Rauhala L, Deen A, Pasonen-Seppanen S, et al. The calcium-activated chloride channel-associated protein rCLCA2 is expressed throughout rat epidermis, facilitates apoptosis and is downmodulated by UVB. *Histochem Cell Biol*. 2021. <https://doi.org/10.1007/s00418-021-01962-5> PMID: 33486586
28. Abdel-Ghany M, Cheng H-C, Elble RC, Pauli BU. The breast cancer  $\beta 4$  integrin and endothelial human CLCA2 mediate lung metastasis. *Journal of Biological Chemistry*. 2001; 276(27):25438–46.
29. Abdel-Ghany M, Cheng H-C, Elble RC, Lin H, DiBiasio J, Pauli BU. The interacting binding domains of the  $\beta 4$  integrin and calcium-activated chloride channels (CLCAs) in metastasis. *Journal of biological chemistry*. 2003; 278(49):49406–16.
30. Gruber AD, Elble RC, Ji H-L, Schreur KD, Fuller CM, Pauli BU. Genomic cloning, molecular characterization, and functional analysis of human CLCA1, the first human member of the family of  $\text{Ca}^{2+}$ -activated  $\text{Cl}^-$  channel proteins. *Genomics*. 1998; 54(2):200–14. <https://doi.org/10.1006/geno.1998.5562> PMID: 9828122
31. Gruber AD, Pauli BU. Tumorigenicity of human breast cancer is associated with loss of the  $\text{Ca}^{2+}$ -activated chloride channel CLCA2. *Cancer research*. 1999; 59(21):5488–91. PMID: 10554024
32. Gruber AD, Schreur KD, Ji H-L, Fuller CM, Pauli BU. Molecular cloning and transmembrane structure of hCLCA2 from human lung, trachea, and mammary gland. *American Journal of Physiology-Cell Physiology*. 1999; 276(6):C1261–C70.
33. Agnel M, Vermet T, Culouscou J-M. Identification of three novel members of the calcium-dependent chloride channel (CaCC) family predominantly expressed in the digestive tract and trachea. *FEBS letters*. 1999; 455(3):295–301. [https://doi.org/10.1016/s0014-5793\(99\)00891-1](https://doi.org/10.1016/s0014-5793(99)00891-1) PMID: 10437792
34. Bustin SA, Li S-R, Dorudi S. Expression of the  $\text{Ca}^{2+}$ -activated chloride channel genes CLCA1 and CLCA2 is downregulated in human colorectal cancer. *DNA and cell biology*. 2001; 20(6):331–8. <https://doi.org/10.1089/10445490152122442> PMID: 11445004
35. Kamada F, Suzuki Y, Shao C, Tamari M, Hasegawa K, Hirota T, et al. Association of the hCLCA1 gene with childhood and adult asthma. *Genes & Immunity*. 2004; 5(7):540–7. <https://doi.org/10.1038/sj.gene.6364124> PMID: 15318163
36. Van Der Doef H, Slieker M, Staab D, Alizadeh B, Seia M, Colombo C, et al. Association of the CLCA1 p. S357N variant with meconium ileus in European patients with cystic fibrosis. *Journal of pediatric gastroenterology and nutrition*. 2010; 50(3):347–9. <https://doi.org/10.1097/MPG.0b013e3181afce6c> PMID: 20179644
37. Pan X, Wang Q, Xu C, Yan L, Pang S, Gan J. Prognostic value of chloride channel accessory mRNA expression in colon cancer. *Oncology letters*. 2019; 18(3):2967–76. <https://doi.org/10.3892/ol.2019.10615> PMID: 31404307
38. Gandhi R, Elble R, Gruber A, Schreur K, Ji H-L, Fuller C, et al. Molecular and functional characterization of a calcium-sensitive chloride channel from mouse lung. *Journal of Biological Chemistry*. 1998; 273(48):32096–101. <https://doi.org/10.1074/jbc.273.48.32096> PMID: 9822685
39. Romio L, Musante L, Cinti R, Seri M, Moran O, Zegarra-Moran O, et al. Characterization of a murine gene homologous to the bovine CaCC chloride channel. *Gene*. 1999; 228(1–2):181–8. [https://doi.org/10.1016/s0378-1119\(98\)00620-9](https://doi.org/10.1016/s0378-1119(98)00620-9) PMID: 10072771
40. Lee D, Ha S, Kho Y, Kim J, Cho K, Baik M, et al. Induction of mouse  $\text{Ca}^{2+}$ -sensitive chloride channel 2 gene during involution of mammary gland. *Biochemical and biophysical research communications*. 1999; 264(3):933–7. <https://doi.org/10.1006/bbrc.1999.1583> PMID: 10544033
41. Komiya T, Tanigawa Y, Hirohashi S. Cloning and identification of the gene gob-5, which is expressed in intestinal goblet cells in mice. *Biochemical and biophysical research communications*. 1999; 255(2):347–51. <https://doi.org/10.1006/bbrc.1999.0168> PMID: 10049711
42. Elble RC, Ji G, Nehrke K, DeBiasio J, Kingsley PD, Kotlikoff MI, et al. Molecular and functional characterization of a murine calcium-activated chloride channel expressed in smooth muscle. *Journal of Biological Chemistry*. 2002; 277(21):18586–91. <https://doi.org/10.1074/jbc.M200829200> PMID: 11896056
43. Evans SR, Thoreson WB, Beck CL. Molecular and functional analyses of two new calcium-activated chloride channel family members from mouse eye and intestine. *Journal of Biological Chemistry*. 2004; 279(40):41792–800. <https://doi.org/10.1074/jbc.M408354200> PMID: 15284223
44. Al-Jumaily M, Kozlenkov A, Mechaly I, Fichard A, Matha V, Scamps F, et al. Expression of three distinct families of calcium-activated chloride channel genes in the mouse dorsal root ganglion. *Neuroscience bulletin*. 2007; 23(5):293–9. <https://doi.org/10.1007/s12264-007-0044-8> PMID: 17952139

45. Gaspar K, Racette K, Gordon J, Loewen M, Forsyth G. Cloning a chloride conductance mediator from the apical membrane of porcine ileal enterocytes. *Physiological genomics*. 2000; 3(2):101–11. <https://doi.org/10.1152/physiolgenomics.2000.3.2.101> PMID: 11015605
46. Plog S, Grötzsch T, Klymiuk N, Kobalz U, Gruber AD, Mundhenk L. The porcine chloride channel calcium-activated family member pCLCA4a mirrors lung expression of the human hCLCA4. *Journal of Histochemistry & Cytochemistry*. 2012; 60(1):45–56. <https://doi.org/10.1369/0022155411426455> PMID: 22205680
47. Plog S, Mundhenk L, Langbein L, Gruber AD. Synthesis of porcine pCLCA2 protein during late differentiation of keratinocytes of epidermis and hair follicle inner root sheath. *Cell and tissue research*. 2012; 350(3):445–53. <https://doi.org/10.1007/s00441-012-1482-9> PMID: 22968961
48. Plog S, Klymiuk N, Binder S, Van Hook MJ, Thoreson WB, Gruber AD, et al. Naturally Occurring Deletion Mutants of the Pig-Specific, Intestinal Crypt Epithelial Cell Protein CLCA4b without Apparent Phenotype. *PLoS one*. 2015; 10(10):e0140050. <https://doi.org/10.1371/journal.pone.0140050> PMID: 26474299
49. Plog S, Mundhenk L, Klymiuk N, Gruber AD. Genomic, tissue expression, and protein characterization of pCLCA1, a putative modulator of cystic fibrosis in the pig. *Journal of Histochemistry & Cytochemistry*. 2009; 57(12):1169–81.
50. Anton F, Leverkoehne I, Mundhenk L, Thoreson WB, Gruber AD. Overexpression of eCLCA1 in small airways of horses with recurrent airway obstruction. *Journal of Histochemistry & Cytochemistry*. 2005; 53(8):1011–21. <https://doi.org/10.1369/jhc.4A6599.2005> PMID: 15879574
51. Mundhenk L, Erickson NA, Klymiuk N, Gruber AD. Interspecies diversity of chloride channel regulators, calcium-activated 3 genes. *PLoS one*. 2018; 13(1):e0191512. <https://doi.org/10.1371/journal.pone.0191512> PMID: 29346439
52. Erickson N, Gruber A, Mundhenk L. 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*. 2020; 174:39–53. <https://doi.org/10.1016/j.jcpa.2019.10.193> PMID: 31955802
53. Leverkoehne I, Gruber AD. The murine mCLCA3 (alias gob-5) protein is located in the mucin granule membranes of intestinal, respiratory, and uterine goblet cells. *Journal of Histochemistry & Cytochemistry*. 2002; 50(6):829–38.
54. KOEGEL H, ALZHEIMER C. Expression and biological significance of Ca<sup>2+</sup>-activated ion channels in human keratinocytes. *The FASEB Journal*. 2001; 15(1):145–54. <https://doi.org/10.1096/fj.00-0055com> PMID: 11149902
55. Connon CJ, Yamasaki K, Kawasaki S, Quantock AJ, Koizumi N, Kinoshita S. Calcium-activated chloride channel-2 in human epithelia. *Journal of Histochemistry & Cytochemistry*. 2004; 52(3):415–8. <https://doi.org/10.1177/002215540405200313> PMID: 14966209
56. Braun J, Bothe MK, Mundhenk L, Beck CL, Gruber AD. Murine mCLCA5 is expressed in granular layer keratinocytes of stratified epithelia. *Histochem Cell Biol*. 2010; 133(3):285–99. <https://doi.org/10.1007/s00418-009-0667-0> PMID: 20012443
57. Gruber AD, Gandhi R, Pauli BU. The murine calcium-sensitive chloride channel (mCaCC) is widely expressed in secretory epithelia and in other select tissues. *Histochem Cell Biol*. 1998; 110(1):43–9. <https://doi.org/10.1007/s004180050263> PMID: 9681688
58. Abdel-Ghany M, Cheng H-C, Eible RC, Pauli BU. Focal adhesion kinase activated by  $\beta$ 4 integrin ligation to mCLCA1 mediates early metastatic growth. *Journal of Biological Chemistry*. 2002; 277(37):34391–400. <https://doi.org/10.1074/jbc.M205307200> PMID: 12110680
59. Furuya M, Kirschbaum SB, Paulovich A, Pauli BU, Zhang H, Alexander JS, et al. Lymphatic endothelial murine chloride channel calcium-activated 1 is a ligand for leukocyte LFA-1 and Mac-1. *The Journal of Immunology*. 2010; 185(10):5769–77. <https://doi.org/10.4049/jimmunol.1002226> PMID: 20937843
60. Roussa E, Wittschen P, Wolff NA, Torchalski B, Gruber AD, Thévenod F. Cellular distribution and sub-cellular localization of mCLCA1/2 in murine gastrointestinal epithelia. *Journal of Histochemistry & Cytochemistry*. 2010; 58(7):653–68. <https://doi.org/10.1369/jhc.2010.955211> PMID: 20385786
61. Leverkoehne I, Horstmeier BA, von Samson-Himmelstjerna G, Scholte BJ, Gruber AD. Real-time RT-PCR quantitation of mCLCA1 and mCLCA2 reveals differentially regulated expression in pre- and post-natal murine tissues. *Histochem Cell Biol*. 2002; 118(1):11–7. <https://doi.org/10.1007/s00418-002-0420-4> PMID: 12122442
62. Bothe MK, Braun J, Mundhenk L, Gruber AD. Murine mCLCA6 is an integral apical membrane protein of non-goblet cell enterocytes and co-localizes with the cystic fibrosis transmembrane conductance regulator. *Journal of Histochemistry & Cytochemistry*. 2008; 56(5):495–509. <https://doi.org/10.1369/jhc.2008.950592> PMID: 18285349

63. Mall M, Gonska T, Thomas J, Schreiber R, Seydewitz HH, Kuehr J, et al. Modulation of Ca<sup>2+</sup>-activated Cl<sup>-</sup> secretion by basolateral K<sup>+</sup> channels in human normal and cystic fibrosis airway epithelia. *Pediatr Res*. 2003; 53(4):608–18. <https://doi.org/10.1203/01.PDR.0000057204.51420.DC> PMID: 12612194
64. Patel AC, Brett TJ, Holtzman MJ. The role of CLCA proteins in inflammatory airway disease. *Annu Rev Physiol*. 2009; 71:425–49. <https://doi.org/10.1146/annurev.physiol.010908.163253> PMID: 18954282
65. Gibson A, Lewis AP, Affleck K, Aitken AJ, Meldrum E, Thompson N. hCLCA1 and mCLCA3 are secreted non-integral membrane proteins and therefore are not ion channels. *Journal of Biological Chemistry*. 2005; 280(29):27205–12.
66. Mundhenk L, Alfalah M, Elble RC, Pauli BU, Naim HY, Gruber AD. Both cleavage products of the mCLCA3 protein are secreted soluble proteins. *Journal of Biological Chemistry*. 2006; 281(40):30072–80. <https://doi.org/10.1074/jbc.M606489200> PMID: 16895902
67. Elble RC, Walia V, Cheng H-c, Cannon CJ, Mundhenk L, Gruber AD, et al. The putative chloride channel hCLCA2 has a single C-terminal transmembrane segment. *Journal of Biological Chemistry*. 2006; 281(40):29448–54.
68. Lenart A, Dudkiewicz M, Grynberg M, Pawłowski K. CLCAs—a family of metalloproteases of intriguing phylogenetic distribution and with cases of substituted catalytic sites. *PLoS One*. 2013; 8(5):e62272. <https://doi.org/10.1371/journal.pone.0062272> PMID: 23671590
69. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, et al. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic acids research*. 2020; 48(D1):D265–D8. <https://doi.org/10.1093/nar/gkz991> PMID: 31777944
70. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 2018 update. *Nucleic acids research*. 2018; 46(W1):W200–W4. <https://doi.org/10.1093/nar/gky448> PMID: 29905871
71. Gough J. The SUPERFAMILY database in structural genomics. *Acta Crystallographica Section D: Biological Crystallography*. 2002; 58(11):1897–900. <https://doi.org/10.1107/s0907444902015160> PMID: 12393919
72. Pandurangan AP, Stahlhacke J, Oates ME, Smithers B, Gough J. The SUPERFAMILY 2.0 database: a significant proteome update and a new webserver. *Nucleic acids research*. 2019; 47(D1):D490–D4. <https://doi.org/10.1093/nar/gky1130> PMID: 30445555
73. Käll L, Krogh A, Sonnhammer EL. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. *Nucleic acids research*. 2007; 35(suppl\_2):W429–W32. <https://doi.org/10.1093/nar/gkm256> PMID: 17483518
74. Hirokawa T, Boon-Chieng S, Mitaku S. SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics (Oxford, England)*. 1998; 14(4):378–9. <https://doi.org/10.1093/bioinformatics/14.4.378> PMID: 9632836
75. Dobson L, Reményi I, Tusnády GE. CCTOP: a Consensus Constrained TOPology prediction web server. *Nucleic acids research*. 2015; 43(W1):W408–W12. <https://doi.org/10.1093/nar/gkv451> PMID: 25943549
76. Armenteros JJA, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature biotechnology*. 2019; 37(4):420–3. <https://doi.org/10.1038/s41587-019-0036-z> PMID: 30778233
77. Pawłowski K, Lepistö M, Meinander N, Sivars U, Varga M, Wieslander E. Novel conserved hydrolase domain in the CLCA family of alleged calcium-activated chloride channels. *Proteins: Structure, Function, and Bioinformatics*. 2006; 63(3):424–39. <https://doi.org/10.1002/prot.20887> PMID: 16470849
78. Bothe MK, Mundhenk L, Kaup M, Weise C, Gruber AD. The murine goblet cell protein mCLCA3 is a zinc-dependent metalloprotease with autoprolytic activity. *Molecules and cells*. 2011; 32(6):535–41. <https://doi.org/10.1007/s10059-011-0158-8> PMID: 22080371
79. Olias P, Adam I, Meyer A, Scharff C, Gruber AD. Reference genes for quantitative gene expression studies in multiple avian species. *PloS one*. 2014; 9(6):e99678. <https://doi.org/10.1371/journal.pone.0099678> PMID: 24926893
80. Bothe MK, Mundhenk L, Beck CL, Kaup M, Gruber AD. Impaired autoprolytic 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*. 2012; 33(3):251–7. <https://doi.org/10.1007/s10059-012-2217-1> PMID: 22350745
81. Plog S, Mundhenk L, Bothe MK, Klymiuk N, Gruber AD. Tissue and cellular expression patterns of porcine CFTR: similarities to and differences from human CFTR. *Journal of Histochemistry & Cytochemistry*. 2010; 58(9):785–97.
82. Sayers EW, Agarwala R, Bolton EE, Brister JR, Canese K, Clark K, et al. Database resources of the national center for biotechnology information. *Nucleic acids research*. 2019; 47(Database issue):D23. <https://doi.org/10.1093/nar/gky1069> PMID: 30395293

83. Yurtsever Z, Sala-Rabanal M, Randolph DT, Scheaffer SM, Roswit WT, Alevy YG, et al. Self-cleavage of human CLCA1 protein by a novel internal metalloprotease domain controls calcium-activated chloride channel activation. *Journal of Biological Chemistry*. 2012; 287(50):42138–49. <https://doi.org/10.1074/jbc.M112.410282> PMID: 23112050
84. Pult F, Fallah G, Braam U, Detro-Dassen S, Niculescu C, Laube B, et al. Robust post-translocational N-glycosylation at the extreme C-terminus of membrane and secreted proteins in *Xenopus laevis* oocytes and HEK293 cells. *Glycobiology*. 2011; 21(9):1147–60. <https://doi.org/10.1093/glycob/cwr013> PMID: 21303814
85. Coudrier E, Arpin M, Dudouet B, Finidori J, Garcia A, Huet C, et al. Villin as a structural marker to study the assembly of the brush border. *Protoplasma*. 1988; 145(2–3):99–105.
86. Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO. The global diversity of birds in space and time. *Nature*. 2012; 491(7424):444–8. <https://doi.org/10.1038/nature11631> PMID: 23123857
87. Lee RM, Ryu RH, Jeong SW, Oh SJ, Huang H, Han JS, et al. Isolation and Expression Profile of the Ca-Activated Chloride Channel-like Membrane Protein 6 Gene in *Xenopus laevis*. *Laboratory animal research*. 2011; 27(2):109–16. <https://doi.org/10.5625/lar.2011.27.2.109> PMID: 21826170
88. Taylor JS, Raes J. Duplication and divergence: the evolution of new genes and old ideas. *Annu Rev Genet*. 2004; 38:615–43. <https://doi.org/10.1146/annurev.genet.38.072902.092831> PMID: 15568988
89. Ouedraogo M, Bettembourg C, Bretaudeau A, Sallou O, Diot C, Demeure O, et al. The duplicated genes database: identification and functional annotation of co-localised duplicated genes across genomes. *PLoS one*. 2012; 7(11):e50653. <https://doi.org/10.1371/journal.pone.0050653> PMID: 23209799
90. Conant GC, Wolfe KH. Turning a hobby into a job: how duplicated genes find new functions. *Nature Reviews Genetics*. 2008; 9(12):938–50. <https://doi.org/10.1038/nrg2482> PMID: 19015656