The Claudinome: Interactome analysis of the human claudin protein family

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry, Pharmacy
of Freie Universität Berlin

by

Lorena Suárez Artiles

2021

Time period: 2016-2021,
Supervisors: Prof. Dr. Dominik Müller, Prof. Dr. Gunnar Dittmar
Institute of doctoral studies: Max Delbrück Center for Molecular Medicine

1st Reviewer: Prof. Dr. Dominik Müller
2nd Reviewer: Prof. Dr. Silke Rickert-Sperling

Date of defense: 06.10.2021

Table of Contents

Declaration of independence

I, Lorena Suárez Artiles, hereby declare that I prepared the submitted dissertation independently and without the support of third parties and that I did not use any unnamed sources or aid.
All parts based on the publications or presentations of other authors are specified as such according to the citing guidelines. Furthermore, I declare that I have marked all texts or parts of texts generated in collaboration with other persons.
I declare that I have not submitted this dissertation or parts of it to another Faculty.
In addition, I declare that the Doctorate Regulations for the Department of Biology, Chemistry, and Pharmacy at Freie Universität Berlin from 31st of May 2018 are known to me. I also declare that I have no cooperation with commercial doctoral researchers and that I shall comply with the regulations of Freie Universität Berlin on ensuring good scientific practice.

Berlin, 27.05.2021

Summary

Claudins constitute a family of tetraspan transmembrane proteins that are ubiquitously expressed in epithelial tissues. They are the major components of tight junctions (TJs) defining the barrier properties in epithelia and regulating the paracellular transport of water and solutes. Claudins can be also found outside the TJ along the basolateral plasma membrane of epithelial cells. It is not yet well understood how claudins regulate the formation of TJs and which functions they exert outside them. The role of their extracellular loops in the regulation of paracellular transport has been widely studied. In contrast, it is not clear how their long and unstructured intracellular C-terminal regions intervene in their functions beyond interacting with tight junction protein ZO-1 (TJP1). Although protein interaction studies can help to answer these two key questions, the disordered nature of the cytosolic tail of claudins makes it challenging. In this work, a large-scale study is presented combining two complementary pull-down techniques followed by mass spectrometry to create an interaction landscape of the claudin protein family.
Co-immunoprecipitation (CoIP) of recombinant claudins overexpressed in MDCK-C7 epithelial cells provided information about interactions beyond the already known TJ proteins. Protein interaction screen on a peptide matrix (PRISMA) allowed for the mapping of interactions along the disordered cytosolic C-terminal region of claudins and enables to study the effect of numerous post-translational modifications (PTMs) in these interactions. We confirmed newly identified claudin tail interactors from our PRISMA approach by proximity ligation assays (PLA) and revealed their possible implication in spatially separated biological processes in epithelial cells. By combining these two complementary approaches we were able to create the Claudinome, a first comprehensive map of interactors for the entire human claudin protein family, which constitutes a valuable resource to improve our understanding of functional connections and regulatory processes for all claudins

Zusammenfassung

Claudine sind eine Familie von Tetraspanin-Transmembranproteinen, die ubiquitär in Epithelgewebe exprimiert werden. Sie sind die Hauptbestandteile von Tight Junctions (TJs), die die Barriereeigenschaften im Epithel bestimmen und den parazellulären Transport von Wasser und gelösten Stoffen regulieren. Außerhalb von TJs kommen Claudine auch entlang der Basalmembran der Epithelzellen vor. Es ist noch wenig darüber bekannt, wie Claudine die Bildung von TJs regulieren und welche weiteren Funktionen sie darüber hinaus erfüllen.

Die Rolle ihrer extrazellulären Schleifen in der Regulation des parazellulären Transports wurde bereits gründlich untersucht, es ist jedoch noch nicht geklärt, welche Funktion die langen, unstrukturierten intrazellulären C-terminalen Regionen über die Interaktion mit dem Tight Junction Protein ZO-1 (TJP1) hinaus haben. Proteininteraktionsstudien können helfen, diese beiden Schlüsselfragen zu beantworten, werden aber durch den variablen c-terminalen Bereich im Zytosol erschwert. In dieser Arbeit wird eine umfangreiche Studie vorgestellt, die zwei komplementäre Pull-down Methoden mit anschließenden massenspektrometrischen Analysen kombiniert, um die Interaktionslandschaft der Proteine der Claudin-Familie darzustellen. Co-Immunpräzipitation (CoIP) rekombinanter Claudine, die on MDCK-C7 Epithelzellen überexprimiert wurden, lieferte Informationen, die über die bereits bekannten TJ Proteine hinausgingen. Ein Proteininteraktionsscreening auf einer Peptidmatrix (protein interaction screen on a peptide matrix, PRISMA) ermöglichte es, die Interaktionen entlang der c-terminalen Region der Claudine und den Einfluss verschiedener post-translationaler Modifikationen (PTMs) zu erforschen. Die Neuentdeckung von Proteinkomplexen, die mit dem zytosolischen Ende verschiedener Claudine interagieren, mithilfe von PRISMA wurde mit Proximity-Ligation-Assays (PLA) bestätigt und offenbart eine mögliche Beteiligung an biologischen Prozessen in Epithelzellen. Durch die Kombination dieser beiden komplementären Herangehensweisen konnten wir das „Claudinom“ erstellen, eine umfangreiche Interaktionslandkarte der Claudin Proteinfamilie, die das gegenwärtige Wissen in diesem Feld erweitert.

1. Introduction

1.1. Epithelia and intercellular junctions

Epithelial cells are attached to each other by a complex of intercellular junctions. The three major types of cell junctions are the tight junctions (TJs), adherens junctions (AJs), and gap junctions. TJs are the most apical cell-cell junctions and they regulate the transepithelial paracellular transport of water and solutes. Tight junctions also restrict the lateral diffusion of membrane proteins maintaining the polarization in epithelial cells. These structures can also be found in endothelial cells specifically those forming the blood-brain barrier, as well as in some tissues without typically polarized cells such as cardiac muscle cells and myelinated neurons (Hagen, 2017).
Tight junctions are multiprotein complexes consisting of a multitude of integral membrane proteins and associated cytoplasmic proteins. The transmembrane proteins fall into three groups: the single transmembrane domain proteins including JAM, Crb3, and CAR; the triple transmembrane domain protein, Bves; and the four transmembrane domain proteins of the claudin, and TAMP families, which include occludin, tricellulin and MARVEL D3. Of these, claudins are the most important protein family and major determinants of paracellular permeability (Günzel and Yu, 2013).

1.1.1. Claudins inside and outside the tight junction.

Claudins are the main structural components of tight junctions. In fact, overexpression of certain claudins in non-tight junction forming cell lines was sufficient to induce the formation of tight junctions (Furuse et al., 1998; Kubota et al., 1999). They are ubiquitously expressed in all epithelial tissues, while only a subset of claudins is simultaneously expressed in the same tissue conferring the particular permeability properties of it. Some claudins form charge- and size-selective pores that regulate the paracellular movement of ions and solutes through epithelial cell layers, whereas other claudin isoforms constitute barriers that block the diffusion of any solutes or water.
Claudins can interact with each other within the same plasma membrane (*cis*) and with claudins in the plasma membrane of adjacent cells (*trans*). These interactions are usually between claudins of the same subtype (homomeric or homotypic) but can

be also between different claudin subtypes (heteromeric or heterotypic) although less common (Coyne et al., 2003; Daugherty et al., 2007; Milatz et al., 2017).

Apart from their central role in the regulation of cell polarity and transepithelial paracellular transport, claudins seem to be also important outside the tight junctions. In her 2017 review, Susan Hagen summarized the evidence suggesting that claudins can be also localized outside of the tight junction complex where they serve important functions outside the regulation of ion permselectivity and paracellular transport. They propose naming these extra-tight junction claudins bCLDN regarding their localization to the basolateral membrane and cnCLDN for those that shuttle between the cytoplasm and the nucleus. Many studies show that there is a pool of basolateral membrane claudins present in epithelial cells of different organs of different species as well as in cultured cells that could represent a pool of claudin molecules available to recycle to the apical tight junction when needed. Other possible explanations are that they may function as an accessory structure to the tight junction for additional permeability barrier; that claudins can freely diffuse along the basolateral membrane and are trapped into tight junction fibrils as they oligomerize within the apical tight junction complex; or that they function as a signaling hub or complex along the basolateral membrane. It has been also described that some claudins (claudin-1, -2, and -7) can be part of the focal adhesion complexes that interact with the extracellular cell matrix (ECM) in the basal membrane of epithelial cells (Hagen, 2017).

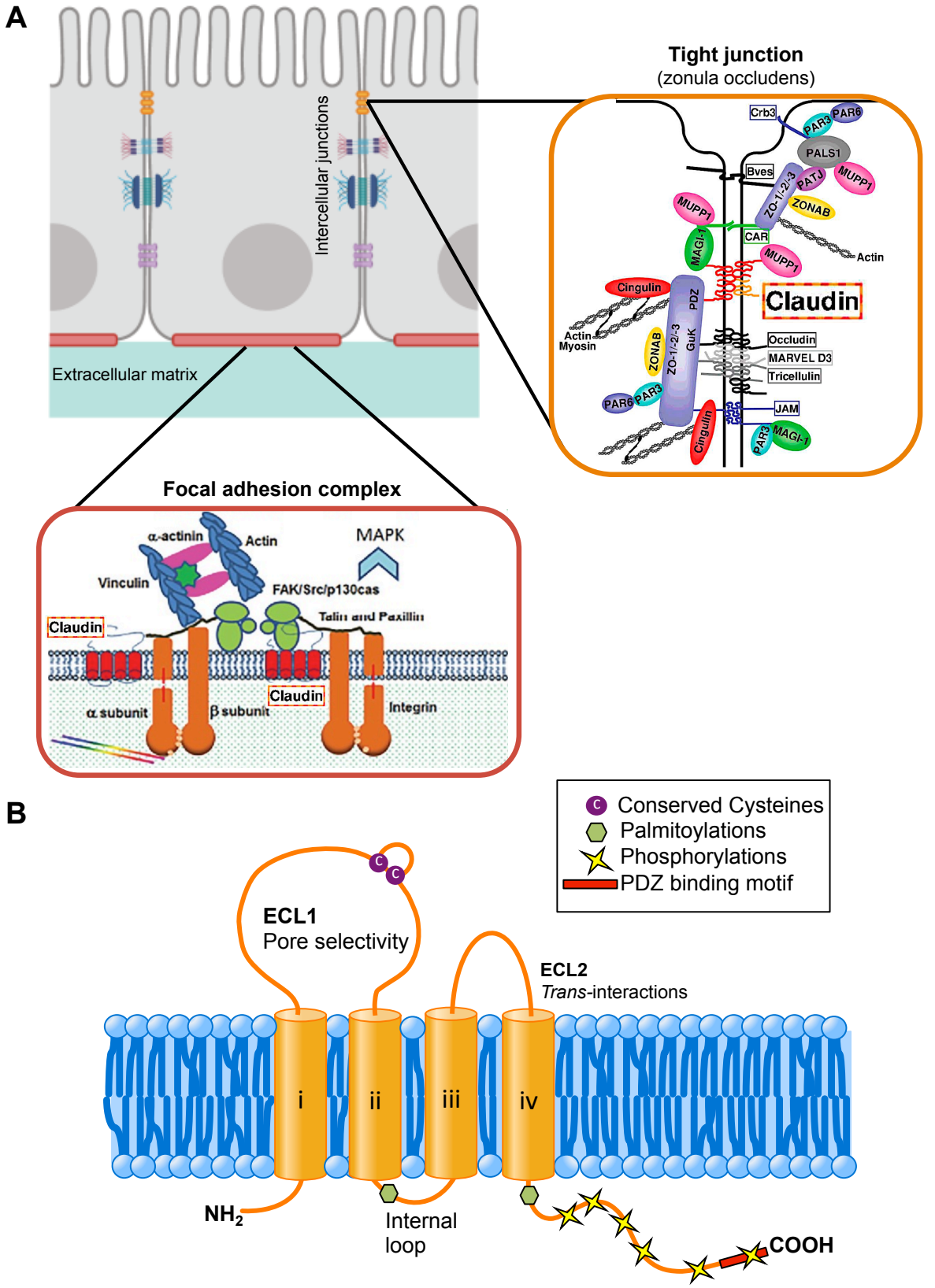


Figure 1 The Claudin family. A. Claudins are the main component of tight junctions and regulate the paracellular transport in epithelia. Claudins also localize to the basolateral plasma membrane outside of the tight junction complex, for example interacting with integrins within the focal adhesion complex. Modified from Günzel and Fromm, 2012; Hagen, 2017. **B.** Claudins are tetraspan transmembrane proteins with 2 extracellular loops and 1 intracellular loop and both N- and C-termini are intracellular. The C-terminal long unstructured regions of claudins contain a PDZ domain-binding motif and manyPTM sites.

1.1.2. Claudins and disease

Mutations in claudin genes are the cause for various rare inherited disorders summarized in Table 1. Additionally, polymorphisms in some claudins are associated with polygenic diseases including claudin-1 with atopic dermatitis and small vessel vascular dementia (De Benedetto et al., 2011; Srinivasan et al., 2017), claudin-5 with schizophrenia (Sun et al., 2004), and claudin-14 with calcium nephrolithiasis (Thorleifsson et al., 2009).

Abnormal claudin expression, regulation, or localization is often observed in intestinal inflammatory disorders such as Crohn's disease and ulcerative colitis (Oshima et al., 2008; Zeissig et al., 2007), and in epithelial cancers. Downregulation of claudins and other tight junction proteins is one of the hallmarks of epithelial-to-mesenchymal transition (EMT) in normal cells. EMT is also an essential step of cancer progression leading to increased migratory ability resulting in invasion and metastasis. Increasing evidence of aberrant claudin expression and localization in several cancers demonstrate their role in EMT regulation and metastatic progression (Kyuno et al., 2021), making claudins relevant for the development of diagnostic tools as well as therapeutic targets in cancer.

Lastly, claudins also play a role in infectious diseases. Claudin-3 and -4 expressed in intestinal epithelial cells are receptors for *Clostridium perfringens* enterotoxin (CPE) responsible for the symptoms of several *C. perfringens* associated gastrointestinal diseases (Katahira et al., 1997; Morita et al., 1999). In the liver, claudin-1, -6, and -9 are co-receptors for hepatitis C virus entry in hepatocytes (Evans et al., 2007; Zheng et al., 2007).

Gene	Phenotype	MIM#	References
CLDN1	Neonatal ichthyosis-sclerosing cholangitis (NISCH) syndrome, or ichthyosis, leukocyte vacuoles, alopecia, and sclerosing cholangitis (ILVASC)	607626	Hadj-Rabia et al., 2004 Feldmeyer et al., 2006
CLDN10	Hypohidrosis, electrolyte imbalance, lacrimal gland dysfunction, ichthyosis, and xerostomia (HELIX syndrome)	617671	Klar et al., 2017 Hadj-Rabia et al., 2018
CLDN14	Autosomal recessive nonsyndromic deafness-29 (DFNB29)	614035	Wilcox et al., 2001 Lee et al., 2012
CLDN16	Familial hypomagnesemia with hypercalciuria and nephrocalcinosis (FHHNC) or renal hypomagnesemia-3, (HOMG3)	248250	Simon et al., 1999; Weber et al., 2000
CLDN19	Familial hypomagnesemia with hypercalciuria, nephrocalcinosis, (FHHNC) and severe ocular involvement or renal hypomagnesemia-5 with ocular involvement (HOMG5)	248190	Konrad et al., 2006

Table 1. Inherited mendelian disorders caused by mutations in claudins. Form the Online Mendelian Inheritane in Man® (OMIM®) database.

1.1.3. The claudin protein family

The claudin family has 27 genes in mammals and at least 23 in humans. They encode proteins with a size range of 20 to 35 kDa characterized by their four helical transmembrane domains, with two extracellular loops, one short intracellular loop, and both N- and C-termini are intracellular. The N-terminal regions are very short with few exceptions like human claudin-16 (73 amino acids). The first extracellular loop (ECL1) is large and in charge of defining the pore selectivity for the paracellular transport by a series of specific charged amino acids that act as binding sites for permeating ions (Günzel and Yu, 2013). It also contains the highly conserved signature sequence of claudins [GN]-L-W-x(2)-C-x(7,9)-[STDENQH]-C (PROSITE ID PS01346) (Van Itallie and Anderson, 2006). The second extracellular loop (ECL2) is smaller than ECL1 and analysis with claudin-5 mutants showed their importance in *trans*-interactions with other claudin monomers in neighboring cells (Piontek et al., 2008). The long, unstructured, intracellular C-terminal region of claudins varies between the 25 and 111 amino acids and its sequence is less conserved among the different isoforms compared to other regions in the proteins (**Figure 2**). It contains the PDZ domain-binding motif (–YV) necessary for the interaction with the PDZ domains of tight junction associated proteins as well as several post-translational modification (PTM) sites.

1.1.4. Importance of the cytosolic tail of claudins

Apart from the interaction with tight junction-associated PDZ domain proteins, the cytosolic C-terminal tail is also implicated in claudin targeting to the plasma membrane and regulation of the turnover and degradation of the different isoforms. Although the mechanisms behind these processes are not well understood yet, several pieces of evidence in the literature show a key role of the C-terminal tail. Truncation of the cytoplasmic tail in several claudin isoforms affects trafficking to the tight junction leading to accumulation in the endoplasmic reticulum followed by proteasomal degradation (Günzel and Yu, 2013). Using domain-swapping chimeras it has also been shown that the half-life of different claudin isoforms is determined by their C-terminal intracellular tails (Van Itallie et al., 2004). The large number of described and predicted phosphorylation sites present in the cytosolic tail of claudins

also suggest that these processes can be highly regulated by post-translational modifications.

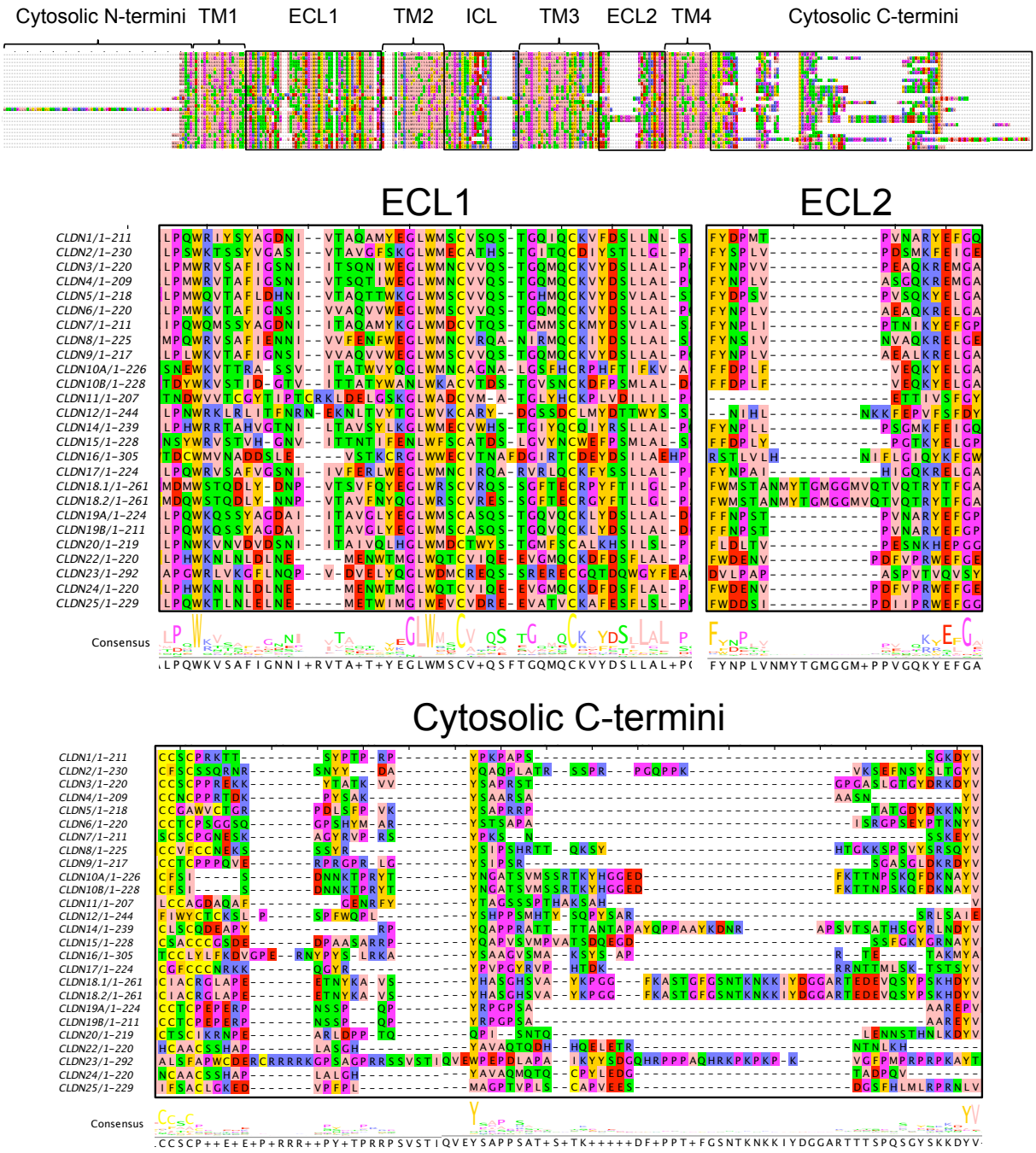


Figure 2. Sequence alignment of all members of the claudin family expressed in humans. The cytosolic C-termini of claudins show a higher variability in lenght and in sequence compared to more conserved regions like transmembrane domains (TM1-TM4) or extracellular loops (ECL1 and ECL2).
TM: transmembrane. ECL: extracellular loop. ICL: intracellular loop.

1.2. Disorganized protein interactions

Protein folding and three-dimensional structure are inherently connected to function. However, a significant fraction of the eukaryotic proteome is either fully disordered or contains disordered regions (intrinsically disordered proteins, IDPs; or intrinsically disordered regions, IDRs) (Wright and Dyson, 1999). The structural disorder is characterized by low sequence complexity and a biased amino acid composition with high enrichment in polar and charged residues and a low proportion of hydrophobic residues, which allows their identification by disorder prediction bioinformatics tools. According to the current estimations, around 15% of all human proteins are fully disordered, and another 35% contain disordered regions of at least 30 amino acids (Ward et al., 2004). In transmembrane proteins, large majority of IDRs are localized to the cytoplasmic side (Bürgi et al., 2016). In addition, IDRs are more frequent in the N- or C- termini than in loop regions of transmembrane proteins (Minezaki et al., 2007; Tusnády et al., 2015). IDRs in terminal regions of transmembrane proteins are often involved in protein-protein interactions, signal transduction across membranes and scaffolding of signaling hub, regulation of vesicle trafficking, and modulation of cell membranes. How IDRs bind and modulate membranes has been extensively described recently in a review by Cornish et al., 2020.

Intrinsically disordered regions present a flexible structure that allows a range of conformations facilitating promiscuous interactions with different target molecules. Protein regions that undergo disorder-to-order transition upon binding events are called molecular recognition features (MoRFs) and are usually around 20 amino acids long. Shorter disordered binding regions consisting of 3-12 residues are called short linear motifs (SLiMs). Recent estimations based on SLiM and MoRF prediction tools calculate that the human proteome may contain more than 100.000 short linear binding motifs located within IDRs (Tompa et al., 2014).

SLiM-mediated interactions are characterized by a low to moderate binding affinity with high dissociation rates and high specificity determined not only by the motif itself but the flanking regions and local sequence context (Ivarsson and Jemth, 2019; Stein and Aloy, 2008). Additionally, IDRs typically contain multiple interaction motifs that mediate binding to multiple targets (Dunker et al., 2005) and often contain sites for post-translational modifications (PTMs) that alter the functionality of motifs (Iakoucheva et al., 2004). This combination of properties allows a dynamic yet highly

controlled regulation of cellular processes such as cell-cell adhesion where the p120 catenin binds the intrinsically disordered cytoplasmic tail of E-cadherin that contains a triple glycine motif as well as a dileucine (LL) motif and tyrosine phosphorylation sites that are also recognized by different endocytosis mediators (Ishiyama et al., 2010).

1.3. Protein-protein interactions studies and mass spectrometry

Proteins frequently carry out their functions by interacting with other proteins; therefore protein-protein interaction (PPI) studies are essential to understand their biological roles. Many different methods have been developed to study various aspects of PPIs. Methods like immunoprecipitation (IP) combined with western blot, crystallography, nuclear magnetic resonance (NMR), or fluorescence resonance energy transfer (FRET) require prior knowledge about the interacting partners and are limited by low throughput. There are also several high throughput methods for the discovery of new PPIs such as yeast-two hybrid and phage display but they lack the biological context (Meyer and Selbach, 2015). In the last decades, mass spectrometry-based proteomics has developed as an important tool for the study of protein-protein interactions.

1.3.1. Mass spectrometry-based proteomics

Mass spectrometry (MS) is an analytical technique that is used to measure the mass-to-charge ratio (m/z) of ions. These measurements are carried out in a mass spectrometer, which, by definition, consists of an ion source, a mass analyzer that measures the m/z of the ionized analytes, and a detector that registers the number of ions at each m/z value (Aebersold and Mann, 2003). For the analysis of biological samples, the most common methods for analyte ionization are electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI), with ESI having the advantage that it can be directly combined with separation techniques such as high-performance liquid chromatography (HPLC) (Fenn et al., 1989; Karas and Hillenkamp, 1988). MS techniques where liquid chromatography as a separation method is combined with two or more steps of m/z analysis of precursor and dissociated fragment ions are also known as liquid chromatography-coupled tandem mass spectrometry (LC-MS/MS).

In shotgun proteomics (also known as bottom-up proteomics), proteins are first digested into peptides using a protease, most commonly trypsin. The resulting peptides are then separated by chromatography (HPLC), ionized by ESI, and transferred directly into a mass spectrometer. The instrument measures the m/z and intensity of the peptides eluting from the HPLC column (MS1) first. Then, the instrument selects individual peptides (precursor ions) for fragmentation and the resulting fragment spectra (MS2) are measured to determine the amino acid sequence (Aebersold and Mann, 2003). Most LC-MS-MS approaches rely on data-dependent acquisition (DDA), in which the most common Top N precursor ions from an MS1 scan are selected for further fragmentation and acquisition of the MS2 spectra. In a DDA mass spectrometry mode, data generated by the instrument is compared to protein databases using a database search algorithm for peptide and protein identification (Eng et al., 2011).

Although mass spectrometry proteomics was initially a qualitative method, various technologies have been developed to enable protein quantification using mass spectrometry. The incorporation of stable heavy isotopes by chemical labeling or metabolic approaches allows different samples to be mixed and analyzed together. The mass shift introduced by the labeling makes them distinguishable and the changes in peptide intensities reflect differences in the abundance of the proteins under different experimental conditions. Alternatively, it is also possible to determine relative protein abundances in a label-free setup.

Intensity-based label-free quantification (LFQ) allows for proteins to be quantified using computational methods based on all the intensities obtained from precursor peptide scans (Bondarenko et al., 2002). This quantitation method is based on the calculation of the area under the curve or peak height for each peptide that elutes from the LC column at an expected retention time. Therefore, variations in LC separation or stability or the electrospray ion source can affect its reproducibility. However, the robustness and accuracy of LFQ approaches can be improved by using highly reproducible sample handling and analysis protocols, sufficient technical replicates per sample (e.g. quadruplicates), and data normalization algorithms such as MaxLFQ (Cox et al., 2014). Additionally, LFQ allows for the analysis of practically an unlimited number of samples in a time- and cost-effective manner since it requires fewer sample preparation steps and no expensive labeling reagents.

1.3.2. Large-scale PPIs studies

Recent advances in the field of MS-based proteomics resulted in more powerful high-resolution mass spectrometers as well as more accurate data analysis algorithms. Combined with the versatility of label-free approaches and making use of protein enrichment techniques, these advances allow for the development of high throughput methods for the analysis of PPI with short sample processing and measuring times, increased sensitivity, and high precision.

1.3.2.1. AP-MS

Affinity purification followed by mass spectrometry (AP-MS) can determine the members of protein complexes in the cellular context in an unbiased manner (Gavin et al., 2002). This approach relies on the expression of recombinant proteins with an in-frame epitope tag that is used as an affinity handle to purify the tagged protein (bait) along with its interacting partners (the prey). AP-MS allows the detection of functional protein interactions under near-physiological conditions and has been used to map the interactome of several organisms (Keilhauer et al., 2015). Previously, protein complexes had to be extensively purified to reduce the number of unspecific binders. With the development of the technology in mass spectrometry with high-resolution instruments and quantitative approaches, it is possible to distinguish true interactors from contaminants. Therefore, less stringent enrichment protocols can be used to preserve weaker interactions. However, AP-MS still has some technical limitations. Although it can be done with endogenous proteins, it often requires the overexpression of recombinant proteins as well as disruption of the cell compartments during the lysis process, which might lead to the identification of false positive interactions. When using the whole protein as bait AP-MS does not provide information about what region of the protein is mediating the interaction. Also, obtaining information about the effect of PTMs on the identified interactions often requires additional experiments using for example PTM mimicking mutant versions of the bait. In the case of the study of transmembrane proteins, due to their hydrophobicity, the use of detergents to solubilize them is necessary and can impact protein-protein interactions and the composition of protein complexes purified by immunoprecipitation (Lee et al., 2018). This often leads to the loss of weak or transient interactions.

Proximity labeling alternatives such as BioID (Roux et al., 2012) or APEX (Rhee et al., 2013; Hung et al., 2014) have been developed to overcome some of the issues of classical AP-MS methods and have been applied to several membrane proteins. In both cases, the protein of interest is fused to an enzyme that labels proteins in close proximity. The advantage of these techniques is that they produce covalent biotinylations that allow for the recovery of weak or transient interactors. The higher affinity between streptavidin and biotin also allows for stringent washing steps with a significant reduction of the background**.** On the other hand, these techniques require extensive cloning steps and also involve overexpression of the bait. Moreover, the BioID tag is larger in size than GFP and it can have a higher interference with the location, function, or stability of the bait (Roux et al., 2018). It is also important to keep in consideration that the labeling of proteins in proximity does not necessarily mean a direct interaction and it could lead to false positives (Kim et al., 2014).

1.3.2.2. Peptide-protein pull-downs

The study of IDR-mediated PPIs is experimentally challenging due to their transient and reversible nature. AP-MS techniques often fail to detect such weak interactions and, even though proximity labeling approaches can overcome that issue, none of those techniques provide precise information about the SLiM mediating the interaction since any given protein with IDRs potentially contains several SLiMs.
A specific characteristic of IDRs is that they can exert their function independently of the context of the full-length protein; therefore interactions mediated by SLiMs can be studied using short peptides. The development of solid-phase peptide synthesis allows for the biochemical study of peptide-mediated interactions with the advantage that PTMs can also be included in peptides during synthesis. Similar to a traditional AP-MS experiment, proteins interacting with immobilized peptides coupled to beads can be identified by mass spectrometry. Using quantitative proteomics it is possible to distinguish interacting proteins from unspecific background binders (Schulze and Mann, 2004; Selbach et al., 2009). This setup allows for the study of SLiM mediated interactions and the impact of PTMs. However, the number of baits studied is limited to the number of individual pull-down experiments performed, which makes high throughput difficult.

With SPOT synthesis (Frank, 1992) multiple peptides can be synthesized and immobilized on cellulose membranes at high density to use them as baits to pull down interacting proteins. The advantage of this high local peptide concentration is that it facilitates the enrichment of proteins with low binding affinities. This type of peptide array is typically probed with a single prey protein subsequently detected by antibodies or by fluorescent/ radioactive labels on the prey protein of interest (Volkmer et al., 2012). In the peptide array X-linking (PAX) assay, synthetic peptide arrays were incubated in cell lysate followed by cross-linking of the interactors. Peptide spots were then excised and analyzed by mass spectrometry for the unbiased identification of the interacting proteins (Okada et al., 2012)

Protein Interaction Screening on Peptide Matrix (PRISMA) is a recently developed technique that similarly combines the high throughput of SPOT synthesized peptide arrays with the identification by mass spectrometry of all interacting proteins in peptide pull-downs without the need of using crosslinking. It has been used to map binding partners along the sequence and PTMS sites of CEBP transcription factors and to study the effect of disease-causing point mutations on protein interactions mediated by disordered regions (Meyer et al., 2018; Dittmar et al., 2019; Ramberger et al., 2020).

1.4. Aim of this study

The vast majority of the advances made in claudin physiology and biology of different organisms are derived from mRNA studies, antibody-based detection, and functional studies made in cell lines and animal models (Liu et al., 2016). Functional studies are essential to understand the physiological implications that claudin alterations or mutations have in the regulation of paracellular transport and related diseases, as well as in many other relevant biological processes where this protein family is implicated. Nevertheless, these types of studies cannot assess the molecular mechanisms underlying these effects. Several studies based on mRNA levels and direct protein detection using antibody-based techniques have been useful to define the claudin expression profile of different tissues and cell lines both in normal conditions and in many diseases. Although knowing the claudin composition of a specific tissue or cell line and how it changes depending on the circumstances is extremely valuable, many of the processes in which claudins are involved are regulated by protein-protein interactions and post-translational modifications and the above-mentioned approaches fail to provide this type of information. Proteomics and more specifically affinity purification-mass spectrometry (AP-MS) techniques allow the study of protein-protein interactions and can therefore contribute to the understanding of the different functions of the claudins and how they are regulated. Additionally, the increasing knowledge on the importance of IDRs in transmembrane proteins reinforces the idea of a potentially relevant role of the C-terminal cytosolic region of claudins in cell signaling and/or regulation of the dynamic changes in tight junctions. Therefore, this study aims to combine CoIPs and PRISMA to develop a comprehensive large-scale interactome network of the claudin family with a focus on the cytosolic C-terminal tails (**Figure 3**).

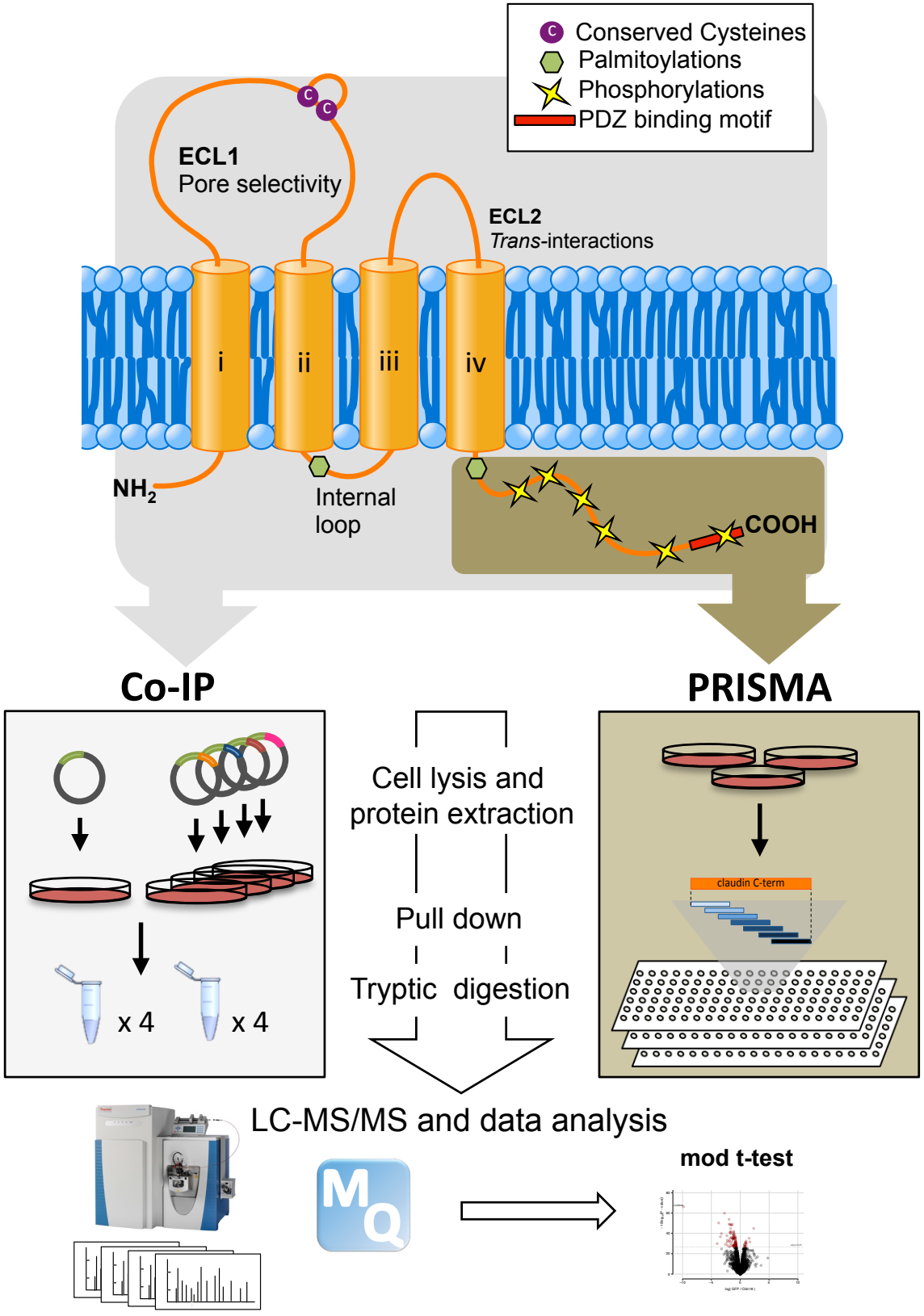


Figure 3. Study of the claudin protein family by pull-down-based techniques combined with mass spectrometry. Co-immunoprecipitation experiments were used to study interactions with the full-length claudin proteins. PRISMA enabled mapping interactions along the unstructured C-terminal cytosolic tail. Samples were subsequently analyzed by LC-MS/MS, and significant interactions defined using moderated t-test.

2. Material and methods

2.1. YFP fused claudin constructs

Plasmids containing cDNAs encoding recombinant human claudins N-terminally fused to either YFP or CFP were kindly provided by Dr. Dorothee Günzel's lab in the Institute of Clinical Physiology (Charité Campus Benjamin Franklin, Berlin). Mouse recombinant claudin isoforms 4 and 9 fused to YFP were cloned by Dr. Tilman Breiderhoff at the Charité Virchow Klinikum. cDNA from mouse lung and stomach tissue was used to clone claudin isoforms 18.1 and 18.2. An aliquot of each construct was submitted to Source BioScience for Sanger sequencing using their stock EGFP C F forward primer (5'-CAT GGT CCT GCT GGA GTT CGT G-3'), to confirm the absence of possible mutations in the different claudin constructs.

2.2. Cell culture

MDCK-C7 cells were cultured in a humidified incubator at 37°C, 5% CO2 in DMEM medium supplemented with 10%FBS and 1% Penicillin/Streptomycin.

2.3. Generation of stable MDCK-C7 cell lines overexpressing claudins

For every claudin isoform-specific cell line, MDCK-C7 cells were seeded at low density (0.125 x10^6 cells/ml) on 6 cm dishes in supplemented DMEM (10%FBS, 1% Penicillin/Streptomycin). The next day, cells were transfected with the YFP-claudin construct using Lipofectamine® 2000 (Invitrogen) following the manufacturer's protocol and leaving one non-transfected dish as a control. First, 6 µg of the DNA construct and 20 µl Lipofectamine® 2000 were separately diluted in 500 µl of serum free DMEM and incubated 5 min at RT. Then, both dilutions were combined in one Eppendorf tube, gently mixed by inverting the tube and incubated another 30 min at RT. The mix was added to the cell dish drop-wise and cells were incubated over night. After incubation, the cell medium was replaced by supplemented DMEM for 24h. Then, cell medium was changed to supplemented DMEM containing G418 (1 mg/ml) for antibiotic selection. Cell medium was replaced every 2 days until the cells on the control plate died (approximately after one week). After antibiotic selection, cells were sorted by FACS to keep those with higher fluorescence signal.

Sorted cells were cultured in supplemented DMEM containing 0.6 mg/ml G418 and 50 µg/ml Gentamicin for at least 12h to avoid contamination, and then kept in supplemented DMEM containing 0.6 mg/ml on a T-25 flask until they were 100% confluent. Stable MDCK C7 cells overexpressing the YFP-tagged claudin isoform were then split into aliquots of approximately $1x10^6$ cells/ml and stored until used for the CoIP experiments.

2.4. Immuno-fluorescence stainings for confocal microscopy

MDCK stable cell lines were grown on Poly-L-lysine coated coverslips in 24-well cell culture plates for a standard immune staining protocol. Briefly, after reaching approximately 90% confluence cells were washed with PBS twice and fixed with 4% PFA. After washing with PBS (3 times, 5 min at RT), cells were permeabilized with 0.1% TritonX-100 in PBS (5 min at RT) and incubated in Blocking buffer (1% BSA 20mM Glycine in PBST (PBS with 0.1% Tween20)) (30 min at RT). Coverslips were then carefully placed on top of parafilm in a humid chamber and incubated with the primary antibody (rabbit anti-ZO-1, 61-7300 Invitrogen) (1:400 in blocking buffer**,** 2h at RT or over night at 4°C) followed by washing steps with PBST to remove the antibody excess (3 times10 min each at RT). The incubation with the secondary antibody was again done in a humid chamber (anti-Rabbit Alexa Fluor® 647)(1:1000 in blocking buffer, 1h at RT) followed by more washing steps with PBST (3 times10 min each at RT). For the counter staining coverslips were incubated with DAPI (0.5 ug/ml, 1 min at RT), washed twice with PBS for 5 min at RT and once with ddH2O, and mounted into a glass slide with a drop of ProLong® Gold Antifade liquid mountant. Samples were stored at 4°C and protected from light.

2.5. Western blotting

As a quality control for the CoIP experiments aliquots from the input, non-bound fraction and output were analyzed by western blotting. For the cell lysates used in PRISMA, immunoblot was used to control the depletion of nuclei in the extracts. In both cases, samples were mixed with 6x loading buffer and boiled for 5 min at 95°C, and then loaded into a 10-12% SDS-polyacrylamide gel and separated by electrophoresis at 100V in running buffer (25mM Tris, 200mM Glycine, 0.1% SDS). Proteins were transferred to a nitrocellulose membrane with the Trans-Blot Turbo

Midi System from Bio-Rad in Towbin running buffer (0.025M Tris, 0.192 M Glycine, pH 8.6, 20% methanol) *(*Towbin et al., 1979), at 100V for 1h. Successful transfer was confirmed by staining the membrane with Ponceau S solution for 5 min. The membrane was then rinsed with TBS-T (50mM Tris-HCl, 150mM NaCl, 0.1% Tween-20) to remove the remaining staining. Free binding sites in the membrane were blocked by incubation in 4% skim milk in TBS-T for 1h at RT and the membrane was then washed twice with TBS-T for 5 min at RT. For detection of actin, GFP/YFP/CFP expression for CoIPs, and histone H3 in the PRISMA cell extracts, membranes were incubated with the corresponding primary antibodies diluted in blocking solution overnight at 4°C (actin: ab179467, histone H3: ab1791, Abcam; GFP: Af1180, guinea-pig anti-GFP, Frontier Institute;). Membranes were then washed 3 times for 5 min in TBS-T and incubated 1h at RT with an HRP-coupled secondary antibody raised against the species of the primary antibody. Membranes were washed 3x for 5 min in TBS-T, immersed in chemiluminescence reaction solution (Pierce™ ECL, Thermo Scientific) for 1 min. Chemiluminescence signal was detected by exposure to an autoradiography film in a dark room and developed with an auto processor. As expected, the highest GFP signal was detected in the output sample as a thick band with an approximate size between 49 and 59 kDa (depending on the claudin) or 27 kDa for the control cell line, indicating a successful immunoprecipitation of the recombinant claudin or the cytosolic GFP with the GFP-Trap®_A nanobodies (Chromotek). In the MDCK C7 cell extracts used for PRISMA, a stronger band of approximately 17 kDa corresponding to histone H3 was detected in the precipitated nuclei sample a faint band of the same size was detected in the PNS samples indicating an effective depletion of nuclear content in the cell lysates (Data not shown).

2.6. Co-immunoprecipitation (CoIP) approach

2.6.1. Cell lysate preparation for CoIP experiments

MDCK-C7 stable cell lines overexpressing YFP/CFP fused claudin isoforms or cytosolic eGFP were grown in 15 cm cell culture dishes in quadruplicates. Once they reached approximately 90% confluence they were ready to be harvested for cell lysis. After removing cell medium, 15 cm dishes were placed on ice and washed twice with ice cold PBS. Using a cell scraper, cells were gently detached from the

culture dish, resuspended in PBS and collected into a pre-chilled 50 ml falcon tube and spun down at 500x g for 5 min at 4ºC, after removing the supernatant cell pellets were collected, snap frozen, and stored at -80°C until the rest of the cell lines were grown and collected following the same procedure. For the cell lysate preparation, cell pellets were thaw on ice, resuspended in 1 ml cold PBS and transferred to a pre-chilled 1.5 ml LoBind Eppendorf were cells were pelleted and resuspended in 400 µl of lysis buffer (150 mM NaCl, 50 mM Tris pH 7.5, 0.5 mM EDTA, 1% IGEPAL-CA-630, 5% Glycerol, protease and phosphatase inhibitors, and 1 µl/ml benzonase), passed through a 23G syringe needle and incubated on ice for 30 min. After the incubation, tubes were centrifuged at 18.000x G, for 10 min at 4°C to remove cell debris. The supernatant cell lysate was transferred to fresh pre-chilled tubes and kept on ice. A small aliquot was taken from each tube to estimate protein concentration by BCA (Pierce™ BCA Protein Assay Kit, ThermoFisher Scientific) following the manufacturer's protocol.

2.6.2. Pull down using GFP-Trap® nanobodies

CoIP experiments were done adapting the protocol from Hubner et al., 2010 to the GFP-Trap®_A nanobodies (Chromotek) manufacturers recommendations. After determination of the protein concentration, the volume equivalent to 1 mg of cell lysate was taken and brought up to 1 ml using dilution/wash (D/W) buffer (150 mM NaCl, 50 mM Tris pH 7.5, 0.5 mM EDTA and 5% Glycerol) to obtain a cell lysate concentration of 1 mg/ml and a detergent final concentration lower than 0.4%. 50 µl of the diluted cell lysate were taken for further immunoblot analysis. In order to condition the nanobodies for the immunoprecipitation, 25 µl of GFP-Trap®_A bead slurry were resuspended in ice-cold D/W buffer and spun down at 2.500xg for 4 min at 4°C, this step was done three times to ensure that the beads are properly washed. After removing the D/W buffer from the last washing step, the cell lysate was added to the equilibrated beads and incubated over night at 4°C (in a cold room) under constant mixing on a rotator. After incubation, tubes were spun down at 2500 xg for 4 min at 4°C, 50 µl of the supernatant was taken for immunoblot analysis and the rest was removed. Beads were then washed as in previous steps but first with 500 µl of D/W buffer + 0.05% IGEPAL-CA-630, then with 500 µl of D/W buffer without detergents and last with 500 µl cold PBS. After centrifugation and removing the PBS

from the last washing step, beads were snap frozen and stored at -80°C until further on-bead protein digestion.

2.6.3. On-bead protein digestion

Frozen beads containing proteins from the CoIP experiments were thawed and incubated in 80µl urea/trypsin buffer (2M urea, 50 mM Tris pH 7.5, 1 mM DTT and 5 µg/ml Trypsin) on a shaker (1h at 25°C, 1000 rpm). 80 µl of the supernatant were transferred to a new tube and beads were washed 2 more times with 60 µl of urea buffer (2M urea, 50 mM Tris pH 7.5). The on-bead digest and washes were combined in one tube with a total volume of 200, spun down at 5000 xg for 1 min to remove the leftover beads, and transfer to a fresh tube. Eluted proteins were reduced by adding 4 mM DTT (30 min incubation on a shaker at 25°C, 1000 rpm), and subsequently alkylated by adding 10 mM IAA (45 min incubation on a shaker at 25°C, 1000 rpm, protected from light). Protein digestion was done adding 0.5 µg of trypsin and incubating overnight (25°C on a shaker, 700 rpm).

2.6.4. Peptide clean-up

After overnight digestion with trypsin, samples were acidified adding 1% FA to reach a pH < 3 for further C18 STAGE tips (STop and Go Extraction tips) desalting (Rappsilber et al., 2003). Stage tips were prepared by packing to disks of Empore 3M C18 material into 200 µl pipette tips. Stage tips were placed in a centrifuge on 2 ml tubes using Glygen adaptors, then washed and equilibrated by sequentially passing trough 100 µl of MeOH (2 times), 100 µl of 50% MeCN/0.1% FA, and 100 µl of 0.1% FA (2 times); for each step stage tips were centrifuged at 3000 xg for 3 min, and 2ml tubes were exchanged after collecting 300 µl. After the washing and equilibrating steps, acidified digests were loaded into the stage tips with the same centrifugation conditions. At this step, acidified peptides were bound to the C18 material from. Stage tips were then washed twice with 100 µl 0.1% FA to remove the remaining salts from the digestion. Desalted peptides were eluted into fresh tubes with 60 µl of 50% MeCN/0.1% FA, and eluates were transferred to a 96-well measuring plate. Samples were snap frozen, lyophilized in a speedvac and stored at - 80°C until measuring by LC-MS/MS.

2.6.5. LC-MS/MS

Dried, desalted peptides were reconstituted in 8 µl of MS sample buffer (3% MeCN/0.1% FA.) and separated online with an Easy-nLCTM 1200 coupled to a Q-Exactive HF-X mass spectrometer equipped with an orbitrap electrospray ion source (Thermo Fisher Scientific). Samples were separated on a 20cm reverse-phase column packed in house with 3 µm C18-Reprosil beads (inner diameter 75µm) with a gradient ramping from 2% to 54% ACN in 35 min, followed by a plateau at 72% ACN for 10 min and a subsequent plateau at 45% ACN for 5 min. MS data was acquired on a Q-Exactive HFX in data dependent acquisition (DDA) with a top20 method. Full can MS spectra were acquired at a resolution of 60000 in the scan range from 350 to 1700 m/z, automated gain control (AGC) target was set to 3e6 and maximum injection time (IT) to 10 ms. MS/MS spectra were acquired at a resolution of 15000, AGC target of 1e5 and maximum IT of 86 ms. Ions were isolated with a 1.3 m/z isolation window and normalized collision energy (NCE) was set to 26. Unassigned charge states and ions with a charge state of one, seven or higher were excluded from fragmentation and dynamic exclusion was set to 20s.

2.7. Protein Interaction Screen on a peptide Matrix (PRISMA) approach

2.7.1. Cell lysate preparation for PRISMA experiments

The goal of PRISMA experiments is to identify cytosolic proteins interacting with the intracellular tail of claudins. Therefore, we selected lysis conditions that enrich for cytosolic content using a modified version of the Schreiber et al., 1989 nuclear-cytoplasmic fractionation protocol. First, confluent cultured MDCK-C7 cells were washed with cold PBS twice and incubated in trypsine for 1h. After cells were completely detached from the surface, they were harvested in a 15 ml tube and spun down at 600 xg, for 10 min at 4°C. Pelleted cells were then resuspended in 5 volumes of hypotonic Buffer A (5mM HEPES pH 8, 0.75mM $MgCl_2$, 5mM KCl, and protease inhibitor cocktail (Roche), supplemented with fresh 1mM DTT) and swollen on ice for 15 min. After the incubation, they are again spun down and resuspended in Buffer A supplemented with 1mM DTT and 0.5% DDM. Cells were passed through a 23G needle and incubated 10 more min. Tubes were then centrifuged at 600xg for 5 min at 4°C to precipitate the nuclei and the post-nuclear supernatant was transferred to a fresh tube. Two small aliquots were taken for estimation of the

protein concentration and for western blot analysis. The rest was snap frozen and shipped to the Luxembourg Institute of Health (LIH) where they were stored at -80°C until the PRISMA experiments were done.

2.7.2. PRISMA pull-downs

Protein interaction screen on a peptide matrix was performed as described before (Dittmar et al., 2019) with slight adaptations of the protocol. Custom PepSpot cellulose membranes including peptides derived from the C-terminal tails of claudins were purchased from JPT (Berlin, Germany). According to the manufacturer, the synthetic peptides are prepared by SPOT-synthesis and each spot contains approximately 5nmol peptide covalently bound to the cellulose-βalanine-membrane.
The experiment was done using three membranes, each of them containing 166 peptide spots (unmodified and phosphorylated) derived from the C-terminal cytosolic tail of claudins (**Supplementary Table 1**). Membranes were pre-conditioned by incubation in membrane binding buffer (MBB, 5 mM HEPES pH 8, 0.75 mM $MgCl_2$, 5 mM KCl, and 1mM DTT), 45 min at RT in a plastic container in a rotator. Then, membranes were blocked with yeast tRNA (1 mg/ml in MBB), 10 min at RT to minimize unspecific binding to the cellulose membrane. To remove the tRNA excess, membranes were washed 5 times with MBB for 5 min at RT. Membranes were then incubated with the MDCK C7 cell extracts (3.5 mg/ml) for 20 min on ice, each membrane was incubated in a separate sealed bag. Prior to the incubation, a 20µg aliquot was taken from each cell extract tube to use them as input samples. After incubation with the cell extracts, membranes were washed 3 times with MBB, 5 min at RT. Membranes were let to dry on a glass surface and then each spot was manually cut and transferred to a 96-well plate containing 20 µl of denaturation buffer (DB, 6M Urea and 2M Thiourea in HEPES, pH 8). Spots containing the interacting proteins pulled down from the cell lysate were then subjected to in solution digestion.

2.7.3. In solution protein digestion

PRISMA samples were reduced and alkylated by incubation in 5mM final concentration of DTT (30 min at 37°C), followed by incubation in 15 mM final concentration IAA (45 min at RT, protected from light). Samples were then diluted in

50 mM ABC buffer to reduce the urea concentration to 0.8 M and digested overnight with 0.5 µg trypsin at 37°C. After digestion, samples were acidified by adding 20 µl of 10%TFA to stop the reaction. The 96-well plates containing the already digested samples were then stored at -20°C until the desalting and peptide clean-up step.

2.7.4. Peptide clean-up

After in solution digestion, PRISMA samples were desalted using Sep-Pak® C18 96-well plates (Waters™) according to the manufacturer's indications. Each step was followed by centrifugation of the 96-well plates for 1 min, at 1000 rpm, and at RT. First, Sep-Pak® columns were pre-conditioned with 300 µl MeOH, followed by washing with 300 µl 80%, and equilibrated twice with 300 µl 0.1% FA. Samples were then loaded and washed 5 times with 300 µl 0.1% FA. Desalted peptides were then eluted with 200 µl 50% ACN, 0.1% FA into Protein LoBind Eppendorf 96-well plates, snap frozen and lyophilized in a speed vac.

2.7.5. LC-MS/MS

Dried, desalted peptides were reconstituted in 10 µl of MS sample buffer (3% MeCN/0.1% FA.) and separated online with a Dionex UltiMate™ 3000 coupled to a Q-Exactive Plus or Q-Exactive HF mass spectrometer equipped with an orbitrap electrospray ion source (Thermo Fisher Scientific). Samples were separated on a 20cm reverse-phase column packed in house with 3 µm C18-Reprosil beads (inner diameter 75µm) with a gradient ramping from 2% to 54% ACN in 7min, followed by a plateau at 72% ACN for 2.5 min. MS data was acquired on a Q-Exactive HF in DDA with a top10 method. Full scan MS spectra were acquired at a resolution of 60000 in the scan range from 375 to 1500 m/z, AGC target was set to 3e6 and maximum IT to 100 ms. MS/MS spectra were acquired at a resolution of 15000, AGC target of 1e5 and maximum IT of 30 ms. Ions were isolated with a 1.2 m/z isolation window and NCE was set to 28. Unassigned charge states and ions with a charge state of 1, 6-8 or higher were excluded from fragmentation and dynamic exclusion was set to 7s.

2.8. MS data processing with MaxQuant

Raw files were analyzed using the MaxQuant version 1.5.2.8 searching against the *Canis lupus familiaris* UniProt database (2018). Settings were kept as default, Deamidation (NQ) was included as a variable modification, and quantitation was done using label-free quantification (Fast LFQ). 'Match between runs' (MBR) was enabled to boost the number of identifications. The analysis of the CoIP raw data was done individually for each claudin isoform versus the GFP control using only unique peptides for quantification. For the PRISMA data, the search was done against an additional second database containing the C-terminal sequence of all human claudin isoforms to detect the synthetic tryptic peptides coming from the membrane. Input samples and groups of peptides from the same claudin were set to non-consecutive fractions so the MBR algorithm works only with runs within the same fraction.

2.9. Statistical analysis of mass spectrometry data

After the MaxQuant analysis of the raw data the protein groups output files were filtered to remove potential contaminants, reverse hit and proteins identified by site. In the case of the PRISMA dataset, C-terminal claudin peptides identified were also removed. The statistical analysis was done using the R software (R version 3.5.0, RStudio version 1.0.143) and the Proteomics Toolset for Integrative Data Analysis Shiny app (ProTIGY, Broad Institute).

2.9.1. CoIP data

Each claudin isoform dataset was filtered for proteins that were detected in at least 3 of the four replicates of each sample group (Claudin pull-down and eGFP control), and with at least two peptides. LFQ missing values were replaced using a downshift imputation approach (Keilhauer et al., 2015). LFQ intensity values of YFP-Claudin pull-downs were then compared against the GFP control by two-sample moderated t-test. First, we applied the standard significance cut-offs of 5% FDR and $\log_2$ fold change >1 for enrichment against the control. Based on the assumption that the eGFP protein has no specific interactors within the MDCK-C7 cells, a more stringent

second level cut-off was applied at an adjusted p-value that leaves only 5% the interactions identified in the GFP control.

2.9.2. PRISMA data

After the initial filtering of the dataset proteins identified only in the input samples were also removed. The following filtering for valid values and imputation of missing values was done as described for CoIP data and separately for the subset of PRISMA spots belonging to each claudin isoform. For the data analysis, moderated t-test pairwise comparisons were done between the unmodified peptides from each claudin isoform on one hand, and between the unmodified and the phosphorylated versions of the same peptides on the other with a significance cut-off of 5% FDR. LFQ values of the significant interactors identified were normalized by z-score and plotted as a heatmap. This representation was use to manually select those proteins showing an intensity profile corresponding to the interaction with a SLiM, which is high LFQ values across 3 consecutive overlapping peptides with a maximum in the middle for the unmodified peptides. For the study of PTMs we selected those proteins that significantly showed a difference in the binding between the modified and the unmodified version of the same peptide.

2.10. Gene Ontology (GO) annotations and enrichment analysis

The list of significant interactors identified for each claudin was submitted to the online functional annotation tool DAVID. Using the gene names we looked for annotations of these interactors in human since the available data for *Canis lupus familiaris* contains many uncharacterized proteins. 708 out of 758 interactions were successfully annotated. Gene ontology terms related to cellular component (GOCC) were used to systematically classify the interactions identified by CoIP following an approach similar to the one used by (Tan et al., 2020). A hierarchical categorization was done looking first for annotations related to tight junction (GO:0005923, GO: 0061689) followed by adherens junction (GO: 0005913, GO:0005912), cell junction (GO:0005911, GO:0030054), apical plasma membrane (GO:0016324), basolateral plasma membrane (GO:0016323), cytoskeleton (GO:0005856, GO:0015629, GO:0015630, GO:0030863, GO:0045111), endosome/caveola/lysosome (GO:000576, GO:0005770, GO:0010008, GO:0005901,GO:0005764), integral

component of plasma membrane, plasma membrane (GO:0005887), Golgi apparatus/vesicle/exosome (GO:0005794, GO:0000139, GO:0032588, GO:0005793, GO:0033116), endoplasmic reticulum (GO:0005783, GO:0005788, GO:0005789, GO:0030176), mitochondria (GO:0005739, GO:0005743, GO:0005741, GO:0005759), and nucleus (GO:0031965, GO:0005634, GO:0005654). Proteins that didn't contain any of these terms were categorized as "others".
Enrichment analysis of the significant interactions identified by CoIP was done using the Metascape gene annotation and analysis resource (Zhou et al., 2019). A multiple gene list was uploaded with the interacting partners identified for each claudin using *H. sapiens* as species. A custom analysis was done selecting GO and KEGG terms in the Pathway and Structural complex sections for the enrichment analysis.

2.11. Proximity ligation assays

Proximity ligation assays (PLA) performed by Dr. Rossana Girardello at the LIH were done using Duolink® PLA reagents (Merck) following manufacturer's instructions. Briefly, three MDCK-C7 stable cell lines overexpressing claudin-1, claudin-3 and claudin-12 respectively were grown in glass slides until confluent. Similarly to a standard immune-fluorescence sample preparation cells were fixed with 4%PFA and permeabilized with 0.1% TritonX100 in PBS. After this pre-treatment, cells were blocked with Duolink® Blocking Solution for 1h at 37°C and subsequently incubated with the primary antibodies for detection of the YFP-caudin and the chaperonin containing TCP1 complex (CCT/Tric) or proteasome complex subunits respectively for 1h at RT (GFP, ab290, Abcam; TCP-1 ß (CCT2), sc-374152; TCP1 Z (CCT6A), sc-514466; 20S proteasome α3 (PSMA3), sc-166205; Santa Cruz Biotechnology). Slides were washed 3 times for 1 min with Wash Buffer for Fluorescence A (Wash Buffer A) and incubated in PLA probe solution (PLUS and MINUS PLA Probes diluted 1:5 in Duolink® Antibody diluent) for 1h at 37°C. Slides were washed again 3 times for 1 min with Wash Buffer A and incubated in ligation solution (Ligase 1:40 in 1x Ligation Buffer) for 30 min at 37°C. After three more washing steps with Wash Buffer A cells were incubated with amplification solution (Polymerase 1:80 in 1x Amplification Buffer) for 100 min at 37°C. Slides were finally washed with Wash Buffer for Fluorescence B (Wash Buffer B) protected from light, 3 times for 1min at RT, and one time with diluted 0.01x Wash Buffer B at RT. Slides were then mounted

with 3-4µl of Duolink® PLA Mounting Medium with DAPI, sealed with transparent nail polish, and stored until taken to the confocal microscope for imaging. Slides were imaged on a Zeiss LSM 880 confocal microscope and images were acquired with a 63x oil immersion Plan-Apochromat objective 1.4 numerical aperture (Zeiss) and standard filter sets. Image analysis was performed with FIJI (ImageJ) according to published protocols (Gomes et al., 2016; Prado Martins et al., 2018). Briefly, single stack images were split in separate channels. The blue channel was used for nuclei counting while the red channel for PLA signal retrieval. The average PLA signal per cell was obtained by dividing the total number of PLA dots by the number of nuclei in each image. 5 different fields were imaged per slide over 3 independent experiments.

3. Results

3.1. Co-immunoprecipitation (CoIP) experiments

For the identification of claudin interactors by CoIP, a total of 24 MDCK-C7 stable cell lines were generated overexpressing recombinant YFP- or CFP-claudins, as well as a control cell line overexpressing cytosolic eGFP. These cells were harvested and lysed to perform pull-downs in quadruplicates using GFP-Trap® nanobodies, followed by tryptic digestion and peptide clean-up. Samples were subsequently analyzed by LC/MS-MS. In parallel, the same stable cell lines were used for confocal microscopy imaging to determine the localization of the recombinant claudins within the cells (**Figure 4A**).

3.1.1. Confocal microscopy images show recombinant YFP- /CFP-claudin localization in the plasma membrane and vesicles

Confocal microscopy images show that the recombinant claudins are correctly targeted to the tight junctions as their YFP/CFP signal co-localizes with the signal from the tight junction protein ZO-1. In some instances, YFP/CFP signal also localized to cytosol depending on the overexpressed recombinant claudin (**Figure 4B, Supplementary Figure 1**). Claudin pull-downs were individually analyzed and compared against the cytosolic control using a moderated t-test to determine significant interactions for each member of the claudin protein family. On top of the consensus significance cut-off of adjusted p-value<0.05, a more stringent second significance level was established for each claudin pull-down corresponding to the adjusted p-value that leaves only 5% of significant interactions for the eGFP control (**Figure 4C, Supplementary Figure 2**).

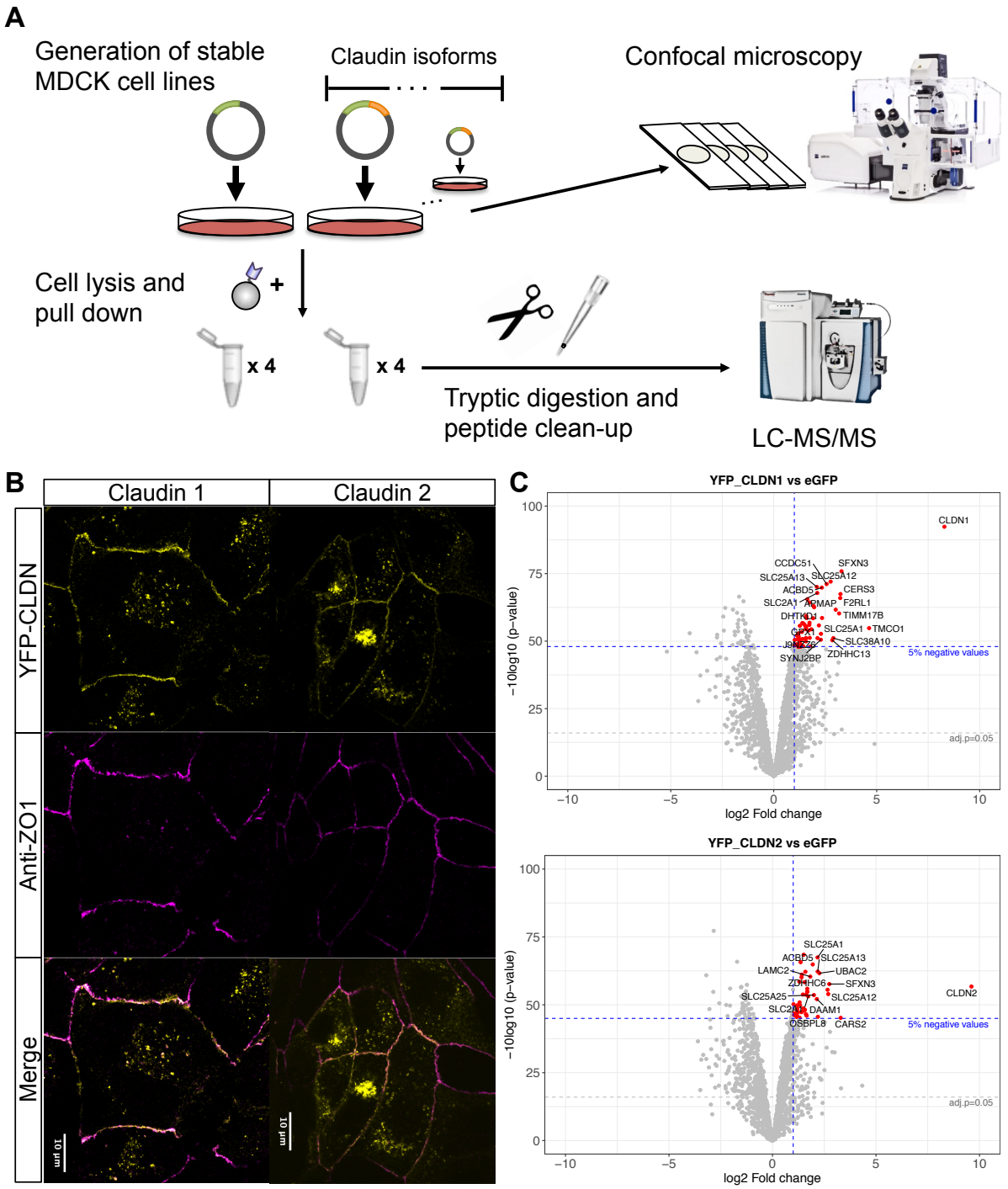


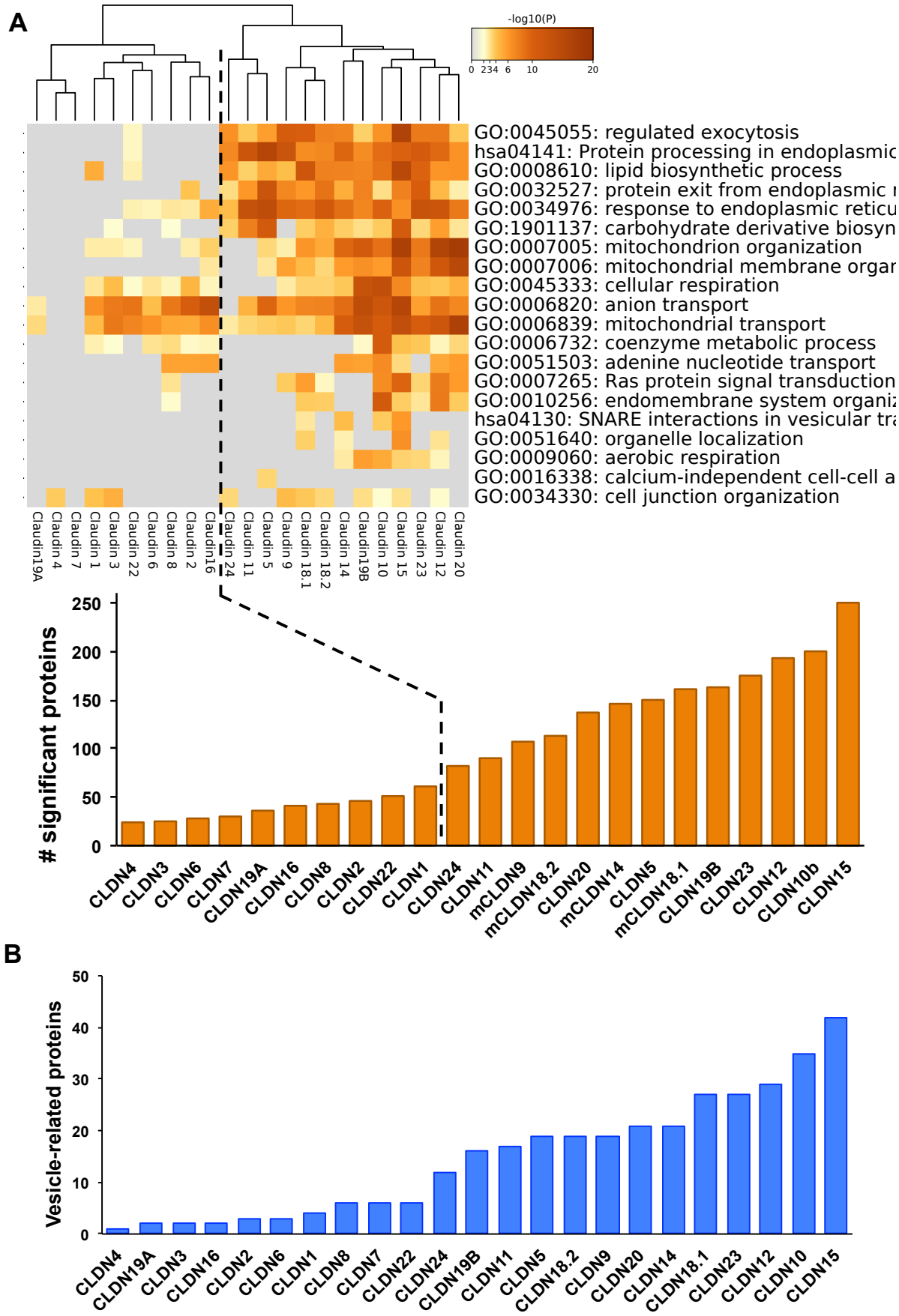
Figure 4. Description of the co-immunoprecipitation experiments. A. Schematic representation of the workflow for co-immunoprecipitation experiments. **B.** Generation of stable cell lines overexpressing recombinant claudins (N-terminal YFP/CFP- fused) for every member of the claudin family and cytosolic eGFP as a control. **C.** Pull-downs using GFP-Trap® nanobodies. The following criteria were used to define significant interactions: first significance level adj-pval<0.05 (grey line), 2nd level cutoff of adj-pval that leaves only 5% of the interactions identified in the GFP control (blue line).

3.1.2. Claudin localization influences the number of significant interactions identified by CoIP.

CoIP experiments resulted in a wide range of significant interactions depending on the claudin bait (from 24 interactions for claudin-4 to 250 for claudin-15) (**Table 2**). An enrichment analysis using the Metascape gene annotation and analysis resource (Zhou et al., 2019) showed two main clusters that separate claudins based on Gene Ontology (GO) terms related to exocytosis and endoplasmic reticulum. These clusters also separate claudins into two groups corresponding to a lower and higher number of significant interactions (**Figure 5A**). The same two groups of claudins are maintained when looking only at proteins with GO Cellular Component (GOCC) annotations related to vesicles: claudins with lower significant interactions identified have also a lower number of vesicle-related proteins and vice versa (**Figure 5B**). According to the confocal microscopy images, claudins with the lowest number of interactions are those expressed solely or mostly in the tight junctions. On the other hand, claudins with a higher number of significant interacting proteins correspond to those localized not only in the tight junctions but in cytosolic vesicles as well (**Figure 5C**). At first, this could look like a result of an artifact of overexpressing exogenous claudins in cell culture, but some claudins are known to be stored in cytosolic vesicles for later incorporation into the tight junction as part of the dynamic of tight junction remodeling. In combination, results from LC-MS/MS and confocal microscopy establish a correlation between the localization of the claudins and the number of significant interactions identified for each of them.

Claudin	Significant interactions	p-value second cutoff
Claudin-1	61	3.16E-04
Claudin-2	46	5.83E-04
Claudin-3	25	1.99E-03
Claudin-4	24	8.56E-03
Claudin-5	150	3.04E-04
Claudin-6	28	7.23E-03
Claudin-7	30	6.66E-02*
Claudin-8	43	5.01E-04
(m)Claudin-9	107	1.24E-03
Claudin-10	200	2.93E-04
Claudin-11	90	1.72E-02
Claudin-12	193	1.31E-03
(m)Claudin-14	146	2.08E-03
Claudin-15	250	4.95E-04
Claudin-16	41	4.45E-02
(m)Claudin-18.1	161	3.26E-04
(m)Claudin-18.2	113	2.93E-04
Claudin-19a	36	5.34E-03
Claudin-19b	163	1.69E-03
Claudin-20	137	2.53E-04
Claudin-22	51	9.71E-04
Claudin-23	175	4.31E-04
Claudin-24	82	3.93E-03

Table 2. Summary of significant interactions found by CoIP. Significant interactions identified on each claudin CoIP and p-value calculated for the second level cutoff. (m): recombinant claudins cloned from mouse cDNA. (*) Claudin-7 second cutoff adj-p > 0.05 leading to the inclusion of 5 more proteins as significant (YIPF3, FLOT1, TES, SEC61G, and NRAS; the latest with the highest adj-pval = 0.057). CFP-CLDN17 construct was neither detected by confocal microscopy nor by mass spectrometry; therefore, it is not present in this list.


A
-log10(P)
0 2 3 4 6 10 20
GO:0045055: regulated exocytosis
hsa04141: Protein processing in endoplasmic
GO:0008610: lipid biosynthetic process
GO:0032527: protein exit from endoplasmic r
GO:0034976: response to endoplasmic reticu
GO:1901137: carbohydrate derivative biosyn
GO:0007005: mitochondrion organization
GO:0007006: mitochondrial membrane organ
GO:0045333: cellular respiration
GO:0006820: anion transport
GO:0006839: mitochondrial transport
GO:0006732: coenzyme metabolic process
GO:0051503: adenine nucleotide transport
GO:0007265: Ras protein signal transduction
GO:0010256: endomembrane system organiz
hsa04130: SNARE interactions in vesicular tra
GO:0051640: organelle localization
GO:0009060: aerobic respiration
GO:0016338: calcium-independent cell-cell a
GO:0034330: cell junction organization
Claudin19A
Claudin 4
Claudin 7
Claudin 1
Claudin 3
Claudin 22
Claudin 6
Claudin 8
Claudin 2
Claudin16
Claudin 24
Claudin 11
Claudin 5
Claudin 9
Claudin 18.1
Claudin 18.2
Claudin 14
Claudin19B
Claudin 10
Claudin 15
Claudin 23
Claudin 12
Claudin 20
significant proteins
0
50
100
150
200
250
CLDN4
CLDN3
CLDN6
CLDN7
CLDN19A
CLDN16
CLDN8
CLDN2
CLDN22
CLDN1
CLDN24
CLDN11
mCLDN9
mCLDN18.2
CLDN20
mCLDN14
CLDN5
mCLDN18.1
CLDN19B
CLDN23
CLDN12
CLDN10b
CLDN15
B
Vesicle-related proteins
0
10
20
30
40
50
CLDN4
CLDN19A
CLDN3
CLDN16
CLDN2
CLDN6
CLDN1
CLDN8
CLDN7
CLDN22
CLDN24
CLDN19B
CLDN11
CLDN5
CLDN18.2
CLDN9
CLDN20
CLDN14
CLDN18.1
CLDN23
CLDN12
CLDN10
CLDN15

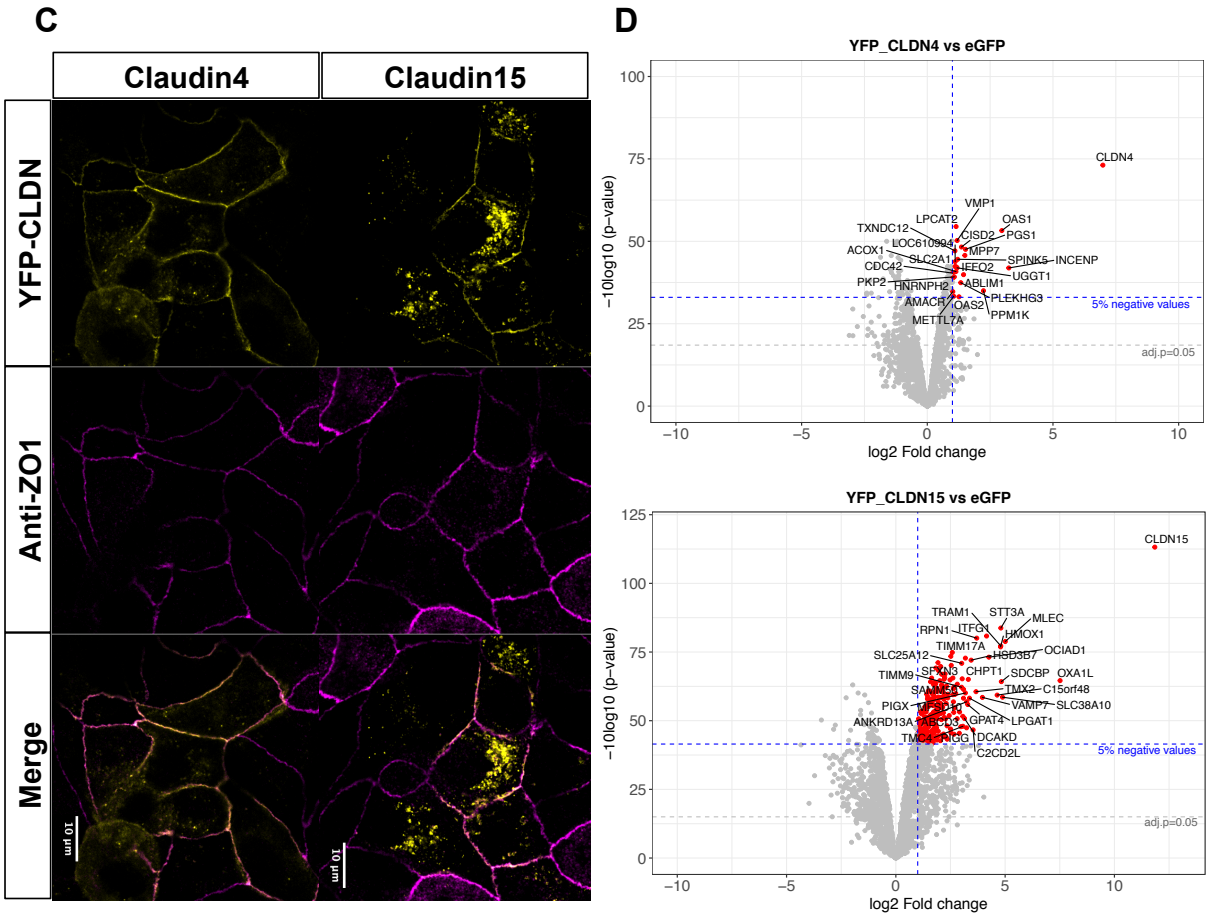


Figure 5. Claudin localization influences the number of significant interactions identified. A. Metascape enrichment analysis of CoIP results revealing two main clusters of claudins. **B.** The main difference between the two clusters is the enrichment of vesicle-related proteins. **C.** Claudin-4 localizes mainly at the cell membrane in contrast to claudin-15 that has a strong signal in the cytosol. **D.** Those differences are also reflected in the number of different significant proteins enriched by CoIP.

3.1.3. CoIP results provide a broad interaction map of the claudin family.

CoIP dataset provides an extensive map of the claudin protein family interactome that expands the current knowledge in protein-protein interactions of claudins inside and outside the tight junction. A total of 758 different proteins were identified that significantly interact with one or more members of the claudin family (**Figure 6A, Supplementary Figure 3**). To make that information more comprehensive, the identified proteins were systematically classified into 13 different groups based on their GOCC annotations extracted from the Database for Annotation, Visualization, and Integrated Discovery (DAVID, Huang et al., 2009a, 2009b). The classification system starts with specific tight junction and adherens junction related GO terms and

moves to broader terms including localization to the plasma membrane and different organelles in the cell (**Figure 6B**).

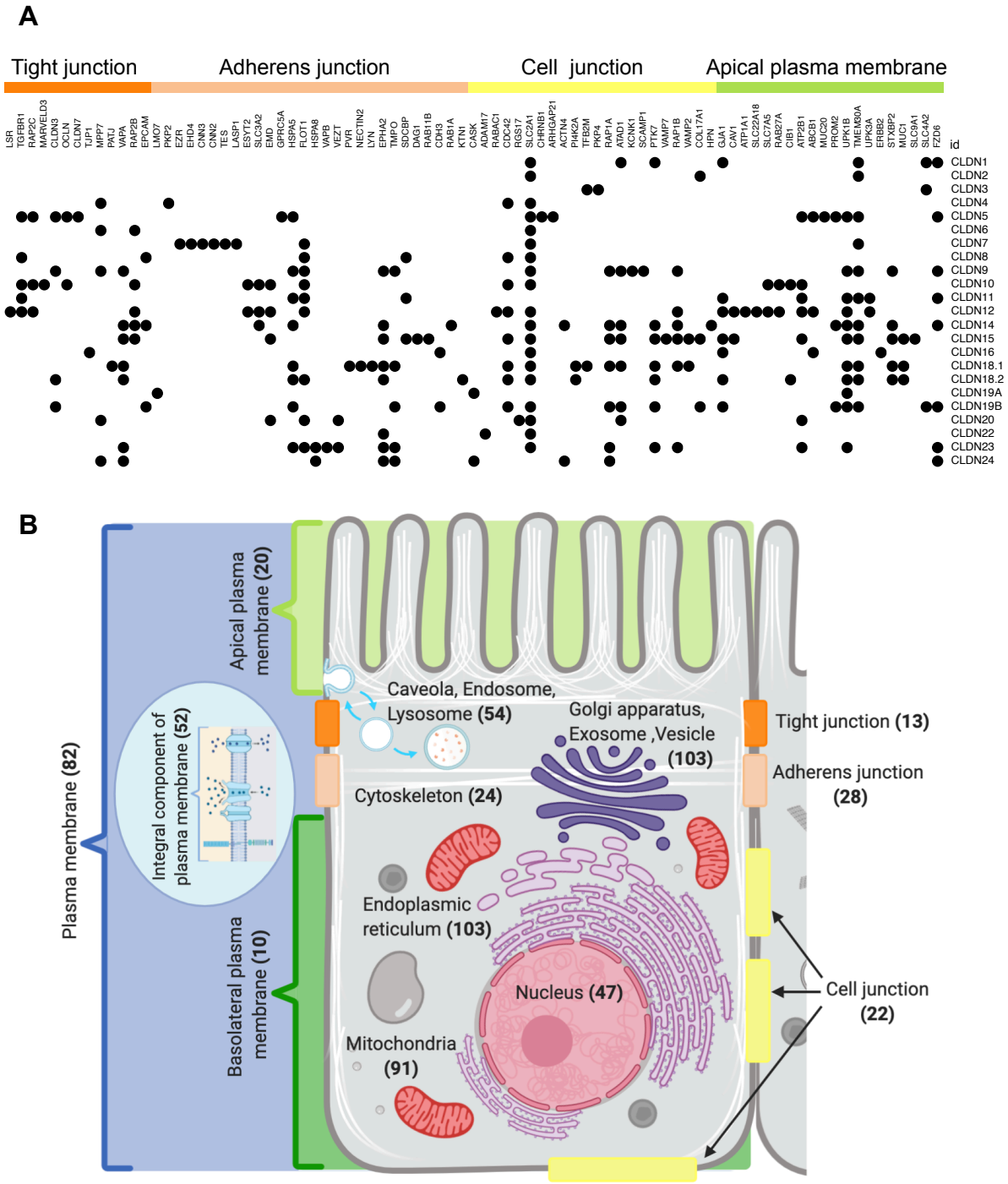


Figure 6. Resource interactome map of the claudin family identified by CoIP A, B. From a total of 758 significant interactions identified by CoIP, 708 were annotated using DAVID. Proteins were classified into 13 different groups based on their Gene Ontology annotation (GOCC).

3.1.4. CoIP allows for the Identification of new heteromeric/heterotypic interactions for claudin-5, claudin-9, and claudin-18.

It was also possible to identify heteromeric/heterotypic interactions of endogenous claudin-3 with the overexpressed recombinant claudin-19b and -5 that are well described in the literature (**Table 3**). Three heterotypic interactions that, to our knowledge, have not been described yet were also found. That is the interaction of endogenous claudin-3 with overexpressed claudin-9 and claudin-18.2, and endogenous claudin-7 with overexpressed claudin-5 (**Figure 7**). Although not significant, the homomeric/homotypic interaction between endogenous and overexpressed claudin-7 was also detected by mass spectrometry (**Supplementary Figure 2**).

Interaction	Method	Reference
Claudin-1, -3	Affinity capture – WB Immunofluorescence , CoIP-WB	Coyne et al., 2003; Daugherty et al., 2007
Claudin-3, -5	Affinity capture – WB Immunofluorescence , CoIP-WB	Coyne et al., 2003; Daugherty et al., 2007
Claudin-2, -3	Immunofluorescence , CoIP-WB	Daugherty et al., 2007
Claudin-3, -19	Live-cell imaging - FRET	Milatz et al., 2017
Claudin-16, -19	Live-cell imaging - FRET	Milatz et al., 2017

Table 3. Claudin-claudin heteromeric/heterotypic interactions described in the literature. Heteromeric and heterotypic interactions occur between different claudins within the same plasma membrane (in *cis*), and in the plasma membrane of adjacent cells (in *trans*) respectively.

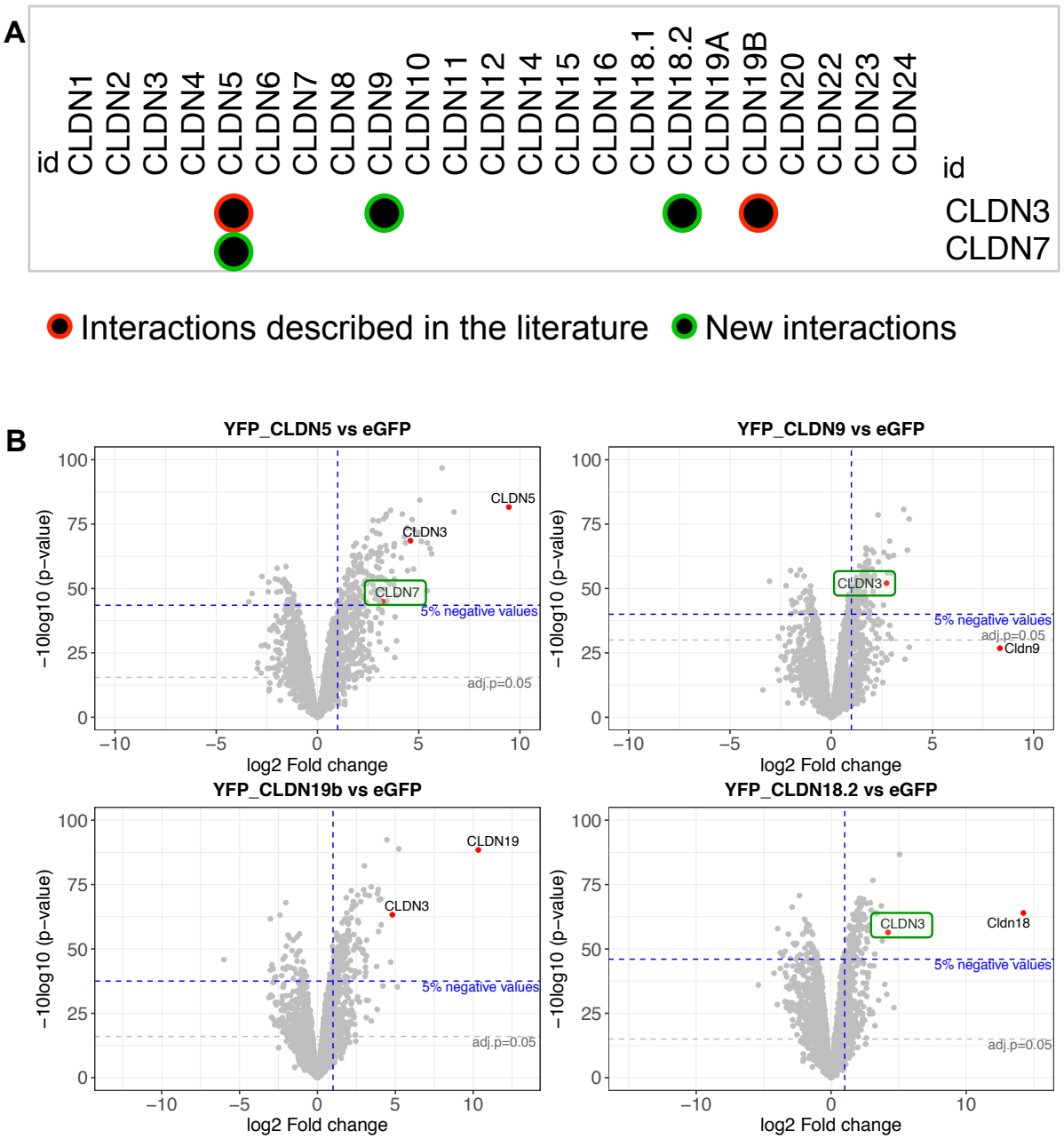


Figure 7. Claudin-claudin interactions identified by CoIP. A. CoIP results show new heteromeric/heterotypic interactions for claudin-5, -9 and -18 (in green). **B.** The interacting claudins (CLDN-3 and CLDN-7) showed in the volcano plots are endogenously expressed by the MDCK-C7 cell line.

3.2. Interactome study of the cytosolic C-terminal tail of claudins using peptide-based pull-downs (PRISMA).

Protein Interaction Screen on a peptide Matrix (PRISMA) allows studying weak interactions taking place in unstructured regions of proteins. These interactions are between globular proteins and short linear motifs. In the case of claudins, the cytosolic C-terminal tails are long and unstructured, and a database search using the Eukaryotic Linear Motif (ELM) resource for Functional Sites in Proteins revealed the presence of several predicted ligand-binding motifs in these regions summarized in Table 4.

ELM Name	Pattern	Functional site class	Claudin
LIG_14-3-3_CanoR1	R[^DE]{0,2}[^DEPG]([ST])(([FWYLMV].)|([^PRIKGN]P)|([^PRIKGN].{2,4}[VILMFWYP]))	14-3-3 binding phosphopeptide motif	1, 6, 7, 8, 10, 11, 12, 14, 16, 17, 19A, 19B, 23
LIG_SH3_3	...[PV]..P	SH3 ligands	2, 5, 9, 14, 19A, 19B, 23, 25
LIG_PDZ_Class3	...[DE].[ACVILF]$	PDZ ligands	1,3, 7, 9, 14, 18, 19B, 20
LIG_PDZ_Class1	...[ST].[ACVILF]$	PDZ ligands	16, 17, 23
LIG_PDZ_Class2	...[VLIFY].[ACVILF]$	PDZ ligands	19A, 25
LIG_SH3_1	[RKY]..P..P	SH3 ligands	5
LIG_SH3_2	P..P.[KR]	SH3 ligands	2
LIG_SH3_4	KP..[QK]…	SH3 ligands	23
LIG_EVH1_2	PP..F	EVH1 ligands	9
LIG_WW3	.PPR.	WW domain ligands	4

Table 4. Ligand binding sites present in the C-terminal region of human claudins. Using the ELM prediction tool from The Eukaryotic Linear Motif (ELM) resource for Functional Sites in Proteins. Additional filters based on taxonomy and cellular compartment were used to select the motifs of interest.

PRISMA was applied for the study of PPIs with the cytosolic C-termini of claudins. The amino acid (aa) sequence of claudins was used to generate a library of 15 aa long overlapping peptides with a 5 aa window shift mapping the cytosolic tails of the different members of the protein family. Modified versions of some peptides were included to investigate the effect that PTMs have in the regulation of certain interactions. The modified peptides contain phosphorylated residues that have been previously described (**Table 5**). These peptides (unmodified and modified) were synthesized into spots on a cellulose membrane that was incubated with MDCK-C7 cell lysate. The different spots were excised from the membrane and transferred into 96-well plates for tryptic digestion and desalting followed by individual analysis by LC-MS/MS. The raw data were analyzed using MaxQuant and peptide intensities were quantified in a label-free manner (LFQ). Data filtering and statistical analysis were done in groups of peptides belonging to the same claudin. Significant interactions were defined by pairwise comparison of the LFQ intensity of proteins identified for the different peptide pull-downs within the same claudin using moderated t-test (**Figure 8.** More details in the methods section).

Claudin	C-terminus length (AA)	Number of peptides		
		Unmodified	Phosphorylated (residue position)	
Claudin-1	27	4	1	(Y210*)
Claudin-2	47	7	2	(Y224§)
Claudin-3	40	6	3	(Y214§, Y219*)
Claudin-4	28	3	1	(Y208*)
Claudin-5	38	5	2	(Y212§, Y217*)
Claudin-6	39	5	2	(Y214§, Y219*)
Claudin-7	30	4	1	(Y210*)
Claudin-8	38	5	-	
Claudin-9	37	5	1	(Y216*)
Claudin-10	47	7	1	(Y227*)
Claudin-11	29	3	-	
Claudin-12	49	7	-	
Claudin-14	56	9	2	(Y233§)
Claudin-15	46	7	-	
Claudin-16	45	7	2	(Y299§)
Claudin-17	39	5	-	
Claudin-18.1/18.2	66	11	1	(Y260*)
Claudin-19a	43	6	-	
Claudin-19b	30	4	-	
Claudin-20	38	5	-	
Claudin-22	35	4	-	
Claudin-23	111	17	-	
Claudin-24	38	5	-	
Claudin-25	44	6	-	

Table 5. Summary of the synthetic peptides used in the PRISMA experiment. 15 aa long peptides were designed with a 5 aa shift to map the entire C-terminal region of each claudin on the list. Phosphorylated versions of the most C-terminal peptides where included based on the literature. *Tyrosine residue in position -1. §Tyrosine residue in position -6.

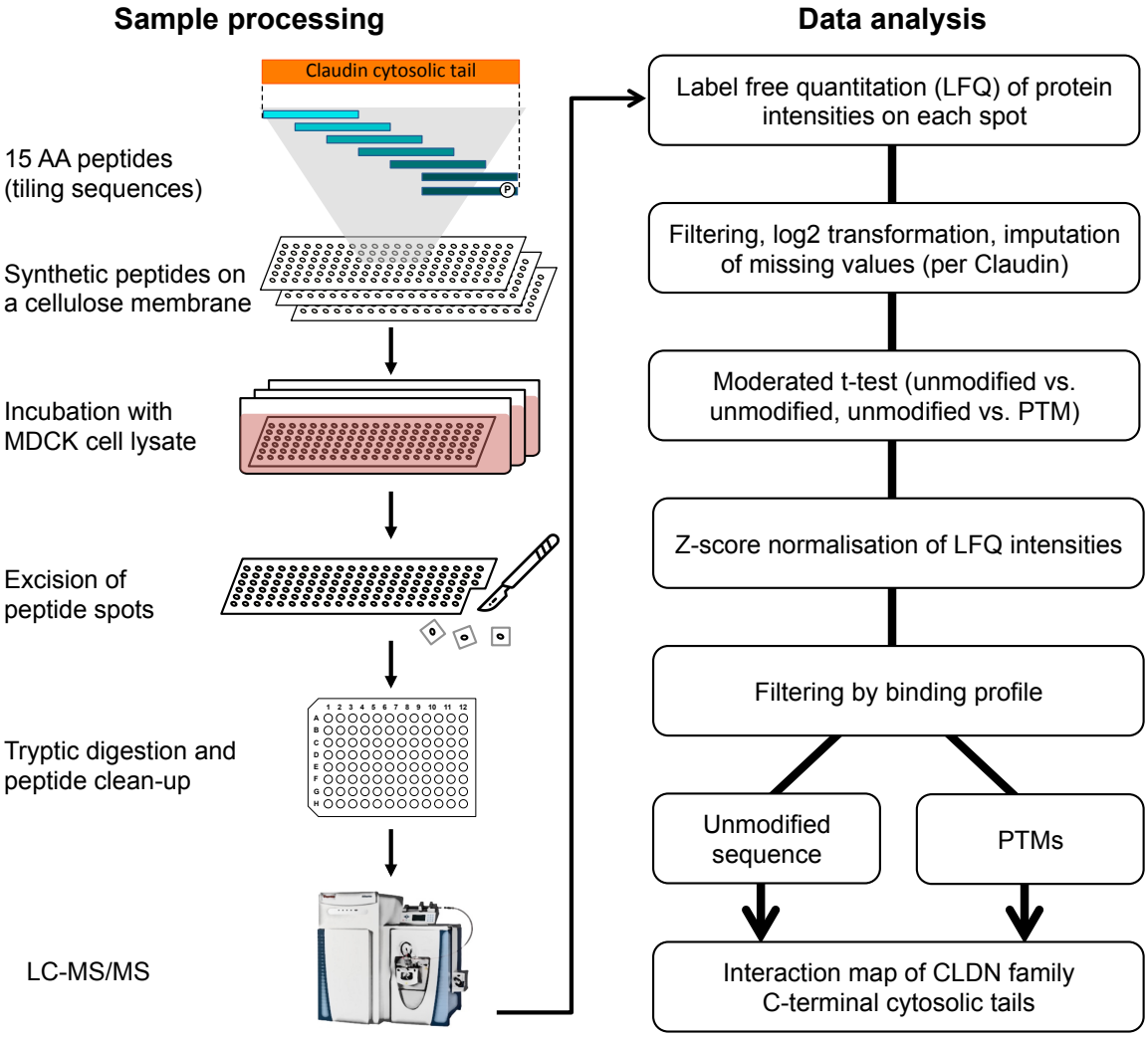


Figure 8. Schematic representation of PRISMA protocol. Cellulose membranes containing synthetic peptides from the C-terminal region of claudins were incubated in MDCK-C7 cell extract. After washing and drying, peptide spots were excised and transferred to a 96-well plate for further in-solution digestion and peptide cleanup. Samples (3 replicates from each peptide spot) were measured by LC-MS/MS. Raw intensity outputs were analyzed with MaxQuant to obtain quantitative data (LFQ). Filtering, imputation of missing values, and statistical analysis were done using Rstudio. Pair-wise comparisons of unmodified peptides from the same claudin and each unmodified-phosphorylated pair of peptides were made applying moderated t-test. After normalization of the intensity values, significant proteins (adj-pval<0.05) were filtered based on: a) intensity binding profile across the unmodified peptides, b) changes in the intensity values of proteins identified in unmodified and phosphorylated peptide pull-downs

3.2.1. PRISMA dataset shows patterns of interactions shared by most claudins, shared only by a subgroup, and specific for certain claudins.

For the unmodified claudin peptides, the significant interactions were filtered according to intensity profiles that fit a SLiM dependent type of interaction. This means interaction across three overlapping peptides with the highest intensity in the middle. As a result of this filtering process, 148 proteins were obtained showing that kind of intensity profile. Based on the normalized LFQ intensities of the significant interacting proteins, it is possible to visually identify patterns of interactions shared by many claudins, shared by only a subset, or exclusive to certain members of the claudin protein family. Some interacting proteins form clusters that show a similar pattern repeated across several members of the claudin family (**Figure 9**).

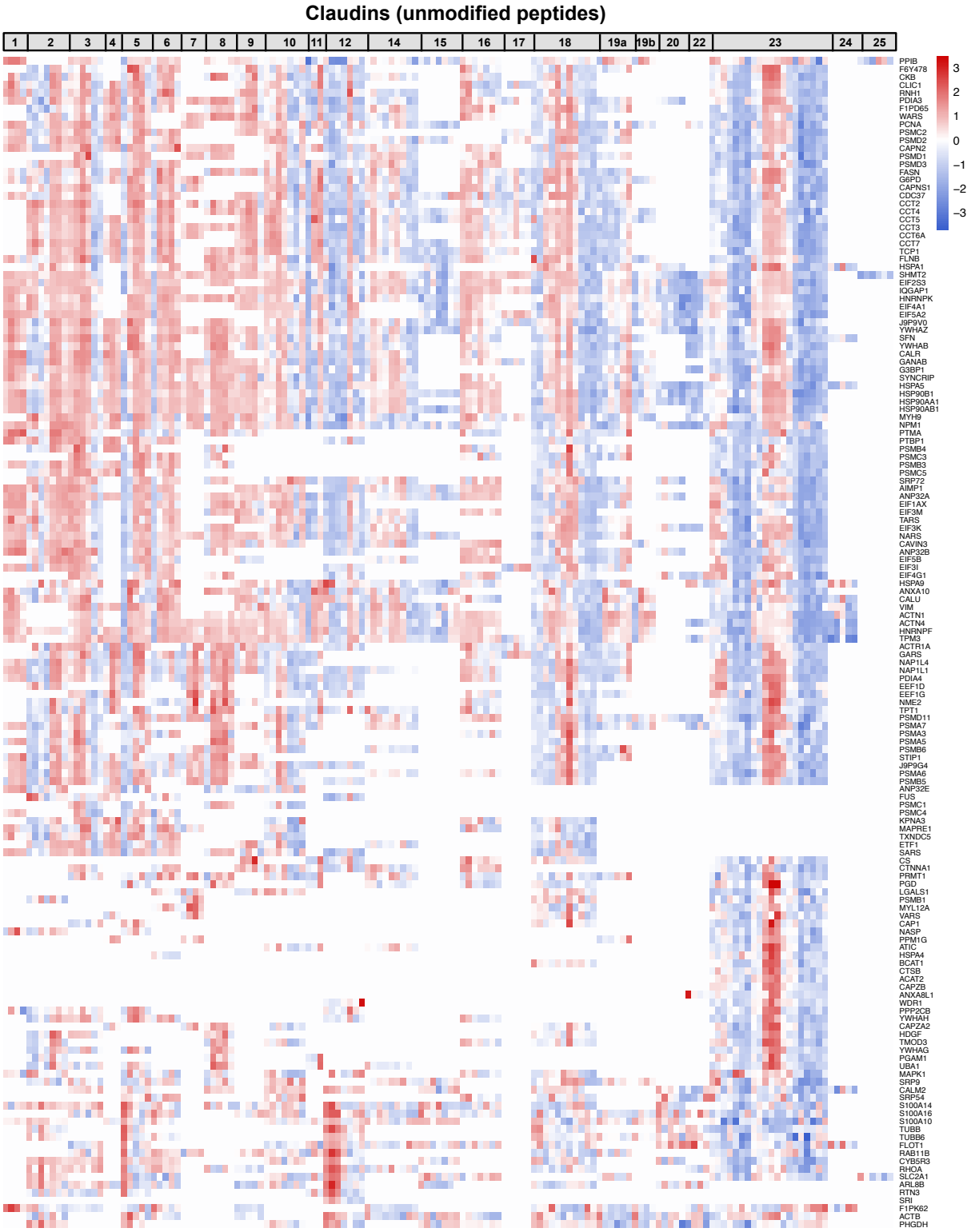


Figure 9. PRISMA maps interactions to the C-terminal cytosolic tail of specific claudins. LFQ normalized intensities (y-axis) of the significant proteins identified interacting with unmodified peptides derived from the C-terminal region of human claudins. PRISMA shows patterns of interactions shared by most members of the claudin protein family, shared only by a subgroup, and specific for certain claudins.

3.2.2. Identification of protein complexes interacting with the cytosolic C-terminal tail of claudins.

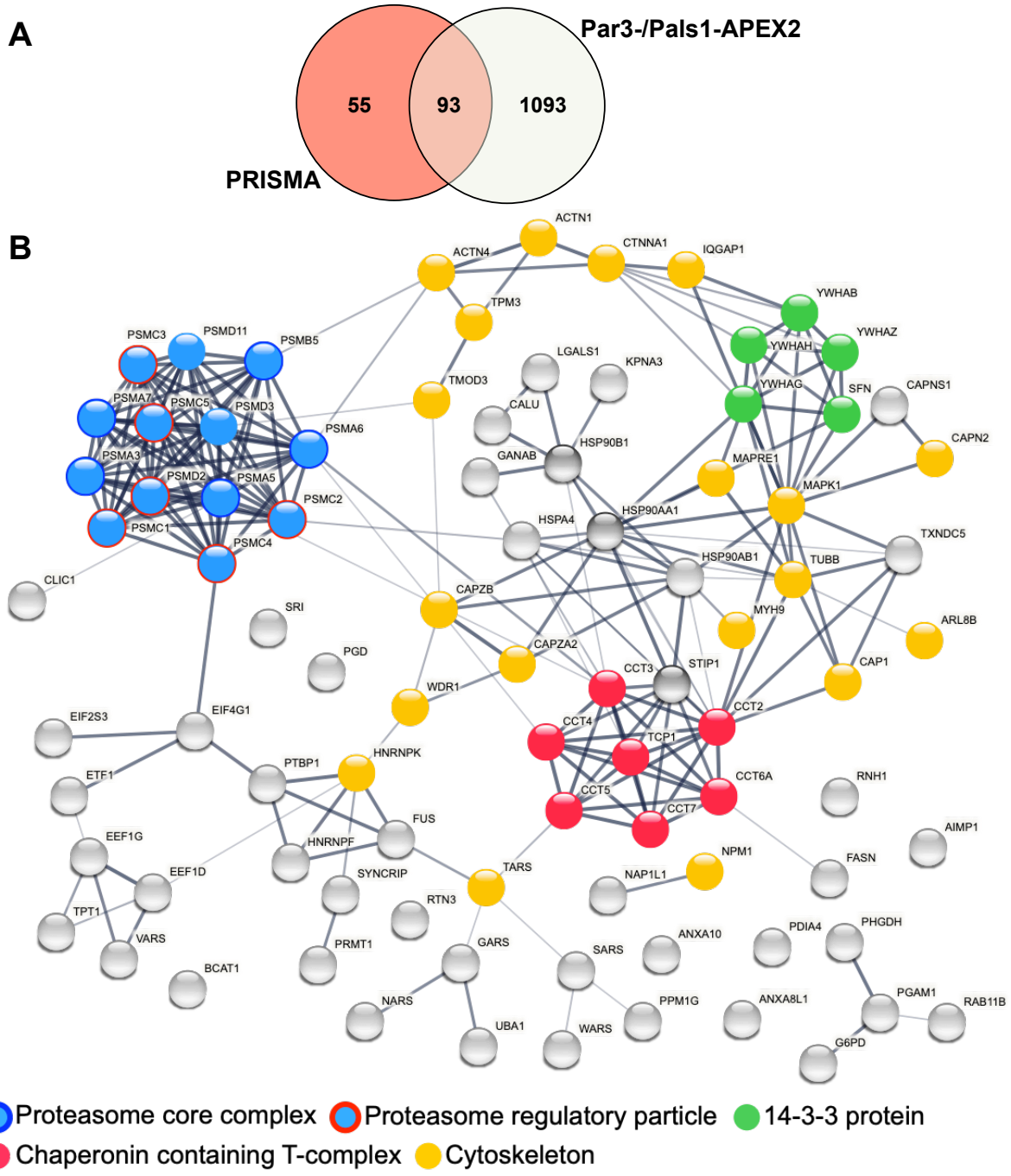
In a previous study, quantitative proximity proteomics was used to resolve the apical junction complex. Using APEX they were able to define the apical and basal limits of the tight junction by the identification of proteins differentially interacting of Pals1 and Par3 respectively (Tan et al., 2020). Comparison of claudin PRISMA dataset with the proteins identified by that previous study showed 89 proteins overlapping with their findings (**Figure 10A**). Among those proteins were 13 members of the proteasome complex, 7 members of the chaperonin containing TCP1 complex (CCT/TriC), 5 members of the 14-3-3 family, and several cytoskeletal proteins (**Figure 10B**). Interestingly, the protein complexes found are differentially interacting with some claudins but not with the others (**Figure 10C**). Tan et al. 2020 also showed that these three groups of proteins seem to localize at the intersection of tight junctions and adherens junctions (**Supplementary Figure 4**).

In the PRISMA experiment, the eight subunits of the CCT/TRiC complex were identified significantly interacting with the cytosolic unstructured tail of several claudins (although CCT8 did not pass the binding profile filter). According to the observed intensity profiles the interactions would be taking place far from the C-terminal PDZ domain-binding motif since the highest intensities are located at the second or third peptide starting from the N-termini.

The proteasome complex (26S) consists of two subcomplexes: a catalytic core particle (CP, or 20S proteasome) and one or two terminal regulatory particles (RP, 19S) that serve as a proteasome activator (Tanaka, 2009). In the PRISMA dataset presented in this work 5 subunits of the proteasome CP (PSMA6, PSMA7, PSMA5, PSMA3, and PSMB5) and 8 subunits of the proteasome RP (PSMC2, PSMC1, PSMC4, PSMC3, PSMC5, PSMD2, PSMD3, and PSMD11) were found interacting with the cytosolic C-terminal tail of several claudins.

Interactions between 14-3-3 proteins and their ligands occur via binding motifs that often, but not always, include a phosphoserine or phosphothreonine (Mrowiec and Schwappach, 2006). The *in silico* analysis of the sequence of the cytosolic tail of claudins using the Eukaryotic Linear Motif (ELM) resource revealed the presence of the canonical 14-3-3 binding phosphopeptide motif (LIG_1433_CanoR_1) in 13 claudins (**Table 4**). Nevertheless, PRISMA dataset shows phosphorylation-

independent interaction of 14-3-3 proteins with unmodified peptides of at least 5 of those claudins (claudin-1, -7, -11, -16, -19A, and -23) and additional claudins with no predicted 14-3-3 binding sites (claudin-2, -3, -4, -5, and -9) (**Figure 10 C**). For claudin-7 and claudin-11, the highest LFQ intensities for the 14-3-3 interactors coincide with the peptides containing the binding motif sequences (-RVPRSYP- and -RFYYTAGSSSP- respectively). In both cases serine and/or threonine residues are present, but according to the PhosphoSitePlus® database (Hornbeck et al., 2015), phosphorylation has been only described for the serine residues of claudin-11 (S196, S197, S198).



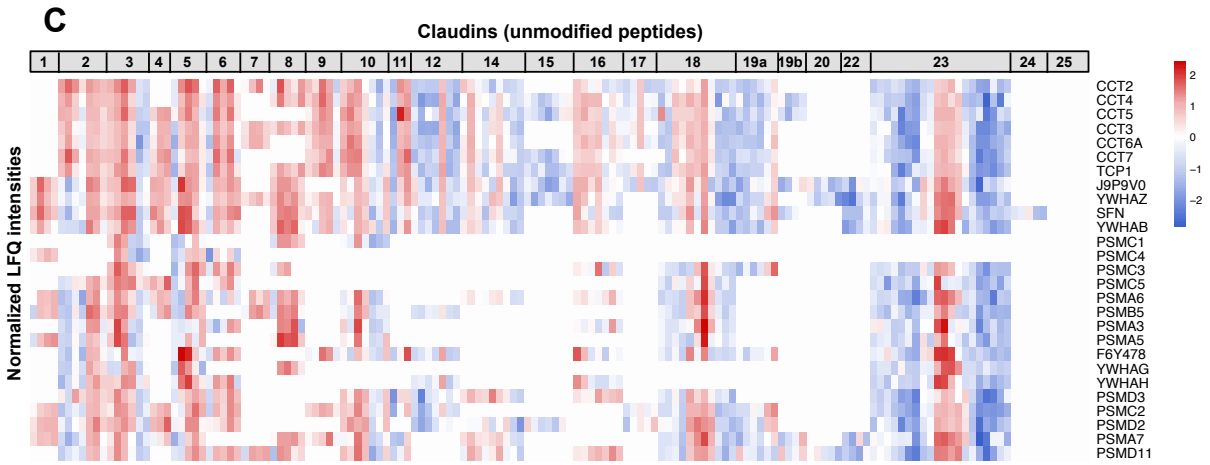


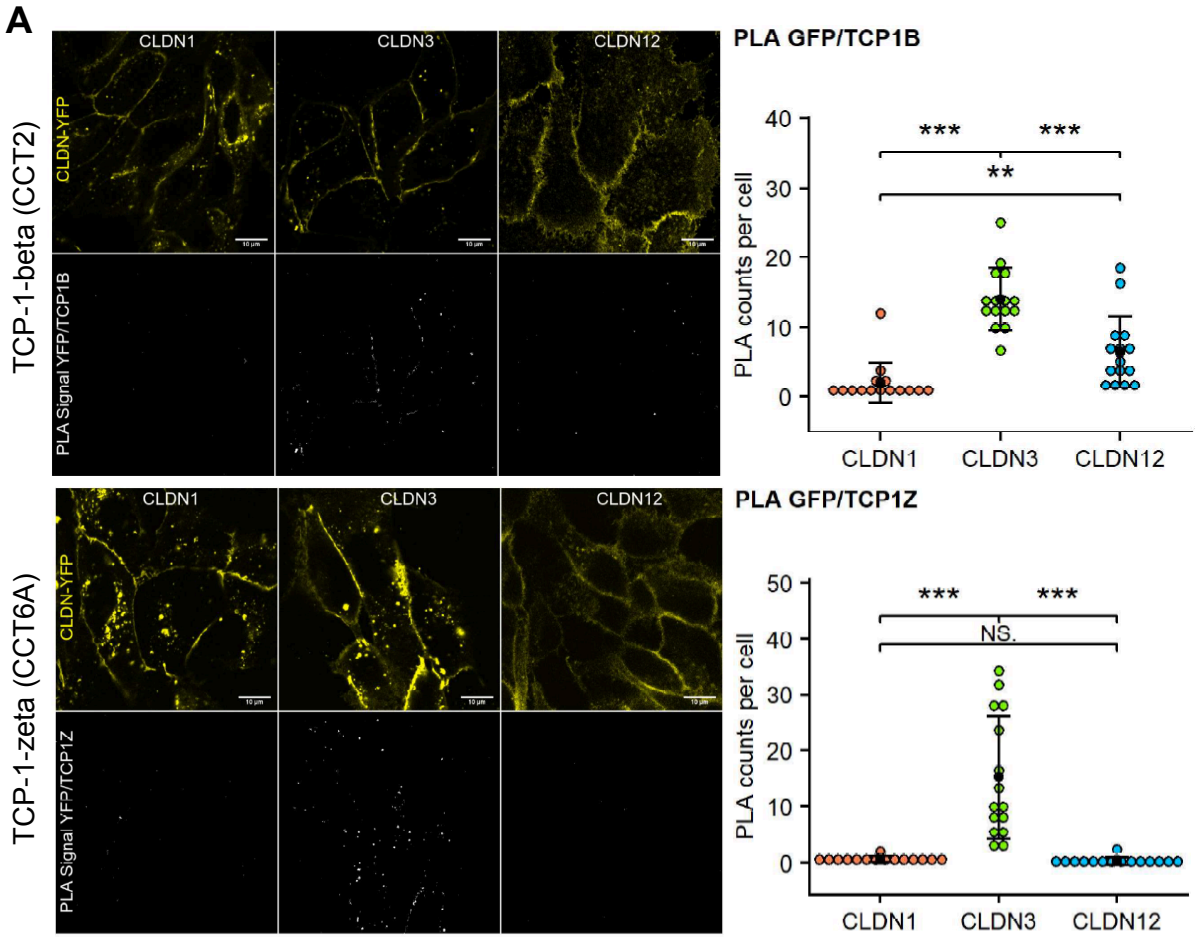
Figure 10. PRISMA shows protein complexes interacting with specific peptides from the C-terminal region of a subset of claudins. **A.** Venn Diagram showing the overlap between the proteins interacting with unmodified claudin peptides (PRISMA) and a previous dataset of proteins interacting with the apical junction complex identified by quantitative proximity proteomics (Tan et al. (2020), Current Biology). **B.** String network of the proteins identified in both datasets. **C.** Interaction map of the 14-3-3 proteins, proteasome complex, and chaperonin containing T-complex. . Uniprot entries J9P9V0 and F6Y478 correspond to 14-3-3 *C. lupus* proteins YWHAE and YWHAQ respectively.

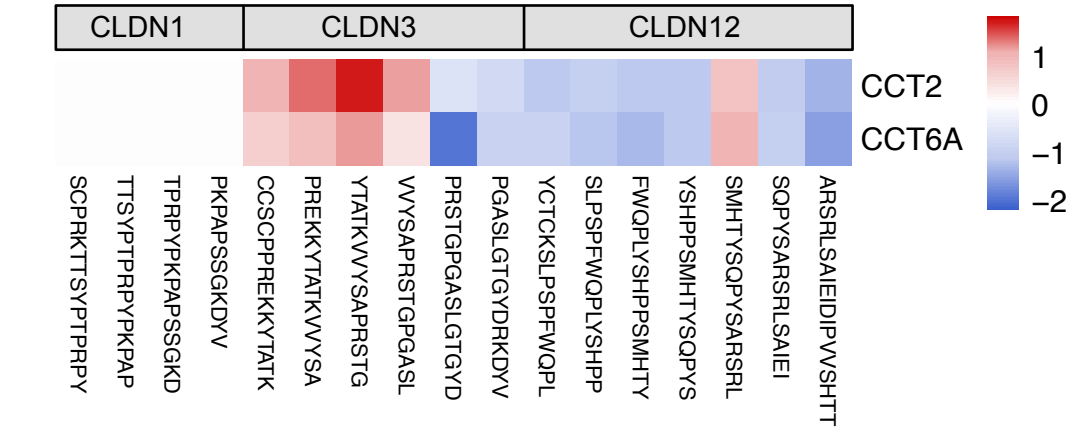
3.2.3. Validation of PRISMA interactions by proximity ligation assays (PLAs)

To corroborate the results obtained by PRISMA we used *in situ* proximity ligation assays (PLA) for three of the proteins identified as differential interactors of the C-terminal cytosolic tail of claudins. Briefly, these assays directly detect protein interactions in a four-step reaction that starts with the detection of the target proteins by primary antibodies (anti-GFP for the recombinant claudins and anti-CCT2, -CCT6 and PSMA3 for the interactors), followed by secondary antibodies (PLA probes) with short specific DNA sequences attached to them. If the proteins of interest are interacting or part of the same complex, the DNA strands attached to the PLA probes can participate in the subsequent rolling circle DNA synthesis that is then detected and quantified using fluorescently labeled complementary oligonucleotides (Fredriksson et al., 2002).

Proximity ligation assays on the same MDCK stable cell lines generated for the CoIP experiments demonstrate the significantly differential interaction of two members of the CCT/TriC chaperonin complex (CCT2 and CCT6) with claudin-3 but not claudin-1 or -12 as shown by the PRISMA results (**Figure 11A, B**). The interaction of the

proteasome complex subunit PSMA3 with claudin-3 but not claudin-1 was also confirmed (**Figure 11C, D**).





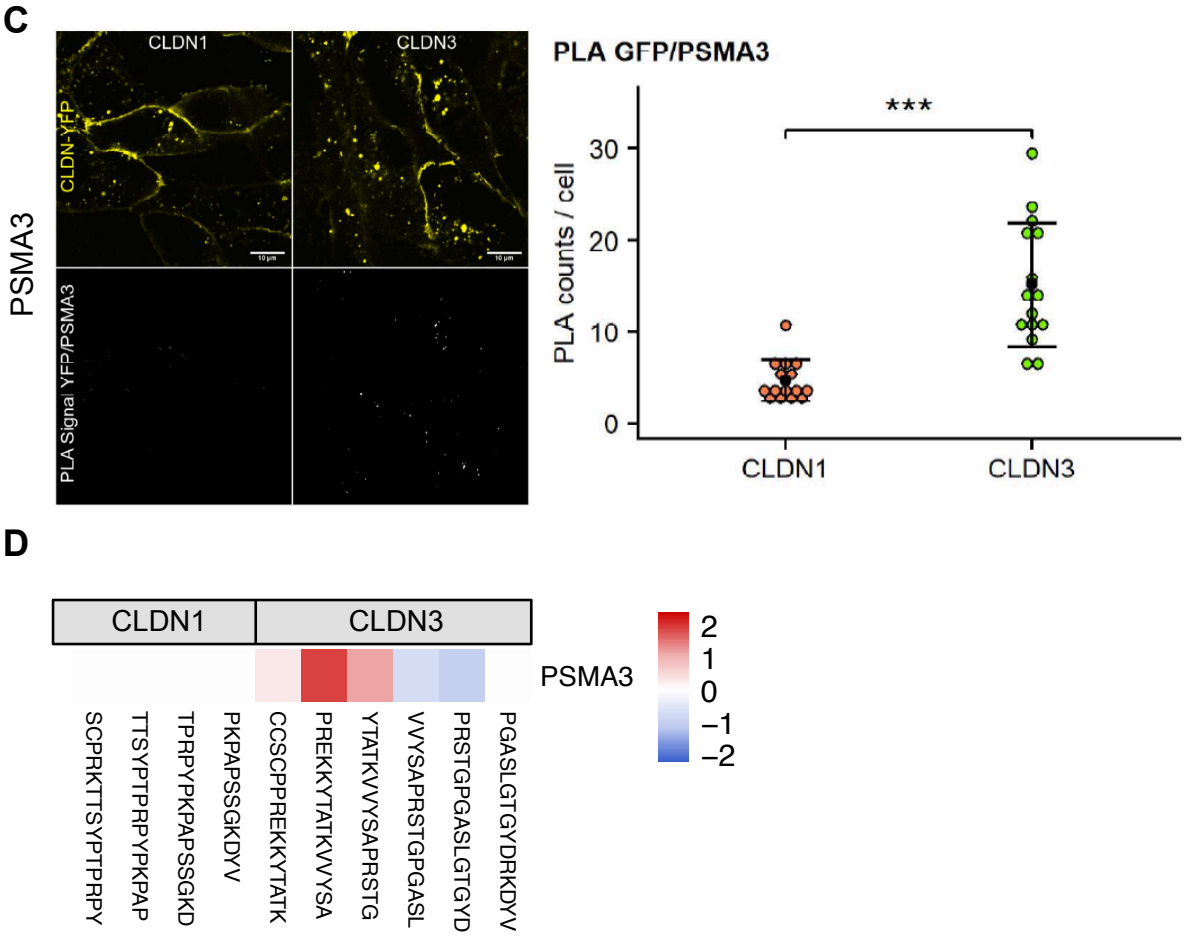


Figure 11. Validation of CCT/TRiC and proteasome subunits interaction with claudins. A, C. CCT/TRiC complex subunits beta and zeta (CCT2 and CCT6A) and proteasome subunit alpha 3 (PSMA3) interact with claudin-3 confirming the results from PRISMA (**B, D**) where these proteins showed an intensity profile that fits a SLiM dependent type of interaction with this particular claudin but claudin-1 or -12. Dot plots represent the average number of PLA signals per cell.

3.2.4. Identification of Interactions with the cytosolic PDZ domain-binding motif of claudins regulated by phosphorylation

PRISMA can be used to study the effect that PTMs have in the regulation of certain protein interactions. In this study, the focus was put on tyrosine phosphorylations described in the C-terminal PDZ domain-binding motif of some claudins.

The majority of claudins have a C-terminal PDZ domain-binding motif defined by the last 3 amino acids (-XYV). This motif interacts with the first PDZ domain of tight junction protein ZO-1, -2, and 3 (ZO-1, -2, -3). This interaction is lost upon phosphorylation of tyrosine residues in positions -1 or -6 in claudins. Therefore, when phosphorylated the PDZ domain-binding motif would be free to interact with other proteins. The disruption of the interaction between claudins and ZO proteins is part of the tight junction remodeling process, an essential feature of this highly dynamic molecular suprastructure (Nomme et al., 2015).

Based on the information present in the PhoshositePlus® database and the literature, PTM modified versions of the most C-terminal peptides from claudins were designed with phosphorylations described in tyrosine residues -1 and/or -6. Interacting proteins significantly binding to the phosphorylated peptides were identified by moderated t-test pairwise comparison against the counterpart unmodified peptides. Using PRISMA 107 proteins were identified differentially binding to the phosphorylated version of the synthetic peptides from 12 members of the claudin family (**Figure 12**).

As mentioned above, an upstream tyrosine in position -6 present in some of the claudins also plays a role in the interaction with ZO-1, -2 and-3, and it can be phosphorylated as well (**Figure 13A**). An example of that is the claudin-2 tyrosine mutant Y224E (Y_{-6} phosphomimetic mutant) that fails to localize to the tight junction and accumulates in cytosolic vesicles in MDCK cells (Nomme et al., 2015). In the case of claudins with phosphorylations described in tyrosine residues -1 and -6 (claudin-3, -5, and -6), the differences in the LFQ normalized intensity of proteins significantly binding to both modified peptides would translate into a difference in the binding strength depending on the position of the phosphorylated tyrosine residue (**Figure 13B**). Therefore, by using PRISMA it is possible to discriminate between the effects that two PTMs in close proximity have on the interaction with a specific protein.

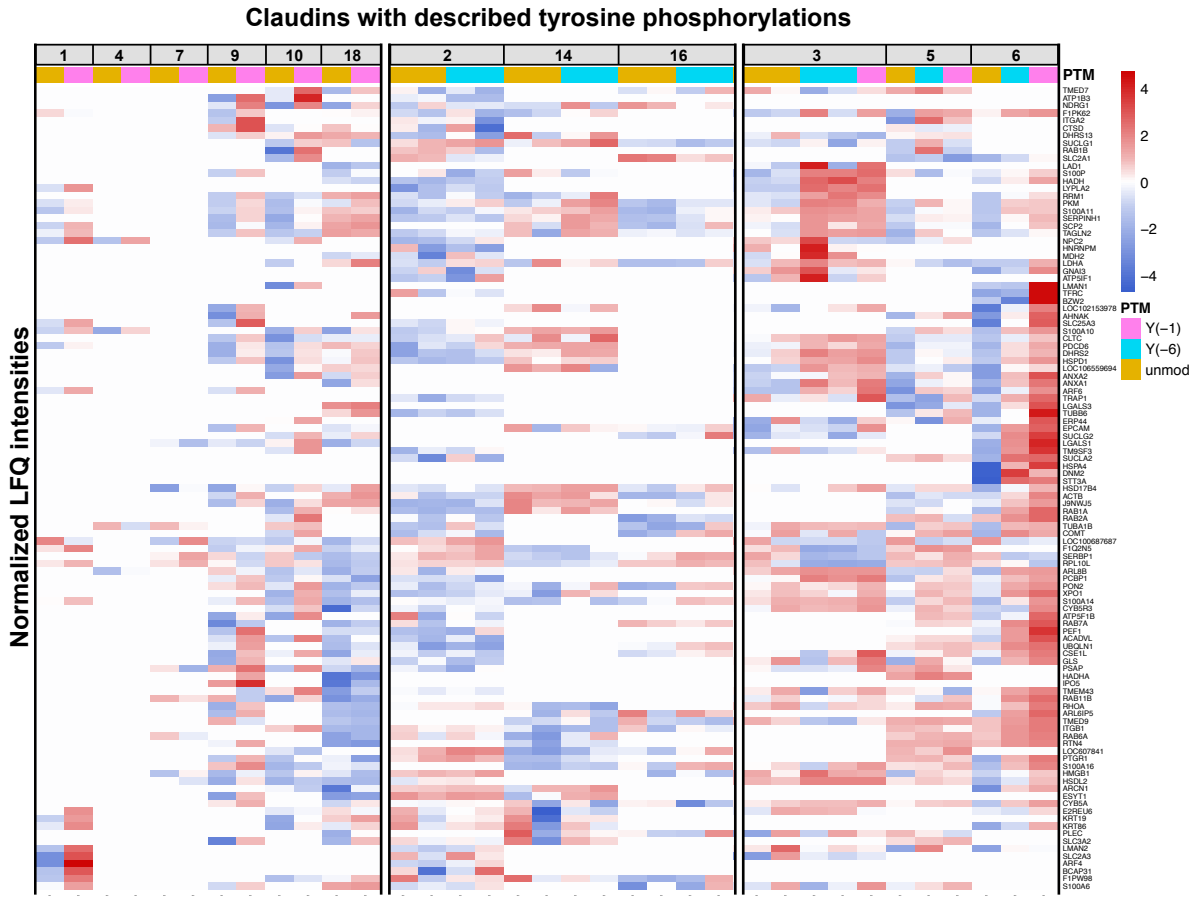


Figure 12. Interactions regulated by phosphorylation of the PDZ domain-binding motif of claudins. PRISMA peptide pull-downs show proteins specifically interacting with the phosphorylated and not with the unmodified peptide versions from several claudins. These peptides belong to the most C-terminal part of the claudins, where the PDZ domain-binding motif is located.

Furthermore, there are several mechanisms of tight junction protein internalization that not only depend on the protein in question but also the stimulus and the cell type (Stamatovic et al., 2017). Clathrin-, caveolae-, and flotillin-dependent endocytosis as well as macropinocytosis, lead to the sorting of tight junction proteins into early endosomes for subsequent recycling or degradation. The dataset obtained by PRISMA shows that the interaction between claudins and certain proteins involved in internalization, sorting and degradation processes take place at the PDZ domain-binding motif. Interestingly enough, these interactions are also regulated by phosphorylations in tyrosines in positions -1 and/or -6. (**Figure 14 A, B**).

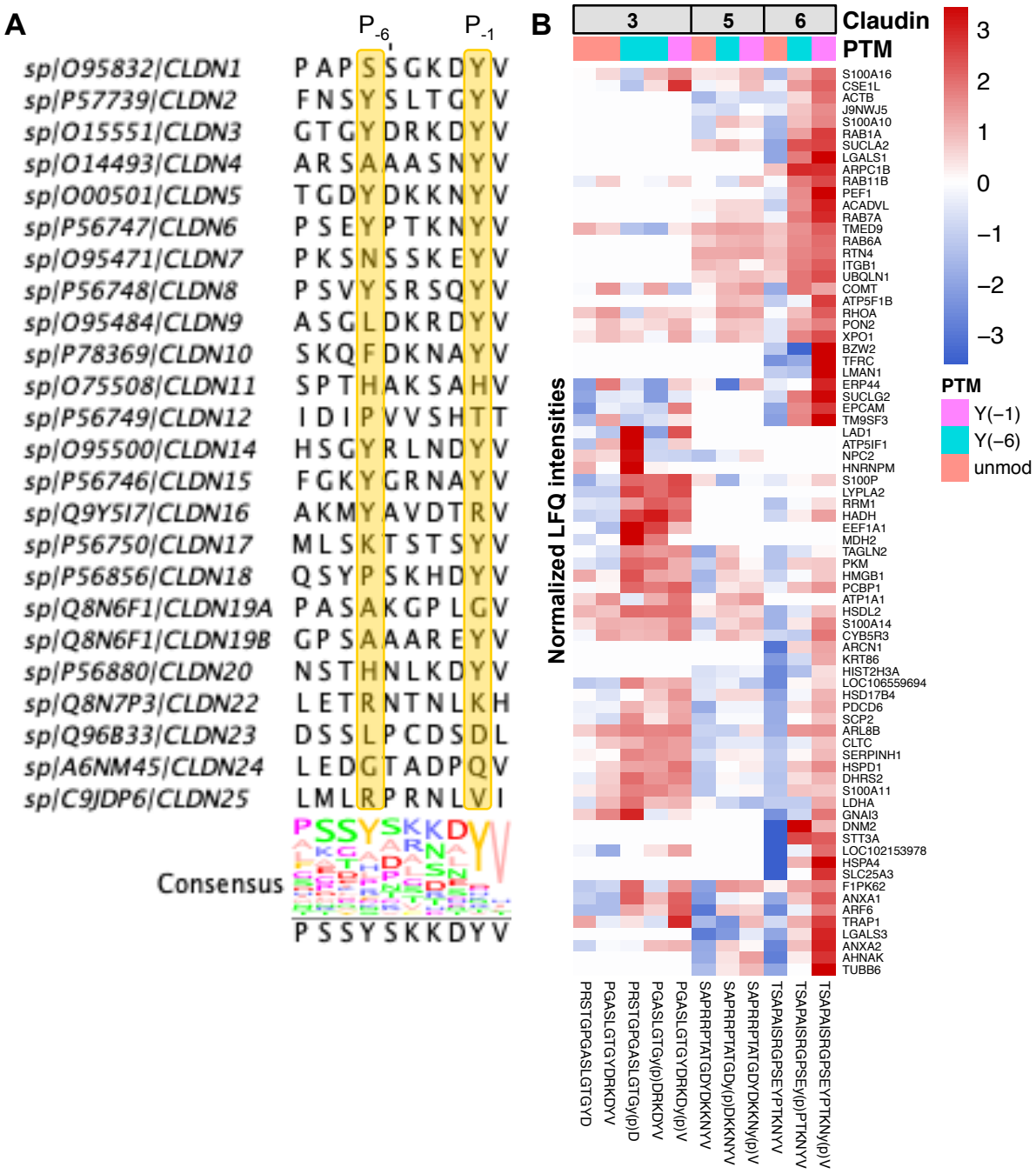


Figure 13. Results from PRISMA link interactions to PTMs on a specific position and show different effects of PTMs located in close proximity. A. Sequence alignment of the most C-terminal amino acids in human claudins where the PDZ domain-binding motif is located. 15 human claudins have a tyrosine residue in position -1, and 8 claudins have a tyrosine residue in position -6 (highlighted in yellow). **B. I**nteractions identified for claudins with phosphorylations described in both tyrosine residues (positions -1 and -6)**.** Proteins interacting with the phosphorylated peptides show different LFQ intensity values depending on the position of the PTM.

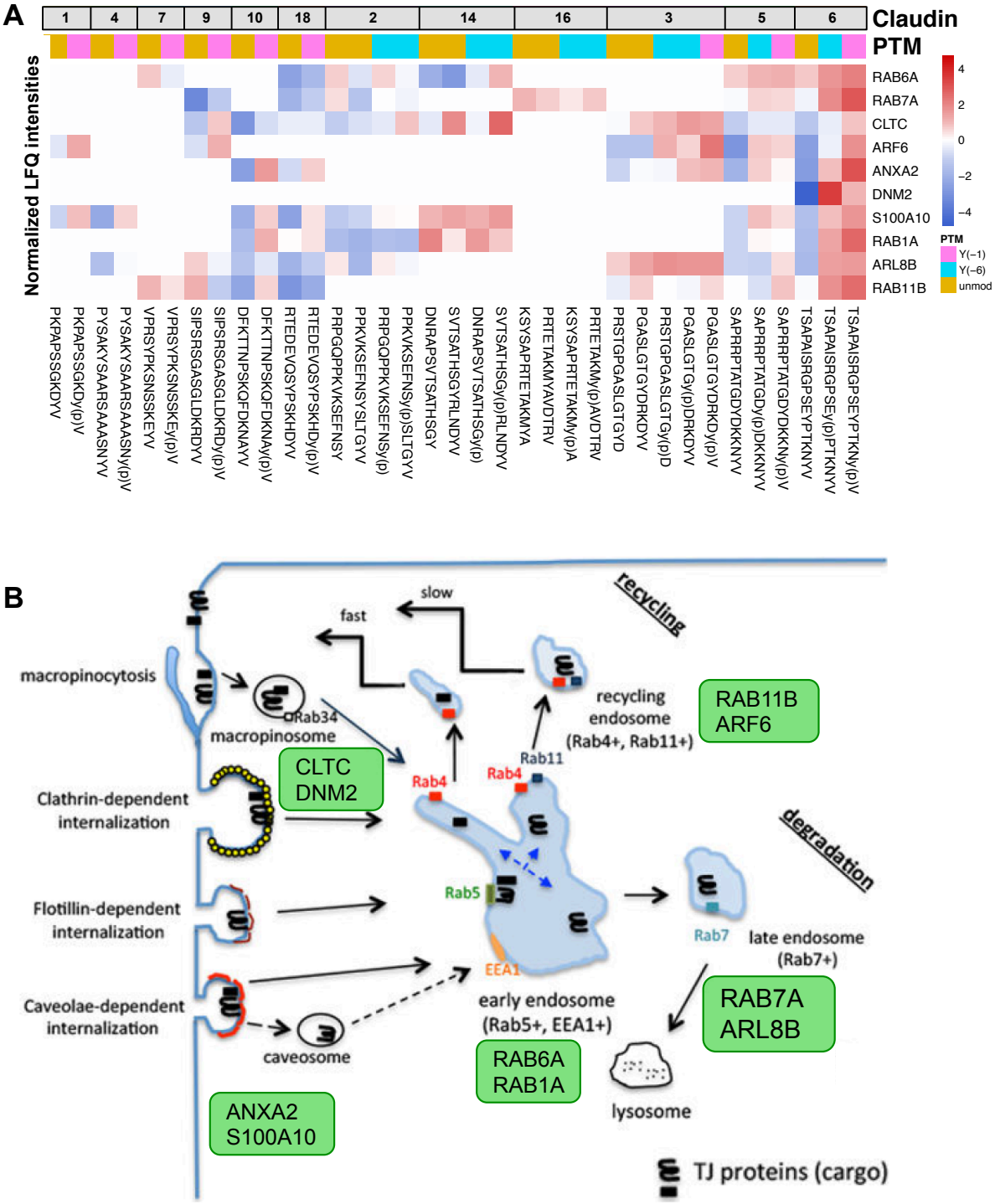


Figure 14. PTMs and claudin endocytosis. A. Phosphorylation of the PDZ domain-binding motif regulates the interaction of claudins with proteins from different internalization mechanisms and sorting pathways. **B.** PRISMA connects proteins related to the endocytic sorting pathways to specific claudins. Figure modified from Stamatovic *et al*. (2017) Ann. N. Y. Acad. Sci.

3.3. CoIP and PRISMA provide complementary information about the claudin family interactome.

Due to the technical differences of CoIP and PRISMA experiments, only 6.6% of the proteins identified in this study were present in both datasets. Thus, by the combination of the two orthogonal techniques, it is possible to obtain complementary information about the claudin family interactome (**Figure 15A**). On one hand, CoIPs are performed in vitro overexpressing the full-length proteins into a system where claudins are normally expressed such as epithelial MDCK-C7 cells. CoIP experiments involve cell lysis and capture with specific anti-GFP nanobodies, allowing for the identification of strong and stable interactions that resist the action of detergents and physical disruption of the cell. On the other hand, PRISMA gives information about less stable interactions that occur between proteins or protein complexes and linear motifs present in the unstructured cytosolic tail of claudins. These interactions are transient and very often regulated by PTMs. The enrichment analysis of both datasets shows the different nature of the interacting proteins detected by each technique (**Figure 15B**). The top 15 most enriched GO Cellular Component categories show that CoIPs mostly allow for the detection of proteins related to membrane-bound organelles whereas PRISMA favors the identification of cytosolic proteins or interactors that are part of biomolecular complexes.

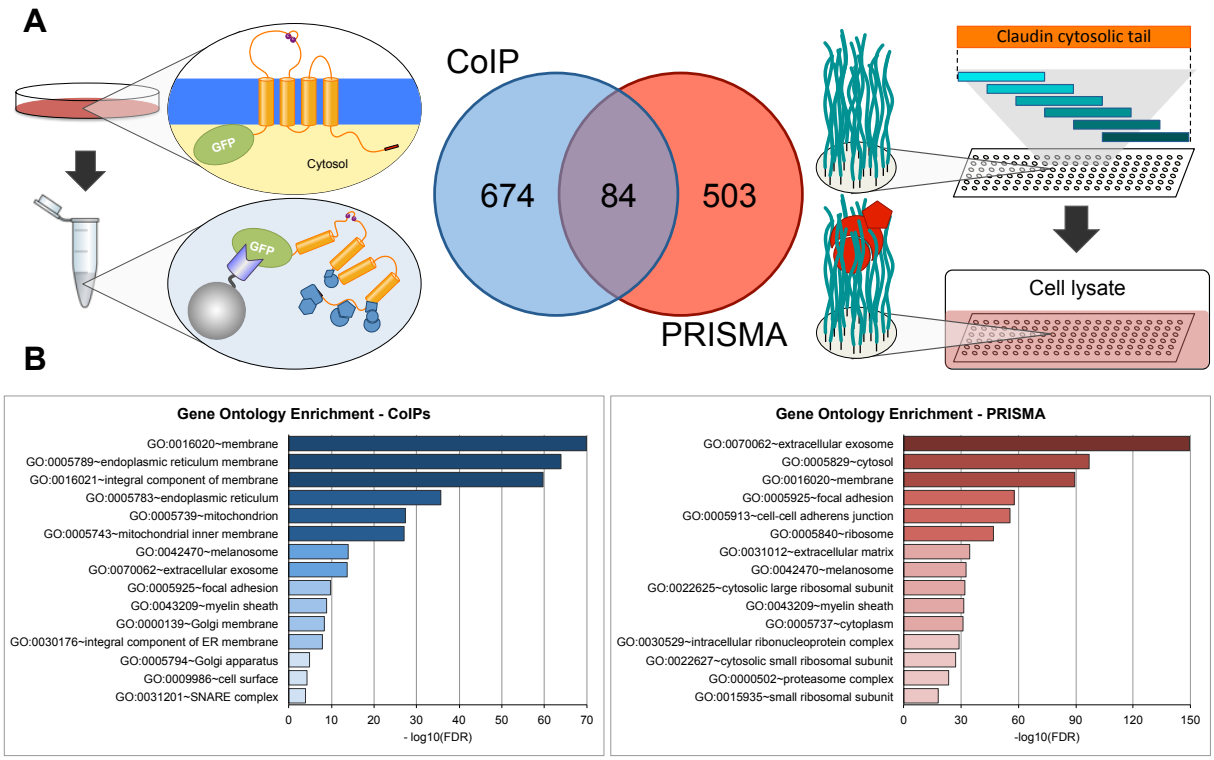


Figure 15. CoIPs and PRISMA provide complementary information about proteins from different cell compartments. A. Differences in the experimental set up of the CoIPs and PRISMA lead to an overlap of a 6.6% (84 proteins) of the significant interactions identified by the two techniques (PRISMA interactions with adj-pval<0.05 before filtering by intensity profile). **B.** Gene Ontology Enrichment Analysis (DAVID Functional Annotation Tool 6.8) of the CoIP and PRISMA significant datasets.

3.4. PRISMA-method as a tool for identification of proteins interacting with disorganized region

In parallel to this project, the PRISMA method was optimized to provide a universal workflow and a more efficient protocol for various applications. Various PRISMA experimental conditions were tested using prototypic unmodified, phosphorylated, and mutated peptides derived from the epidermal growth factor receptor (EGFR) as a reference. The C-terminus of EGFR has been shown to interact with the adaptor protein GRB2 upon tyrosine (Y1092) phosphorylation in bead-based peptide-protein pull-downs and this interaction was lost by a mutation located two amino acids downstream the phosphorylation site (N1094A) (Schulze and Mann, 2004). These results were reproduced using PRISMA and the protocol was optimized in a stepwise manner by modifying experimental parameters including protein lysate concentration, incubation time, LC gradient, and washing temperature. The protocol with the optimal settings was applied to a larger group of peptides previously used in peptide pull-downs to confirm the reproducibility of the new conditions (**Figure 16**).

The first parameter tested was the effect of protein lysate concentration on the number of identified peptides and on the GRB2 LFQ intensity, as an indication of background binding versus the specific GRB2 signal of the EGFR-pY 1092 peptide. A protein concentration of 4 mg/ml was sufficient to obtain strong signals with a good signal-to-noise ratio (**Figure 17A, left panels**). The effect of the membrane incubation time was also tested since in previous studies it varies between 30 and 120 min. This variation had little effect in terms of GRB2 LFQ intensity, which denotes the robustness of the method. In consequence, it was decided that an incubation time of 20 min to maintain the overall protocol short and reproducible in terms of handling (**Figure 17A, central panels**). Pull-downs are low complexity proteome samples; therefore the length of the LC-gradient influences the number of identified peptides and proteins. However, the GRB2 LFQ intensity and the total number of peptides did not improve after a gradient length of 20 min (**Figure 17A, right panels**).

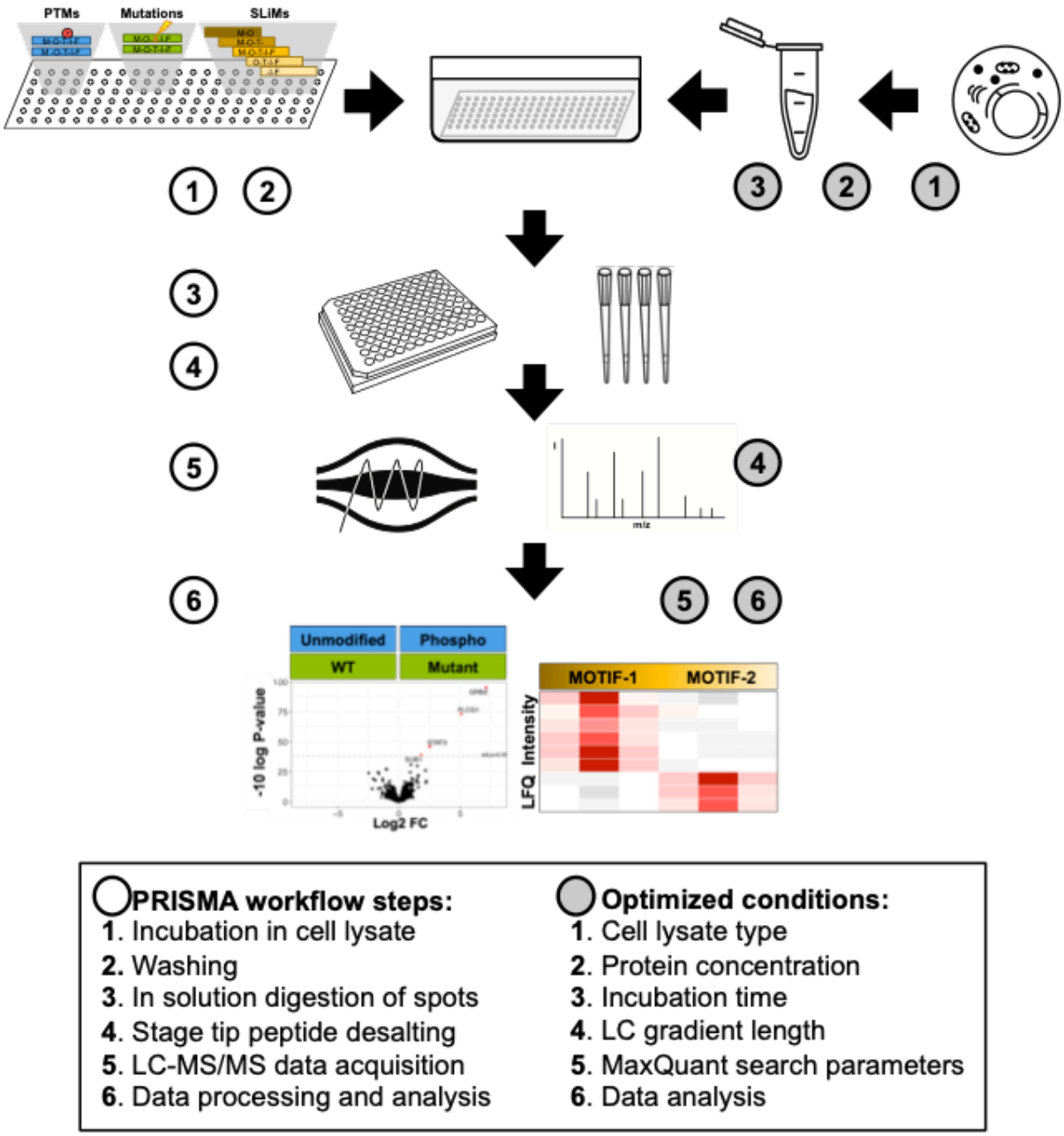


Figure 16. PRISMA workflow and the conditions optimized in the protocol. Modified, mutated, and overlapping peptides were used to optimize the conditions of different steps of the PRISMA protocol.

Additionally, the effect of different washing temperatures was tested for those cases in which cooling is not possible like in automated workflows. Although the LFQ signal of GRB2 was slightly lower than at 4°C, the differential interaction with the EGFR peptides was still identified washing at room temperature (data not shown). In summary, the optimized conditions for the PRISMA protocol are with 4 mg/ml protein lysate, 20 min membrane incubation time, washing at 4°C, and an LC gradient length of 20 min (**Figure 17B**).

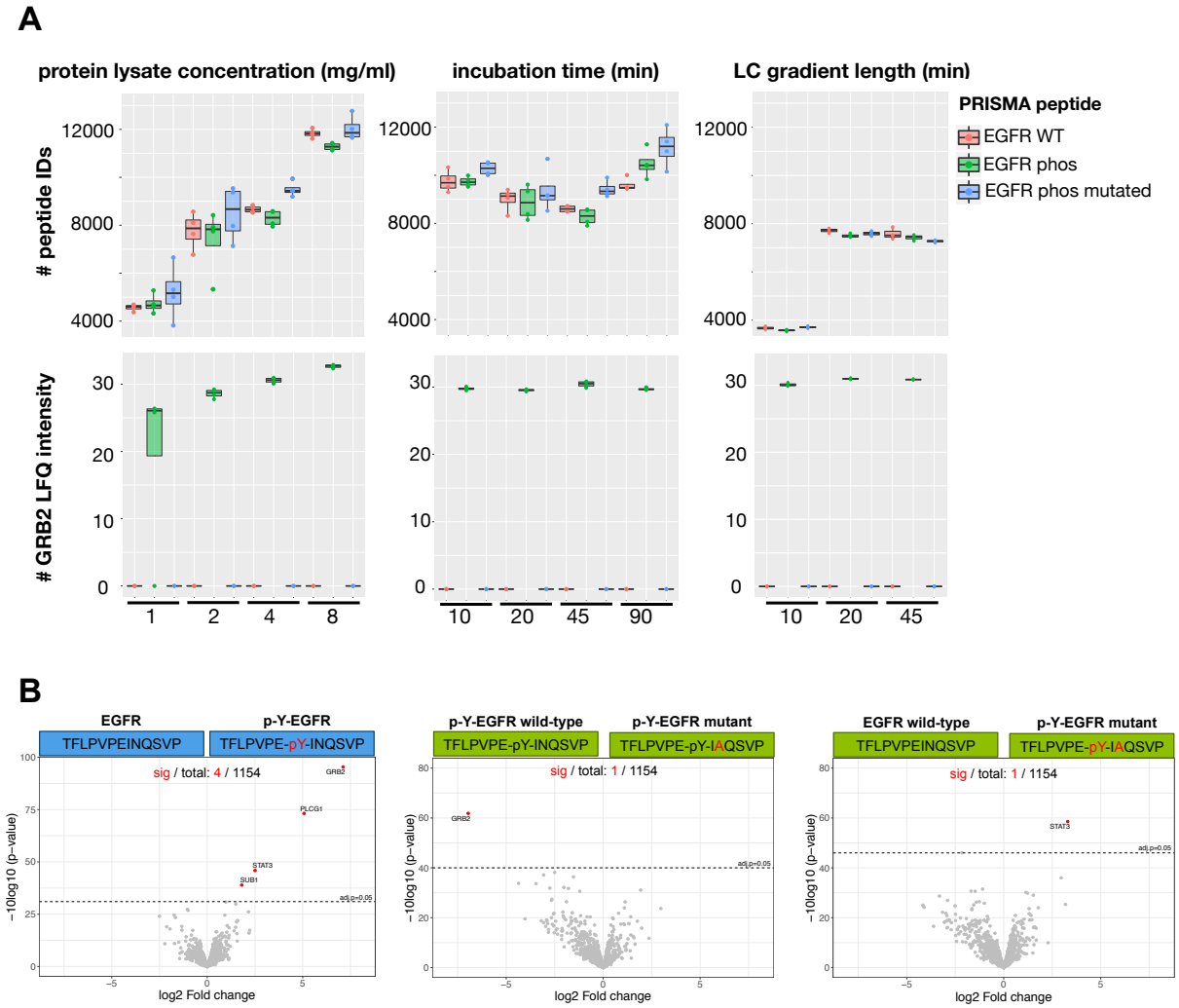


Figure 17. Peptides from EGFR were used to optimize the experimental conditions of PRISMA. A. Number of identified peptides and GRB2 LFQ intensity across peptides from wild-type EGFR, phosphorylated EGFR (Y1092), and phosphorylated mutated EGFR (N1094 -> A) dependent on the conditions tested (protein lysate concentration, incubation time, and LC gradient length). **B**. Volcano plots showing the results of EGFR PRISMA pull-downs with optimal conditions.

Next, the optimized PRISMA protocol was benchmarked using previously described PPI examples including SOS1, GLUT-1, and CEBPB, and extended to map interaction partners of kinase activation loops. The peptide-protein interactions tested were those affected by single amino acid changes in the bait sequence and affected by PTMs. The protocol was also used for mapping SLiM –mediated protein interaction sites by tiling peptides along the unstructured conserved regions of CEBPB.

For the mapping of protein interaction motifs, PRISMA was performed with peptides covering the conserved regions (CR) CR2 and CR7 of the transcription factor CEBPB. Similarly to the study of the interactome of claudins, peptides with overlapping sequences were included. In this case, peptides were designed with a sequence overlap of four amino acids as previously described (Dittmar et al., 2019) and the experiments were performed with commercial nuclear extracts since the majority of the expected CEBPB interactors are nuclear proteins. The data analysis was done following the same pipeline using a t-test-based approach to determine the significant interactions followed by an additional filtering step based on the intensity profile of the interacting proteins across neighboring peptides. Using this optimized workflow we were able to reproduce the findings from Dittmar et al. These results showed components of the mediator of transcription complex and anaphase-promoting complex differentially binding to CEBPB CR2 and CR7 respectively (**Figure 18**).

Finally, the optimized protocol was also implemented on an Agilent Bravo liquid handling platform to further increase the throughput of PRISMA that now allows for the analysis of 60 samples per 24h of machine time in a semi-automated mode.

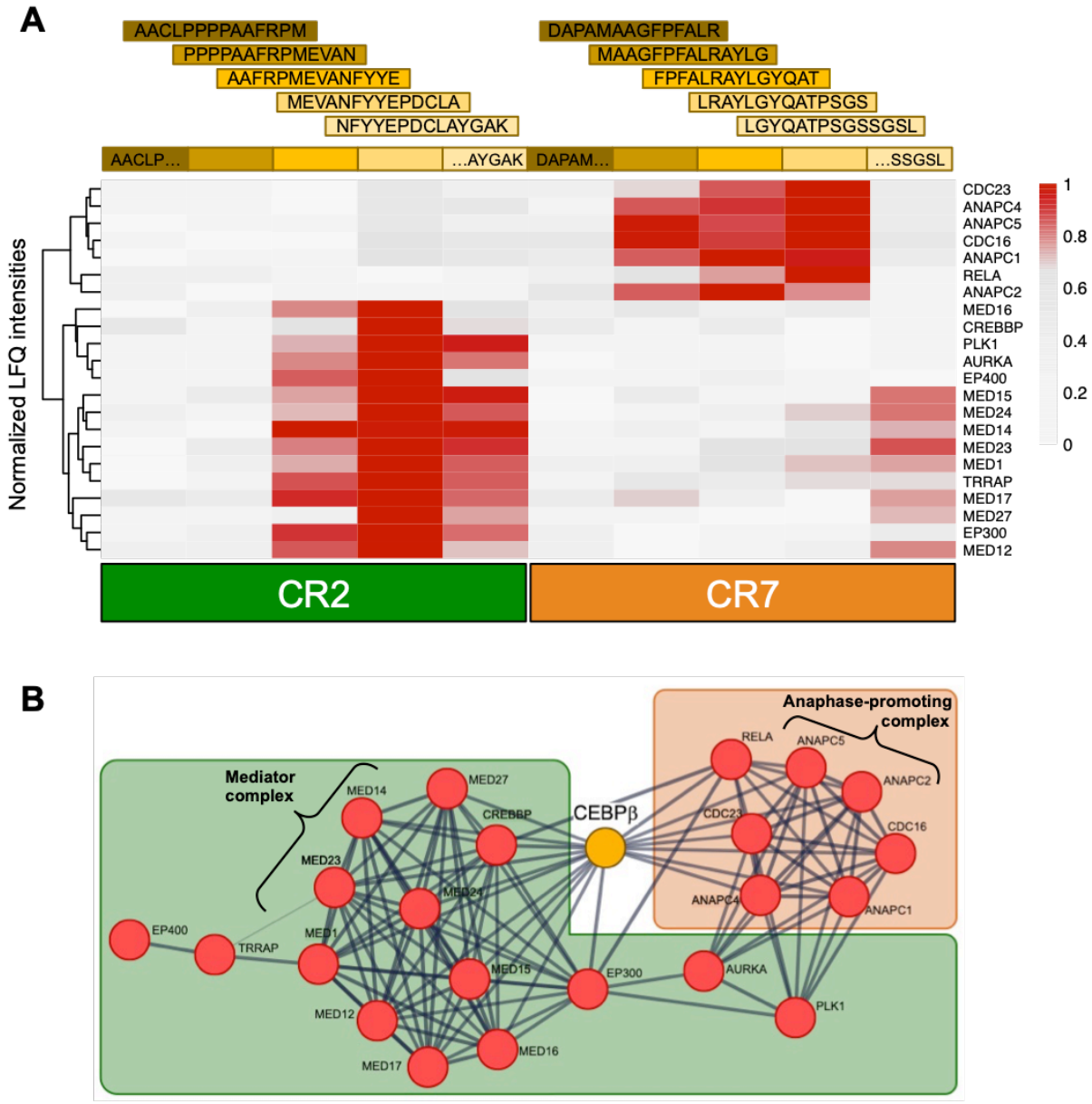


Figure 18. **Results from the optimized PRISMA conditions tested on a larger group of peptides (CEBPβ example). A.** Different groups of proteins specifically interact with overlapping peptides derived from CEBPβ conserved regions (CR2 and CR7) as previously described by Dittmar *et al.*, 2019. **B.** String BD network of CEBPβ known interactions including members of the transcription Mediator complex and the anaphase-promoting complex/cyclosome (APC/C).

4. Discussion

4.1. Importance of MDCK cell lines in TJ studies

The Madin-Darby Canine Kidney (MDCK) cell line is a prototypical polarized epithelial cell line widely used to study epithelial development and function, and one of the few immortalized renal epithelial cells. MDCK cells were established by Madin and Darby in 1958 and characterized for the first time in 1966 (Gaush et al, 1966). The canine origin of the cell line can be a disadvantage when working with antibodies (usually developed against human or mouse proteins) or protein databases like Uniprot where the *Canis lupus* proteome has 41% less entries (protein counts) than the *Homo sapiens* proteome (The UniProt Consortium, 2021). However, studies on epithelial cell polarity, trafficking, and junctions in vertebrates rely on experiments using MDCK cells due to their clear apico-basolateral polarity, well defined cell junctions, and rapid growth rate. The presence of proper TJs can be crucial for the formation of certain protein complexes or protein-protein interactions. Additionally, MDCK cells are suitable for confocal imaging and will polarize in 2D and 3D cell culture.

There are several (sub)strains of MDCK cells available including the parental heterogeneous cell line derived from the renal collecting duct of a female cocker spaniel, and type I and type II MDCK strains (MDCK I, MDCK II) derived from a low and high passage parental cell line respectively (Dukes et al 2011). These two strains display differences in their transepithelial resistance (TER) indicating differences in the composition of their tight junctions. High TER values (> 4000 $\Omega \cdot cm^2$) in MDCK I cells indicate very tight TJs, whereas lower TER values in MDCK II point at the presence of “leaky” tight junctions (Barker et al, 1981). In this study, MDCK-C7 cells were used to perform both CoIP and PRISMA experiments. The MDCK-C7 cell line corresponds to a subtype cloned from the heterogeneous MDCK parental cell line at passage 57. This cell line shares properties with the MDCK I high resistance subtype and resembles the principal cells (PC) of the renal collecting duct involved in K^+ secretion and Na^+ reabsorption (Geckle et al 1994). Therefore, it is important to keep in mind that some interactions described in this study may be directly connected to the characteristics of this particular phenotype.

4.2. Claudin family interactome identified by CoIP

Co-IP results provide a long list of new interactors for the different claudins. Within this long list, 13 proteins annotated as tight junction proteins are found that serve as positive controls such as tight junction protein ZO-1 (ZO-1 or TJP1), occludin (OCLN), MAGUK p55 subfamily member 7 (MPP7), InaD-like protein (PATJ), and claudin-3 and -7 (CLDN3, CLDN7) among others. Epithelial cell adhesion molecule (EpCAM, EPCAM) is also annotated as TJ protein and interacts with claudin-7 in both non-transformed tissues of the gastrointestinal tract and metastasizing tumor cell lines. However, co-localization of the two proteins was described as more prominent at the basolateral membranes (Ladwein et al., 2005). Several studies showed claudin-7 as part of macromolecular complexes forming focal adhesions along the basolateral membranes (Hagen, 2017), and intestinal cells of EpCAM-deficient mice and claudin-7 deficient mice show similar phenotypes with intestinal epithelial cells unable to attach to the underlying mucosa (Kozan et al., 2015). In this study, EpCAM was found interacting with claudins-8, -14, and -19B suggesting that these claudins may be part of similar complexes outside TJs. The fact that in this study EpCAM was not found as a significant interactor of claudin-7 (adj.p-val=0.08) may be a result of the overexpressed claudin-7 having to compete with the endogenous claudin-7 present in MDCK-C7 cells.

Since claudins can be also found in regions of the plasma membrane outside the tight junctions like in the apical plasma membrane, the basolateral plasma membrane, and as part of the focal adhesion complex, GOCC annotations related to these terms are also relevant for the analysis of claudin interactors. In the CoIP dataset 50 of the identified proteins are annotated as part of other cell-cell junctions besides tight junction, and a total of 184 proteins are annotated as plasma membrane proteins. It is worth mentioning the identification of the glucose transporter GLUT-1 (SLC2A1) as an interactor of almost all claudins (except claudin-3, -11, -19A, and -24)) annotated as cell junction protein. This facilitative glucose transporter is in charge of the basal glucose intake, mostly expressed in the blood-brain barrier (BBB) where it assures glucose transport into the brain (Klepper et al., 1999). Therefore, the interaction between GLUT-1 and claudins expressed In the BBB (mainly claudin-5) is somehow expected. In the kidney, glucose uptake is mostly regulated by the two Na+-glucose co-transporters SGLT2 and SGLT1

(SLC5A2 and SLC5A1), expressed in the apical brush border of epithelial cells in the early and late proximal tubule of the nephron respectively, and the facilitative glucose transporter GLUT2 (SLC2A2) in the basolateral membrane in the proximal convoluted tubules. The Na^+-glucose cotransport is electrogenic and requires transcellular K^+ secretion or Cl^- and K^+ paracellular reabsorption (regulated by claudins in the TJs) to stabilize the membrane potential. In this process, basolateral GLUT-1 has been proposed to support glucose reabsorption in the proximal tube and, more importantly, take up glucose in further distal tubule segments for energy supply (Vallon, 2020). This idea is consistent with the strongest renal expression of GLUT-1 being found in the basolateral membrane of further distal tubule segments, with the highest levels detected in connective segments and collecting ducts in rat kidneys (Thorens et al., 1990). The interaction of GLUT-1 with the majority of claudins might be specific for the MDCK-C7 cell line that, as previously mentioned, resembles the phenotype of the principal cells in the collecting duct. Furthermore, these interactions point in the direction of a possible role of claudins in glucose homeostasis beyond the paracellular control of Cl^- and K^+ reabsorption and outside the TJ.

As part of the tight junction recycling and remodeling as well as claudin degradation process, claudins can be also localized on cytosolic vesicles. In this work, 54 different proteins were found related to internalization and lysosomal degradation. 12 of these proteins were also present in the proximity quantitative proteomics dataset from (Tan et al., 2020) including 5 small GTPases of the Rab family (RAB5B, RAB5C, RAB7A, RAB11FIP1, and RAB35) involved in different internalization pathways that lead to vesicle recycling or degradation by the lysosome and can be now connected to specific claudins.

Being tetraspan transmembrane proteins, claudins are translated into and processed by the endoplasmic reticulum and Golgi apparatus and then transported in vesicles to different regions of the cell membrane or directly to the tight junctions. A total of 206 proteins related to this process were found interacting with different claudins.

Interestingly enough, interactors from unexpected cell compartments such as the nucleus and mitochondria were also identified by CoIP (47 and 91 proteins respectively). Due to the nature of the CoIP experiments, this could be an artifact derived from cell lysis. However, certain claudins like claudin-10b can be found in the invaginations of the basal membrane of epithelial cells in the thick ascending limb of

Henle (TAL) of kidneys where a high number of mitochondria are required for energy-dependent transport (Milatz and Breiderhoff, 2017). Therefore, the interaction of claudins with some of the mitochondrial proteins identified may have an actual biological relevance.

In the case of nuclear proteins, there is solid evidence of claudins localizing to the nucleus in cancer cells. For example, claudin-1 nuclear localization has been shown in tissues from patients with primary colon cancer, in transformed nasopharyngeal carcinoma cells, in melanoma cells, and in thyroid carcinoma cells (Dhawan et al., 2005; Lee et al., 2010; French et al., 2009; Zwanziger et al., 2015). Other claudins found localizing to the nucleus are claudin-2 in human lung adenocarcinoma cells (Ikari et al., 2014), claudin-3 in breast cancer cells (Todd et al., 2015), and claudin-4 in endometrial cancer cells (Cuevas et al., 2015). Also, molecular studies demonstrate a direct transcriptional role for nuclear claudin-1 in E-cadherin expression in cultured colon cancer cell lines (Dhawan et al., 2005). Additionally, an *in silico* analysis with cNLS mapper (Kosugi et al., 2009), a prediction tool for importin α-dependent nuclear localization signals (NLS), showed that many mouse claudins contain putative NLS (Hagen, 2017). Among the 47 proteins identified annotated as nuclear, three are nuclear pore complex proteins (NUP88, NUP160, and NUP205) that might be involved in the recognition of such NLS and translocation of claudins to the nucleus.

The study of the claudin family using CoIPs provides a wide map of interactions that expands the current knowledge in the field. This will serve as a resource for future studies focused on any of the different aspects of the claudins whether it is their role as tight junction proteins, their implications in different cell-cell contact zones outside the tight junctions, their processing in the endoplasmic reticulum, vesicular trafficking, turn-over and degradation through different pathways, and their least canonical functions in nucleus and related to mitochondria.

4.3. Interactions with the cytosolic tail of claudins

PRISMA dataset revealed different interaction patterns between 148 proteins and the cytosolic tail of the different members of the claudin family. These patterns not only show the claudin but also the region within the cytosolic tail where the interaction is taking place with a resolution of 5 aa. By including PTM containing

peptides to the PRISMA matrix it was also possible to identify interactions that are regulated by phosphorylation of the PDZ domain-binding motif present in a subset of the claudins.

4.3.1. Biological relevance of protein complexes interacting with the cytosolic tail of claudins

After comparing the present dataset with a previous quantitative proximity proteomics study (Tan et al., 2020) the attention was focused into three groups of interacting proteins that differentially bind to the claudins: two protein complexes (proteasome and CCT/TRiC) and the 14-3-3 protein family.

4.3.1.1. The CCT/TriC chaperonin complex

The CCT/TriC chaperonin complex is in charge of the folding of approximately 10% of the cytosolic proteins including cytoskeleton proteins from the actin and tubulin families, cell cycle regulators, and signaling proteins (Jin et al., 2019). It comprises two hetero-oligomeric rings with eight subunits each (CCT1 to CCT8) and each ring provides a central cavity for ATP-driven protein folding.

Although the CCT/TRiC complex was found interacting with several claudins and PLA experiments confirm the differential interaction of the subunits beta and zeta (CCT2 and CCT6A) with claudin-3 versus claudin-1 and -12, the role of this interaction in the biology of the cell remains unknown.

WD-repeat proteins have been described as an important family of CCT/TriC substrates. The common feature among the members of this large protein family is coordinating multi-protein complex assemblies where the repeating units of the core sequence, of approximately 40 amino acids (WD40-repeats, WD, or beta-transducin repeats), serve as a scaffold for protein interactions that can occur simultaneously with different proteins (Li and Roberts, 2001). WD-repeat proteins are involved in a wide range of cellular functions including signal transduction, cytoskeletal assembly, and regulation of vesicular trafficking among many others. One of the cytoskeletal proteins found as interacting with claudins in the PRISMA dataset is the WD repeat-containing protein 1 (WDR1 or Actin-interacting protein (Aip) 1), involved in the assembly and maintenance of apical cell junctions in epithelial cells by regulating the F-actin dynamics (Lechuga et al., 2015). Looking at the pre-filtered PRISMA

significant interacting proteins WD-repeat proteins involved in signal transduction were also found, such as G-protein subunits beta 1 (GNB1, where the repeating unit was first described), GNB2, and receptor of activated protein C kinase 1 (RACK1); as well as the coatomer subunits alpha and beta (COPA, COPB2) which are known WD-repeat proteins that mediate intracellular vesicular trafficking (Li and Roberts, 2001). Other significant interactors annotated as WD-repeat proteins (InterPro domain IPR001680) are the Actin-related protein 2/3 complex subunit 1B (ARPC1B), component of the multiprotein complex Arp2/3 that mediates actin polymerization; the eukaryotic translation initiation factor 3 subunit I (EIF3I); the Serine-threonine kinase receptor-associated protein (STRAP) which plays a role in the cellular distribution of the spliceosome-related SMN complex; and the histone binding protein RBBP7.

The identification of CCT/TRiC complex and known substrate proteins suggests that in the context of the tight junctions, the presence of this cytosolic chaperonin complex could use the cytosolic tail of claudin-2, -3, -4, -5, -6, -9, -10, and -23 as physical support for performing its biological task of folding of newly synthesized cytosolic proteins that take part in the many dynamic cell processes related to apical junction complex formation and remodeling

4.3.1.2. The proteasome complex

The proteasome complex is responsible for the degradation of intracellular proteins in a process that requires ubiquitination. Ubiquitin (Ub) is a 76 amino acid protein expressed in all tissues and ubiquitination is the process of ligating one or more Ub molecules to a substrate protein. This process consists of an ATP-dependent reaction that requires the action of three different enzymes: E1 (Ub-activating), E2 (Ub-conjugating), and E3 (Ub-ligating). Ub E3 ligases selectively mark a protein substrate for degradation by attachment of Ub to its lysine (K) residues. Monoubiquitination is involved in transcriptional regulation, DNA damage repair, and membrane-associated endocytosis (Nakagawa and Nakayama, 2015) whereas polyubiquitination (attachment of a chain of Ub) is the signal for protein degradation by the proteasome (Chau et al., 1989; Thrower, 2000).

The ubiquitin/proteasome system (UPS) has been shown to regulate the fate of various membrane proteins by mono- or polyubiquitination leading to endocytosis

and recycling/degradation via the lysosome or degradation by the proteasome (Cai et al., 2018). Adherens junction (AJ) and tight junction (TJ) transmembrane proteins such as cadherins and come claudins are known to be monoubiquitinated and targeted for lysosomal degradation. Within the claudin family, only the ubiquitination of claudin-1, -2, -4, -5, -8, and -16 have been characterized, with E3 ubiquitin ligases LNX1p80 (claudin-1, -2, and -4) and KLHL3 (claudin-8) involved in their targeting to lysosomal degradation in MDCK cells and collecting duct of mouse kidney respectively (Takahashi et al., 2009; Gong et al., 2015). Ub E3 ligase PDZRN3 catalyzes claudin-16 mono-ubiquitination and translocalization regulating paracellular Mg^{2+} permeability in the kidney (Marunaka, 2017). In the cases of claudin-5 and occludin, it appears that they can be degraded both in an Ub-proteasome-dependent manner and by the Ub-independent lysosomal pathway (Traweger et al., 2002; Mandel et al., 2012).

On the other hand, cytosolic proteins that are part of the cell-cell junctions such as linkers are mostly polyubiquitinated and degraded via the proteasome. One example is p120-catenin, which plays an essential role in the maintenance of the integrity of AJs and TJs and its reduction correlates with the progression of different tumors as well as inflammatory diseases. p120-catenin can be regulated by either calpain-1-mediated degradation or phosphorylation-dependent ubiquitination and proteasomal degradation (Cai et al., 2018). Tight junction protein 1 (TJP1, ZO-1) also undergoes ubiquitination by E3 ligase component N-recognin-1 (Ubr1) followed by protein degradation in response to interleukin-6 (IL-6) leading to disruption of the endothelial barrier integrity in cultured brain microvascular endothelial cells (Chen et al., 2014)

The identification of various subunits of the proteasome complex interacting with unmodified peptides from the C-terminal region of claudins indicates that the role of these interactions is not the degradation of the claudin itself but cytosolic ubiquitinated proteins in the vicinity. In addition, there is evidence showing the role of UPS and local protein degradation in synaptic plasticity (Hegde, 2004). Similarly, tight junctions are very dynamic structures with constant trafficking and recycling of proteins between the cytoplasm and the cell surface. The present findings suggest that the proteasome complex uses the cytosolic tail of claudin-1 to -10, -18, and -23 as a scaffold for localized degradation of tight junction cytosolic proteins contributing to the fast turnover that regulates paracellular transport in epithelia in response to physiological variations.

4.3.1.3. The 14-3-3 protein family

14-3-3 proteins are a family of highly conserved acidic proteins with at least 7 isoforms in mammals. They are mainly cytosolic proteins but can move freely to the nucleus and they are involved in regulating signal transduction pathways, apoptosis, adhesion, cellular proliferation, differentiation, and survival. 14-3-3 proteins can form homo- or heterodimers and interact with various cellular proteins both in a phosphorylation-dependent or -independent manner. Some of the 14-3-3 binding proteins are involved in the regulation of the cytoskeleton, GTPase function, membrane signaling, and cell fate determination (Jin et al., 2004).

One of the many roles in which 14-3-3 proteins are involved in membrane protein transport. In their review on the topic Mrowiec and Schwappach summarize the numerous studies that show the interaction of the different 14-3-3 isoforms to membrane proteins and the motifs involved in these interactions. The effect of these interactions varies from correlation with the cell surface expression of the ligand protein (for membrane channels or cell surface receptors) to activation or inhibition of pump activity in the case of plasma membrane ATPases. All the proteins listed are multimers and, in addition to a 14-3-3 binding motif, they also contain a coatomer protein complex I- (COPI-) interacting signal. In some cases, the COPI interacting signal overlaps or is very close to the 14-3-3 binding motif whereas in others it can be on the opposite end of the protein. This is the case of the KCNK3 potassium channel where the COPI interacting motif consists of two basic amino acids at the cytosolic N-terminus of the membrane protein and the 14-3-3 binding motif is present at the distal C-terminus of the protein (O'Kelly et al., 2002; Rajan et al., 2002).

As already mentioned in the results section, 14-3-3 proteins and their ligands often interact via motifs that include phosphorylated Ser or Thr residues (Mrowiec and Schwappach, 2006), and the canonical 14-3-3 binding phosphopeptide motif (LIG_1433_CanoR_1) is found in 13 claudins. However, PRISMA shows that the 7 isoforms of the 14-3-3 protein family can interact with the cytosolic C-terminal tail of claudin-1 to -10, -18, and -23 in a phosphorylation-independent manner through novel interaction motifs in most cases.

The molecular models of action for 14-3-3 proteins include clamping (stabilization of a certain conformation of the ligand), masking (blocking the access of another interacting protein), and scaffolding (recruiting additional proteins o molecules acting

as a backbone for protein complex assembly). Comparing their PRISMA interaction patterns with the ones from the protein complexes previously discussed it seems logical to think that the different 14-3-3 isoforms could serve as a scaffold to link the proteasome and the CCT/TRiC chaperonin complexes (as well as cytoskeletal proteins also identified by PRISMA), to the unstructured cytosolic tail of claudins. On the other hand, as it has been shown for other transmembrane proteins, the 14-3-3 family could be also implicated in the targeting of some claudins to the plasma membrane. Thus, further studies are necessary to elucidate the role that these interactions play in the biology of claudins and tight junctions.

4.3.2. Local protein folding and degradation may influence dynamic changes in tight junctions

Local translation and degradation play a crucial role in highly polarized cells like oocytes, early embryos, and neurons. This allows for a rapid response in subcellular domains with highly dynamic processes such as synapsis and embryonic development. However, this phenomenon is not exclusive to highly polarized cells. Localization of ZO-1 and β-actin mRNAs to TJ and AJ respectively regulate diverse aspects of cell adhesion (Katz et al., 2016; Nagaoka et al., 2012), indicating that spatially controlled translation also takes place in epithelial cells where acute cellular responses regulate the dynamic changes that control TJ remodeling and paracellular transport. To maintain the steady state of concentrated proteins necessary to create such subcellular domains, mRNA and protein complexes involved have to be either anchored in place or subjected to continuous active transport. In dendritic spines of rat hippocampal neurons, the proteasome complex is sequestered by the actin cytoskeleton (Bingol and Schuman, 2006). Considering the previous evidence and the identification by PRISMA and subsequent confirmation of the interaction of the cytosolic tail of some claudins with the proteasome and the CCT/TRiC chaperonin complexes, it seems logical to hypothesize that localized protein folding and protein degradation by the proteasome also occur in TJs. The cytosolic tails of certain claudins would therefore act as a scaffold facilitating a contact point for the protein complexes involved.

4.3.3. PTM regulated interactions

Many aspects of the biological activity of claudins are regulated by PTMs including protein-protein interactions, oligomer assembly, subcellular localization, trafficking, and net claudin homeostasis (Findley and Koval, 2009). Although there is evidence of regulation by palmitoylation and ubiquitination, the vast majority of observed PTMs in claudins are phosphorylations. Prediction studies indicate that there are up to 10 possible phosphorylation sites depending on the claudin and most of them are located in the C-terminal cytosolic tail (González-Mariscal et al., 2010).

The Tyr residue present at the position -1 of most claudins is a conserved putative Eph phosphorylation site (González-Mariscal et al., 2010). When phosphorylated, this residue regulated the binding activity of the PDZ domain-binding motif of claudins. For example, in HT29 colon carcinoma cells, EphA2 activation leads to claudin-4 Tyr -1 phosphorylation and as a result, claudin-4 can no longer integrate efficiently to the cell borders delaying the formation of TJs (Tanaka et al., 2005). Another tyrosine residue located a few amino acids upstream (Y_{-6}) also influences the activity of the PDZ domain-binding motif when phosphorylated leading to internalization of some claudins (Nomme et al., 2015). PRISMA shows how phosphorylation of the two residues not only disrupts the interaction with tight junction protein ZO-1/-2 or -3 but also allows the PDZ domain-binding motif to differentially interact with new proteins depending on the phosphorylated site. Some of the new interacting partners are involved in claudin internalization by different pathways. Eph receptors control tissue development and are frequently overexpressed in cancerous tissues, and cell-cell junction disruption is one of the hallmarks of EMT in epithelial cancers that ultimately end in tumor metastasis. Therefore, the phosphorylation state of Tyr in positions -1 and -6 of claudins could be used as markers for cancer progression and tumor prognosis, and potential therapeutic targets.

4.4. The Claudinome: claudin family interactome landscape

Since their discovery in 1998 by Furuse and collaborators, claudins have been thoroughly studied especially in the context of tight junctions and control of the paracellular transport. However, the mechanisms underlying the biology of claudins are still not well understood. The extended literature in the field shows the high versatility of this protein family that depends on the context in which they are studied. Moreover, many studies focus on how a particular subset of claudins behaves in a particular setting, which is often difficult to extrapolate to the rest of the family. Therefore, a systematic study of the entire claudin protein family with a standard setup for all of them provides a general overview that helps understand the differences and similarities between the different members while giving additional insight into the least studied ones.

The combination of the complementary information obtained from the two different approaches applied in this study generates a first comprehensive interactome landscape of the claudin protein family now termed the claudinome. This allows for a better understanding of their implication in several processes within the cell. The new claudinome can serve as a resource for future studies related to the many aspects of their biology inside and outside the tight junction, as well as their role in the various pathologies they are related to.

4.5. Conclusions and outlook

Even though the importance of claudins as modulators of epithelial barrier integrity as well as transepithelial transport has been proven in the past, the underlying mechanisms of their regulation are still not fully understood. Their implication in pathologies of different nature makes it essential to further investigate and comprehend their biology and how they behave inside and outside the tight junctions.

In this work, a combination of peptide-based (PRISMA) and full-length protein (CoIP) interaction proteomics creates an extended map of interactions across the entire claudin protein family that provides complementary information about their regulation. New interacting proteins from different cell compartments can be now connected to individual members of the claudin family and new possible functions. Novel claudin-claudin heterotypic interactions were also identified. Protein complexes interacting with the cytosolic unstructured tail of specific claudins were found by PRISMA, leading to a new hypothesis about their possible implication in localized biological processes in epithelial cells. PRISMA also provided information about new interactions with the PDZ domain-binding motif of claudins that are regulated by PTMs.

Finally, this study also shows how PRISMA is a valuable tool that can be applied to study the different aspects of protein interactions with disordered regions in a robust and automated manner.

5. References

Aebersold, R., Mann, M., 2003. Mass spectrometry-based proteomics. Nature 422, 198–207. https://doi.org/10.1038/nature01511

Bingol, B., Schuman, E.M., 2006. Activity-dependent dynamics and sequestration of proteasomes in dendritic spines. Nature 441, 1144–1148. https://doi.org/10.1038/nature04769

Bondarenko, P.V., Chelius, D., Shaler, T.A., 2002. Identification and Relative Quantitation of Protein Mixtures by Enzymatic Digestion Followed by Capillary Reversed-Phase Liquid Chromatography−Tandem Mass Spectrometry. Anal. Chem. 74, 4741–4749. https://doi.org/10.1021/ac0256991

Bürgi, J., Xue, B., Uversky, V.N., van der Goot, F.G., 2016. Intrinsic Disorder in Transmembrane Proteins: Roles in Signaling and Topology Prediction. PLoS ONE 11, e0158594. https://doi.org/10.1371/journal.pone.0158594

Cai, J., Culley, M.K., Zhao, Y., Zhao, J., 2018. The role of ubiquitination and deubiquitination in the regulation of cell junctions. Protein Cell 9, 754–769. https://doi.org/10.1007/s13238-017-0486-3

Chau, V., Tobias, J.W., Bachmair, A., Marriott, D., Ecker, D.J., Gonda, D.K., Varshavsky, A., 1989. A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. Science 243, 1576–1583. https://doi.org/10.1126/science.2538923

Chen, C.-J., Ou, Y.-C., Li, J.-R., Chang, C.-Y., Pan, H.-C., Lai, C.-Y., Liao, S.-L., Raung, S.-L., Chang, C.-J., 2014. Infection of Pericytes In Vitro by Japanese Encephalitis Virus Disrupts the Integrity of the Endothelial Barrier. Journal of Virology 88, 1150–1161. https://doi.org/10.1128/JVI.02738-13

Cornish, J., Chamberlain, S.G., Owen, D., Mott, H.R., 2020. Intrinsically disordered proteins and membranes: a marriage of convenience for cell signalling? Biochemical Society Transactions 48, 2669–2689. https://doi.org/10.1042/BST20200467

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., Mann, M., 2014. Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. Molecular & Cellular Proteomics 13, 2513–2526. https://doi.org/10.1074/mcp.M113.031591

Coyne, C.B., Gambling, T.M., Boucher, R.C., Carson, J.L., Johnson, L.G., 2003. Role of claudin interactions in airway tight junctional permeability. American Journal of Physiology-Lung Cellular and Molecular Physiology 285, L1166–L1178. https://doi.org/10.1152/ajplung.00182.2003

Cuevas, M.E., Gaska, J.M., Gist, A.C., King, J.M., Sheller, R.A., Todd, M.C., 2015. Estrogen-dependent expression and subcellular localization of the tight junction protein claudin-4 in HEC-1A endometrial cancer cells. Int J Oncol 47, 650–656. https://doi.org/10.3892/ijo.2015.3030

Daugherty, B.L., Ward, C., Smith, T., Ritzenthaler, J.D., Koval, M., 2007. Regulation of Heterotypic Claudin Compatibility. Journal of Biological Chemistry 282, 30005–30013. https://doi.org/10.1074/jbc.M703547200

De Benedetto, A., Rafaels, N.M., McGirt, L.Y., Ivanov, A.I., Georas, S.N., Cheadle, C., Berger, A.E., Zhang, K., Vidyasagar, S., Yoshida, T., Boguniewicz, M., Hata, T., Schneider, L.C., Hanifin, J.M., Gallo, R.L., Novak, N., Weidinger, S., Beaty, T.H., Leung, D.Y., Barnes, K.C., Beck, L.A., 2011. Tight Junction Defects in Atopic Dermatitis. J Allergy Clin Immunol 127, 773-786.e7. https://doi.org/10.1016/j.jaci.2010.10.018

Dhawan, P., Singh, A.B., Deane, N.G., No, Y., Shiou, S.-R., Schmidt, C., Neff, J., Washington, M.K., Beauchamp, R.D., 2005. Claudin-1 regulates cellular transformation and metastatic behavior in colon cancer. J Clin Invest 115, 1765–1776. https://doi.org/10.1172/JCI24543

Dittmar, G., Hernandez, D.P., Kowenz-Leutz, E., Kirchner, M., Kahlert, G., Wesolowski, R., Baum, K., Knoblich, M., Hofstätter, M., Muller, A., Wolf, J., Reimer, U., Leutz, A., 2019. PRISMA: Protein Interaction Screen on Peptide Matrix Reveals Interaction Footprints and Modifications- Dependent Interactome of Intrinsically Disordered C/EBPβ. iScience 13, 351–370. https://doi.org/10.1016/j.isci.2019.02.026

Dunker, A.K., Cortese, M.S., Romero, P., Iakoucheva, L.M., Uversky, V.N., 2005. Flexible nets. The roles of intrinsic disorder in protein interaction networks. FEBS J 272, 5129–5148. https://doi.org/10.1111/j.1742-4658.2005.04948.x

Eng, J.K., Searle, B.C., Clauser, K.R., Tabb, D.L., 2011. A face in the crowd: recognizing peptides through database search. Mol Cell Proteomics 10, R111.009522. https://doi.org/10.1074/mcp.R111.009522

Evans, M.J., von Hahn, T., Tscherne, D.M., Syder, A.J., Panis, M., Wölk, B., Hatziioannou, T., McKeating, J.A., Bieniasz, P.D., Rice, C.M., 2007. Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. Nature 446, 801–805. https://doi.org/10.1038/nature05654

Feldmeyer, L., Huber, M., Fellmann, F., Beckmann, J.S., Frenk, E., Hohl, D., 2006. Confirmation of the origin of NISCH syndrome. Human Mutation 27, 408–410. https://doi.org/10.1002/humu.20333

Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., Whitehouse, C.M., 1989. Electrospray ionization for mass spectrometry of large biomolecules. Science 246, 64–71. https://doi.org/10.1126/science.2675315

Findley, M.K., Koval, M., 2009. Regulation and roles for claudin-family tight junction proteins. IUBMB Life 61, 431–437. https://doi.org/10.1002/iub.175

Frank, R., 1992. Spot-synthesis: an easy technique for the positionally addressable, parallel chemical synthesis on a membrane support. Tetrahedron 48, 9217–9232. https://doi.org/10.1016/S0040-4020(01)85612-X

Fredriksson, S., Gullberg, M., Jarvius, J., Olsson, C., Pietras, K., Gústafsdóttir, S.M., Östman, A., Landegren, U., 2002. Protein detection using proximity-dependent DNA ligation assays. Nat Biotechnol 20, 473–477. https://doi.org/10.1038/nbt0502-473

French, A.D., Fiori, J.L., Camilli, T.C., Leotlela, P.D., O'Connell, M.P., Frank, B.P., Subaran, S., Indig, F.E., Taub, D.D., Weeraratna, A.T., 2009. PKC and PKA phosphorylation affect the subcellular localization of claudin-1 in melanoma cells. Int J Med Sci 6, 93–101. https://doi.org/10.7150/ijms.6.93

Furuse et al. - 1998 - A Single Gene Product, Claudin-1 or -2, Reconstitu.pdf, n.d.

Furuse, M., Sasaki, H., Fujimoto, K., Tsukita, S., 1998. A Single Gene Product, Claudin-1 or -2, Reconstitutes Tight Junction Strands and Recruits Occludin in Fibroblasts. Journal of Cell Biology 143, 391–401. https://doi.org/10.1083/jcb.143.2.391

Gavin, A.-C., Bösche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.-M., Cruciat, C.-M., Remor, M., Höfert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.-A., Copley, R.R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G., Superti-Furga, G., 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 415, 141–147. https://doi.org/10.1038/415141a

Gomes, I., Sierra, S., Devi, L.A., 2016. Detection of Receptor Heteromerization Using In Situ Proximity Ligation Assay. Curr Protoc Pharmacol 75, 2.16.1-2.16.31. https://doi.org/10.1002/cpph.15

Gong, Y., Wang, J., Yang, J., Gonzales, E., Perez, R., Hou, J., 2015. KLHL3 regulates paracellular chloride transport in the kidney by ubiquitination of claudin-8. Proc Natl Acad Sci USA 112, 4340–4345. https://doi.org/10.1073/pnas.1421441112

González-Mariscal, L., Garay, E., Quirós, M., 2010. Chapter 6 - Regulation of Claudins by Posttranslational Modifications and Cell-Signaling Cascades, in: L. Yu, A.S. (Ed.), Current Topics in Membranes. Academic Press, pp. 113–150. https://doi.org/10.1016/S1063-5823(10)65006-5

Günzel, D., Yu, A.S.L., 2013. Claudins and the Modulation of Tight Junction Permeability. Physiological Reviews 93, 525–569. https://doi.org/10.1152/physrev.00019.2012

Hadj-Rabia, S., Baala, L., Vabres, P., Hamel-Teillac, D., Jacquemin, E., Fabre, M., Lyonnet, S., de Prost, Y., Munnich, A., Hadchouel, M., Smahi, A., 2004. Claudin-1 gene mutations in neonatal sclerosing cholangitis associated with ichthyosis: A tight junction disease. Gastroenterology 127, 1386–1390. https://doi.org/10.1053/j.gastro.2004.07.022

Hadj-Rabia, S., Brideau, G., Al-Sarraj, Y., Maroun, R.C., Figueres, M.-L., Leclerc-Mercier, S., Olinger, E., Baron, S., Chaussain, C., Nochy, D., Taha, R.Z., Knebelmann, B., Joshi, V., Curmi, P.A., Kambouris, M., Vargas-Poussou, R., Bodemer, C., Devuyst, O., Houillier, P., El-Shanti, H., 2018. Multiplex epithelium dysfunction due to CLDN10 mutation: the HELIX syndrome. Genet Med 20, 190–201. https://doi.org/10.1038/gim.2017.71

Hagen, S.J., 2017. Non-canonical functions of claudin proteins: Beyond the regulation of cell-cell adhesions. Tissue Barriers 5, e1327839. https://doi.org/10.1080/21688370.2017.1327839

Hegde, A.N., 2004. Ubiquitin-proteasome-mediated local protein degradation and synaptic plasticity. Progress in Neurobiology 73, 311–357. https://doi.org/10.1016/j.pneurobio.2004.05.005

Hornbeck, P.V., Zhang, B., Murray, B., Kornhauser, J.M., Latham, V., Skrzypek, E., 2015. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Research 43, D512–D520. https://doi.org/10.1093/nar/gku1267

Huang, D.W., Sherman, B.T., Lempicki, R.A., 2009a. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4, 44–57. https://doi.org/10.1038/nprot.2008.211

Huang, D.W., Sherman, B.T., Lempicki, R.A., 2009b. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37, 1–13. https://doi.org/10.1093/nar/gkn923

Hubner, N.C., Bird, A.W., Cox, J., Splettstoesser, B., Bandilla, P., Poser, I., Hyman, A., Mann, M., 2010. Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. Journal of Cell Biology 189, 739–754. https://doi.org/10.1083/jcb.200911091

Hung, V., Zou, P., Rhee, H.-W., Udeshi, N.D., Cracan, V., Svinkina, T., Carr, S.A., Mootha, V.K., Ting, A.Y., 2014. Proteomic Mapping of the Human Mitochondrial Intermembrane Space in Live Cells via Ratiometric APEX Tagging. Molecular Cell 55, 332–341. https://doi.org/10.1016/j.molcel.2014.06.003

Iakoucheva, L.M., Radivojac, P., Brown, C.J., O'Connor, T.R., Sikes, J.G., Obradovic, Z., Dunker, A.K., 2004. The importance of intrinsic disorder for protein phosphorylation. Nucleic Acids Res 32, 1037–1049. https://doi.org/10.1093/nar/gkh253

Ikari, A., Watanabe, R., Sato, T., Taga, S., Shimobaba, S., Yamaguchi, M., Yamazaki, Y., Endo, S., Matsunaga, T., Sugatani, J., 2014. Nuclear distribution of claudin-2 increases cell proliferation in human lung adenocarcinoma cells. Biochim Biophys Acta 1843, 2079–2088. https://doi.org/10.1016/j.bbamcr.2014.05.017

Ishiyama, N., Lee, S.-H., Liu, S., Li, G.-Y., Smith, M.J., Reichardt, L.F., Ikura, M., 2010. Dynamic and Static Interactions between p120 Catenin and E-Cadherin Regulate the Stability of Cell-Cell Adhesion. Cell 141, 117–128. https://doi.org/10.1016/j.cell.2010.01.017

Ivarsson, Y., Jemth, P., 2019. Affinity and specificity of motif-based protein-protein interactions. Curr Opin Struct Biol 54, 26–33. https://doi.org/10.1016/j.sbi.2018.09.009

Jin, J., Smith, F.D., Stark, C., Wells, C.D., Fawcett, J.P., Kulkarni, S., Metalnikov, P., O'Donnell, P., Taylor, P., Taylor, L., Zougman, A., Woodgett, J.R., Langeberg, L.K., Scott, J.D., Pawson, T., 2004. Proteomic, Functional, and Domain-Based Analysis of In Vivo 14-3-3 Binding Proteins Involved in Cytoskeletal Regulation and Cellular Organization. Current Biology 14, 1436–1450. https://doi.org/10.1016/j.cub.2004.07.051

Jin, M., Liu, C., Han, W., Cong, Y., 2019. TRiC/CCT Chaperonin: Structure and Function, in: Harris, J.R., Marles-Wright, J. (Eds.), Macromolecular Protein Complexes II: Structure and Function, Subcellular Biochemistry. Springer International Publishing, Cham, pp. 625–654. https://doi.org/10.1007/978-3-030-28151-9_19

Karas, M., Hillenkamp, F., 1988. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem 60, 2299–2301. https://doi.org/10.1021/ac00171a028

Katahira, J., Sugiyama, H., Inoue, N., Horiguchi, Y., Matsuda, M., Sugimoto, N., 1997. Clostridium perfringens Enterotoxin Utilizes Two Structurally Related Membrane Proteins as Functional Receptors in Vivo. Journal of Biological Chemistry 272, 26652–26658. https://doi.org/10.1074/jbc.272.42.26652

Katz, Z.B., English, B.P., Lionnet, T., Yoon, Y.J., Monnier, N., Ovryn, B., Bathe, M., Singer, R.H., 2016. Mapping translation “hot-spots” in live cells by tracking single molecules of mRNA and ribosomes. eLife 5, e10415. https://doi.org/10.7554/eLife.10415

Keilhauer, E.C., Hein, M.Y., Mann, M., 2015. Accurate Protein Complex Retrieval by Affinity Enrichment Mass Spectrometry (AE-MS) Rather than Affinity Purification Mass Spectrometry (AP-MS). Mol Cell Proteomics 14, 120–135. https://doi.org/10.1074/mcp.M114.041012

Kim, D.I., Birendra, K.C., Zhu, W., Motamedchaboki, K., Doye, V., Roux, K.J., 2014. Probing nuclear pore complex architecture with proximity-dependent biotinylation. Proc Natl Acad Sci U S A 111, E2453-2461. https://doi.org/10.1073/pnas.1406459111

Klar, J., Piontek, J., Milatz, S., Tariq, M., Jameel, M., Breiderhoff, T., Schuster, J., Fatima, A., Asif, M., Sher, M., Mäbert, K., Fromm, A., Baig, S.M., Günzel, D., Dahl, N., 2017. Altered paracellular cation permeability due to a rare CLDN10B variant causes anhidrosis and kidney damage. PLoS Genet 13, e1006897. https://doi.org/10.1371/journal.pgen.1006897

Klepper, J., Wang, D., Fischbarg, J., Vera, J.C., Jarjour, I.T., O’Driscoll, K.R., De Vivo, D.C., 1999. Defective glucose transport across brain tissue barriers: a newly recognized neurological syndrome. Neurochem Res 24, 587–594. https://doi.org/10.1023/a:1022544131826

Konrad, M., Schaller, A., Seelow, D., Pandey, A.V., Waldegger, S., Lesslauer, A., Vitzthum, H., Suzuki, Y., Luk, J.M., Becker, C., Schlingmann, K.P., Schmid, M., Rodriguez-Soriano, J., Ariceta, G., Cano, F., Enriquez, R., Jüppner, H., Bakkaloglu, S.A., Hediger, M.A., Gallati, S., Neuhauss, S.C.F., Nürnberg, P., Weber, S., 2006. Mutations in the Tight-Junction Gene Claudin 19 (CLDN19) Are Associated with Renal Magnesium Wasting, Renal Failure, and Severe Ocular Involvement. The American Journal of Human Genetics 79, 949–957. https://doi.org/10.1086/508617

Kosugi, S., Hasebe, M., Tomita, M., Yanagawa, H., 2009. Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. PNAS 106, 10171–10176. https://doi.org/10.1073/pnas.0900604106

Kozan, P.A., McGeough, M.D., Peña, C.A., Mueller, J.L., Barrett, K.E., Marchelletta, R.R., Sivagnanam, M., 2015. Mutation of EpCAM leads to intestinal barrier and ion transport dysfunction. J Mol Med (Berl) 93, 535–545. https://doi.org/10.1007/s00109-014-1239-x

Kubota, K., Furuse, M., Sasaki, H., Sonoda, N., Fujita, K., Nagafuchi, A., Tsukita, S., 1999. Ca(2+)-independent cell-adhesion activity of claudins, a family of

integral membrane proteins localized at tight junctions. Curr Biol 9, 1035–1038. https://doi.org/10.1016/s0960-9822(99)80452-7

Kyuno, D., Takasawa, A., Kikuchi, S., Takemasa, I., Osanai, M., Kojima, T., 2021. Role of tight junctions in the epithelial-to-mesenchymal transition of cancer cells. Biochimica et Biophysica Acta (BBA) - Biomembranes 1863, 183503. https://doi.org/10.1016/j.bbamem.2020.183503

Ladwein, M., Uf, P., Ds, S., M, S., S, F., L, L., Ww, F., G, M., M, Z., 2005. The cell-cell adhesion molecule EpCAM interacts directly with the tight junction protein claudin-7. Exp Cell Res 309, 345–357. https://doi.org/10.1016/j.yexcr.2005.06.013

Lechuga, S., Baranwal, S., Ivanov, A.I., 2015. Actin-interacting protein 1 controls assembly and permeability of intestinal epithelial apical junctions. American Journal of Physiology-Gastrointestinal and Liver Physiology 308, G745–G756. https://doi.org/10.1152/ajpgi.00446.2014

Lee, J.-W., Hsiao, W.-T., Chen, H.-Y., Hsu, L.-P., Chen, P.-R., Lin, M.-D., Chiu, S.-J., Shih, W.-L., Hsu, Y.-C., 2010. Upregulated claudin-1 expression confers resistance to cell death of nasopharyngeal carcinoma cells. Int J Cancer 126, 1353–1366. https://doi.org/10.1002/ijc.24857

Lee, K., Ansar, M., Andrade, P.B., Khan, B., Santos-Cortez, R.L.P., Ahmad, W., Leal, S.M., 2012. Novel CLDN14 mutations in Pakistani families with autosomal recessive non-syndromic hearing loss. American Journal of Medical Genetics Part A 158A, 315–321. https://doi.org/10.1002/ajmg.a.34407

Lee, Y.-C., Bååth, J.A., Bastle, R.M., Bhattacharjee, S., Cantoria, M.J., Dornan, M., Gamero-Estevez, E., Ford, L., Halova, L., Kernan, J., Kürten, C., Li, S., Martinez, J., Sachan, N., Sarr, M., Shan, X., Subramanian, N., Rivera, K., Pappin, D., Lin, S.-H., 2018. Impact of Detergents on Membrane Protein Complex Isolation. J. Proteome Res. 17, 348–358. https://doi.org/10.1021/acs.jproteome.7b00599

Li, D., Roberts, R., 2001. Human Genome and Diseases: Review 58, 13.

Liu, F., Koval, M., Ranganathan, S., Fanayan, S., Hancock, W.S., Lundberg, E.K., Beavis, R.C., Lane, L., Duek, P., McQuade, L., Kelleher, N.L., Baker, M.S., 2016. Systems Proteomics View of the Endogenous Human Claudin Protein Family. J. Proteome Res. 15, 339–359. https://doi.org/10.1021/acs.jproteome.5b00769

Mandel, I., Paperna, T., Volkowich, A., Merhav, M., Glass-Marmor, L., Miller, A., 2012. The ubiquitin–proteasome pathway regulates claudin 5 degradation. Journal of Cellular Biochemistry 113, 2415–2423. https://doi.org/10.1002/jcb.24118

Marunaka, K., 2017. The RING finger- and PDZ domain-containing protein PDZRN3 controls localization of the Mg2+ regulator claudin-16 in renal tube epithelial cells 11.

Meyer, K., Kirchner, M., Uyar, B., Cheng, J.-Y., Russo, G., Hernandez-Miranda, L.R., Szymborska, A., Zauber, H., Rudolph, I.-M., Willnow, T.E., Akalin, A., Haucke, V., Gerhardt, H., Birchmeier, C., Kühn, R., Krauss, M., Diecke, S., Pascual, J.M., Selbach, M., 2018. Mutations in Disordered Regions Can Cause Disease by Creating Dileucine Motifs. Cell 175, 239-253.e17. https://doi.org/10.1016/j.cell.2018.08.019

Meyer, K., Selbach, M., 2015. Quantitative affinity purification mass spectrometry: a versatile technology to study protein–protein interactions. Front. Genet. 6. https://doi.org/10.3389/fgene.2015.00237

Milatz, S., Breiderhoff, T., 2017. One gene, two paracellular ion channels—claudin-10 in the kidney. Pflugers Arch - Eur J Physiol 469, 115–121. https://doi.org/10.1007/s00424-016-1921-7

Milatz, S., Himmerkus, N., Wulfmeyer, V.C., Drewell, H., Mutig, K., Hou, J., Breiderhoff, T., Müller, D., Fromm, M., Bleich, M., Günzel, D., 2017. Mosaic expression of claudins in thick ascending limbs of Henle results in spatial separation of paracellular Na^{+} and Mg^{2+} transport. Proc Natl Acad Sci USA 114, E219–E227. https://doi.org/10.1073/pnas.1611684114

Minezaki, Y., Homma, K., Nishikawa, K., 2007. Intrinsically Disordered Regions of Human Plasma Membrane Proteins Preferentially Occur in the Cytoplasmic Segment. Journal of Molecular Biology 368, 902–913. https://doi.org/10.1016/j.jmb.2007.02.033

Morita, K., Furuse, M., Fujimoto, K., Tsukita, S., 1999. Claudin multigene family encoding four-transmembrane domain protein components of tight junction strands. Proceedings of the National Academy of Sciences 96, 511–516. https://doi.org/10.1073/pnas.96.2.511

Mrowiec, T., Schwappach, B., 2006. 14-3-3 proteins in membrane protein transport. Biological Chemistry 387. https://doi.org/10.1515/BC.2006.152

Nagaoka, K., Udagawa, T., Richter, J.D., 2012. CPEB-mediated ZO-1 mRNA localization is required for epithelial tight-junction assembly and cell polarity. Nat Commun 3, 675. https://doi.org/10.1038/ncomms1678

Nakagawa, T., Nakayama, K., 2015. Protein monoubiquitylation: targets and diverse functions. Genes Cells 20, 543–562. https://doi.org/10.1111/gtc.12250

Nomme, J., Antanasijevic, A., Caffrey, M., Van Itallie, C.M., Anderson, J.M., Fanning, A.S., Lavie, A., 2015. Structural Basis of a Key Factor Regulating the Affinity between the Zonula Occludens First PDZ Domain and Claudins. Journal of Biological Chemistry 290, 16595–16606. https://doi.org/10.1074/jbc.M115.646695

Okada, H., Uezu, A., Soderblom, E.J., Moseley, M.A., Gertler, F.B., Soderling, S.H., 2012. Peptide array X-linking (PAX): a new peptide-protein identification approach. PLoS One 7, e37035. https://doi.org/10.1371/journal.pone.0037035

O'Kelly, I., Butler, M.H., Zilberberg, N., Goldstein, S.A.N., 2002. Forward Transport: 14-3-3 Binding Overcomes Retention in Endoplasmic Reticulum by Dibasic Signals 12.

Oshima, T., Miwa, H., Joh, T., 2008. Changes in the expression of claudins in active ulcerative colitis. Journal of Gastroenterology and Hepatology 23, S146–S150. https://doi.org/10.1111/j.1440-1746.2008.05405.x

Piontek, J., Winkler, L., Wolburg, H., Müller, S.L., Zuleger, N., Piehl, C., Wiesner, B., Krause, G., Blasig, I.E., 2008. Formation of tight junction: determinants of homophilic interaction between classic claudins. FASEB J 22, 146–158. https://doi.org/10.1096/fj.07-8319com

Prado Martins, R., Findakly, S., Daskalogianni, C., Teulade-Fichou, M.-P., Blondel, M., Fåhraeus, R., 2018. In Cellulo Protein-mRNA Interaction Assay to Determine the Action of G-Quadruplex-Binding Molecules. Molecules 23, 3124. https://doi.org/10.3390/molecules23123124

Rajan, S., Preisig-Müller, R., Wischmeyer, E., Nehring, R., Hanley, P.J., Renigunta, V., Musset, B., Schlichthörl, G., Derst, C., Karschin, A., Daut, J., 2002. Interaction with 14-3-3 proteins promotes functional expression of the potassium channels TASK-1 and TASK-3. The Journal of Physiology 545, 13–26. https://doi.org/10.1113/jphysiol.2002.027052

Ramberger, E., Sapozhnikova, V., Kowenz-Leutz, E., Zimmermann, K., Nicot, N., Nazarov, P.V., Perez-Hernandez, D., Reimer, U., Mertins, P., Dittmar, G., Leutz, A., 2020. A comprehensive motifs-based interactome of the C/EBPα transcription factor (preprint). Molecular Biology. https://doi.org/10.1101/2020.12.28.424569

Rhee, H.-W., Zou, P., Udeshi, N.D., Martell, J.D., Mootha, V.K., Carr, S.A., Ting, A.Y., 2013. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. Science 339, 1328–1331. https://doi.org/10.1126/science.1230593

Roux, K.J., Kim, D.I., Burke, B., May, D.G., 2018. BioID: A Screen for Protein-Protein Interactions. Curr Protoc Protein Sci 91, 19.23.1-19.23.15. https://doi.org/10.1002/cpps.51

Roux, K.J., Kim, D.I., Raida, M., Burke, B., 2012. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J Cell Biol 196, 801–810. https://doi.org/10.1083/jcb.201112098

Schreiber, E., Matthias, P., Müller, M.M., Schaffner, W., 1989. Rapid detection of octamer binding proteins with “mini-extracts”, prepared from a small number of cells. Nucleic Acids Res 17, 6419.

Schulze, W.X., Mann, M., 2004. A Novel Proteomic Screen for Peptide-Protein Interactions. Journal of Biological Chemistry 279, 10756–10764. https://doi.org/10.1074/jbc.M309909200

Selbach, M., Paul, F.E., Brandt, S., Guye, P., Daumke, O., Backert, S., Dehio, C., Mann, M., 2009. Host cell interactome of tyrosine-phosphorylated bacterial proteins. Cell Host Microbe 5, 397–403. https://doi.org/10.1016/j.chom.2009.03.004

Simon, D.B., Lu, Y., Choate, K.A., Velazquez, H., Al-Sabban, E., Praga, M., Casari, G., Bettinelli, A., Colussi, G., Rodriguez-Soriano, J., McCredie, D., Milford, D., Sanjad, S., Lifton, R.P., 1999. Paracellin-1, a Renal Tight Junction Protein Required for Paracellular Mg2+ Resorption. Science 285, 103–106. https://doi.org/10.1126/science.285.5424.103

Srinivasan, V., Braidy, N., Xu, Y.H., Xie, P., Kancherla, K., Chandramohan, S., Chan, E.K.W., Chan, D.K., 2017. Association of genetic polymorphisms of claudin-1 with small vessel vascular dementia. Clinical and Experimental Pharmacology and Physiology 44, 623–630. https://doi.org/10.1111/1440-1681.12747

Stamatovic, S.M., Johnson, A.M., Sladojevic, N., Keep, R.F., Andjelkovic, A.V., 2017. Endocytosis of tight junction proteins and the regulation of degradation and

recycling: Endocytic sorting of tight junction proteins. Ann. N.Y. Acad. Sci. 1397, 54–65. https://doi.org/10.1111/nyas.13346

Stein, A., Aloy, P., 2008. Contextual Specificity in Peptide-Mediated Protein Interactions. PLoS ONE 3, e2524. https://doi.org/10.1371/journal.pone.0002524

Sun, Z.-Y., Wei, J., Xie, L., Shen, Y., Liu, S.-Z., Ju, G.-Z., Shi, J.-P., Yu, Y.-Q., Zhang, X., Xu, Q., Hemmings, G.P., 2004. The CLDN5 locus may be involved in the vulnerability to schizophrenia. European Psychiatry 19, 354–357. https://doi.org/10.1016/j.eurpsy.2004.06.007

Takahashi, S., Iwamoto, N., Sasaki, H., Ohashi, M., Oda, Y., Tsukita, S., Furuse, M., 2009. The E3 ubiquitin ligase LNX1p80 promotes the removal of claudins from tight junctions in MDCK cells. Journal of Cell Science 122, 985–994. https://doi.org/10.1242/jcs.040055

Tan, B., Yatim, S.M.J.M., Peng, S., Gunaratne, J., Hunziker, W., Ludwig, A., 2020. The Mammalian Crumbs Complex Defines a Distinct Polarity Domain Apical of Epithelial Tight Junctions. Curr Biol 30, 2791-2804.e6. https://doi.org/10.1016/j.cub.2020.05.032

Tanaka, K., 2009. The proteasome: Overview of structure and functions. Proc. Jpn. Acad., Ser. B 85, 12–36. https://doi.org/10.2183/pjab.85.12

Tanaka, M., Kamata, R., Sakai, R., 2005. EphA2 Phosphorylates the Cytoplasmic Tail of Claudin-4 and Mediates Paracellular Permeability. Journal of Biological Chemistry 280, 42375–42382. https://doi.org/10.1074/jbc.M503786200

The UniProt Consortium, 2021. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Research 49, D480–D489. https://doi.org/10.1093/nar/gkaa1100

Thorens, B., Lodish, H.F., Brown, D., 1990. Differential localization of two glucose transporter isoforms in rat kidney. Am J Physiol 259, C286-294. https://doi.org/10.1152/ajpcell.1990.259.2.C286

Thorleifsson, G., Holm, H., Edvardsson, V., Walters, G.B., Styrkarsdottir, U., Gudbjartsson, D.F., Sulem, P., Halldorsson, B.V., de Vegt, F., d'Ancona, F.C.H., den Heijer, M., Franzson, L., Christiansen, C., Alexandersen, P., Rafnar, T., Kristjansson, K., Sigurdsson, G., Kiemeney, L.A., Bodvarsson, M., Indridason, O.S., Palsson, R., Kong, A., Thorsteinsdottir, U., Stefansson, K., 2009. Sequence variants in the CLDN14 gene associate with kidney stones and bone mineral density. Nat Genet 41, 926–930. https://doi.org/10.1038/ng.404

Thrower, J.S., 2000. Recognition of the polyubiquitin proteolytic signal. The EMBO Journal 19, 94–102. https://doi.org/10.1093/emboj/19.1.94

Todd, M.C., Petty, H.M., King, J.M., Piana Marshall, B.N., Sheller, R.A., Cuevas, M.E., 2015. Overexpression and delocalization of claudin-3 protein in MCF-7 and MDA-MB-415 breast cancer cell lines. Oncol Lett 10, 156–162. https://doi.org/10.3892/ol.2015.3160

Tompa, P., Davey, N.E., Gibson, T.J., Babu, M.M., 2014. A Million Peptide Motifs for the Molecular Biologist. Molecular Cell 55, 161–169. https://doi.org/10.1016/j.molcel.2014.05.032

Traweger, A., Fang, D., Liu, Y.-C., Stelzhammer, W., Krizbai, I.A., Fresser, F., Bauer, H.-C., Bauer, H., 2002. The Tight Junction-specific Protein Occludin Is a Functional Target of the E3 Ubiquitin-protein Ligase Itch *. Journal of Biological Chemistry 277, 10201–10208. https://doi.org/10.1074/jbc.M111384200

Tusnády, G.E., Dobson, L., Tompa, P., 2015. Disordered regions in transmembrane proteins. Biochimica et Biophysica Acta (BBA) - Biomembranes 1848, 2839–2848. https://doi.org/10.1016/j.bbamem.2015.08.002

Vallon, V., 2020. Glucose transporters in the kidney in health and disease. Pflugers Arch - Eur J Physiol 472, 1345–1370. https://doi.org/10.1007/s00424-020-02361-w

Van Itallie, C.M., Anderson, J.M., 2006. CLAUDINS AND EPITHELIAL PARACELLULAR TRANSPORT. Annu. Rev. Physiol. 68, 403–429. https://doi.org/10.1146/annurev.physiol.68.040104.131404

Van Itallie, C.M., Colegio, O.R., Anderson, J.M., 2004. The Cytoplasmic Tails of Claudins Can Influence Tight Junction Barrier Properties through Effects on Protein Stability. J Membrane Biol 199, 29–38. https://doi.org/10.1007/s00232-004-0673-z

Volkmer, R., Tapia, V., Landgraf, C., 2012. Synthetic peptide arrays for investigating protein interaction domains. FEBS Lett 586, 2780–2786. https://doi.org/10.1016/j.febslet.2012.04.028

Ward, J.J., Sodhi, J.S., McGuffin, L.J., Buxton, B.F., Jones, D.T., 2004. Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. Journal of Molecular Biology 337, 635–645. https://doi.org/10.1016/j.jmb.2004.02.002

Weber, S., Hoffmann, K., Jeck, N., Saar, K., Boeswald, M., Kuwertz-Broeking, E., Meij, I.I., Knoers, N.V., Cochat, P., Šuláková, T., Bonzel, K.E., Soergel, M., Manz, F., Schaerer, K., Seyberth, H.W., Reis, A., Konrad, M., 2000. Familial hypomagnesaemia with hypercalciuria and nephrocalcinosis maps to chromosome 3q27 and is associated with mutations in the PCLN-1 gene. Eur J Hum Genet 8, 414–422. https://doi.org/10.1038/sj.ejhg.5200475

Wilcox, E.R., Burton, Q.L., Naz, S., Riazuddin, Saima, Smith, T.N., Ploplis, B., Belyantseva, I., Ben-Yosef, T., Liburd, N.A., Morell, R.J., Kachar, B., Wu, D.K., Griffith, A.J., Riazuddin, Sheikh, Friedman, T.B., 2001. Mutations in the Gene Encoding Tight Junction Claudin-14 Cause Autosomal Recessive Deafness DFNB29. Cell 104, 165–172. https://doi.org/10.1016/S0092-8674(01)00200-8

Wright, P.E., Dyson, H.J., 1999. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. J Mol Biol 293, 321–331. https://doi.org/10.1006/jmbi.1999.3110

Zeissig, S., Burgel, N., Gunzel, D., Richter, J., Mankertz, J., Wahnschaffe, U., Kroesen, A.J., Zeitz, M., Fromm, M., Schulzke, J.-D., 2007. Changes in expression and distribution of claudin 2, 5 and 8 lead to discontinuous tight junctions and barrier dysfunction in active Crohn's disease. Gut 56, 61–72. https://doi.org/10.1136/gut.2006.094375

Zheng, A., Yuan, F., Li, Y., Zhu, F., Hou, P., Li, J., Song, X., Ding, M., Deng, H., 2007. Claudin-6 and Claudin-9 Function as Additional Coreceptors for Hepatitis C Virus. J. Virol. 81, 12465–12471. https://doi.org/10.1128/JVI.01457-07

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., Chanda, S.K., 2019. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun 10, 1523. https://doi.org/10.1038/s41467-019-09234-6

Zwanziger, D., Badziong, J., Ting, S., Moeller, L.C., Schmid, K.W., Siebolts, U., Wickenhauser, C., Dralle, H., Fuehrer, D., 2015. The impact of CLAUDIN-1 on follicular thyroid carcinoma aggressiveness. Endocr Relat Cancer 22, 819–830. https://doi.org/10.1530/ERC-14-0502

Appendix

List of Figures

List of Tables

Supplementary figures and tables:

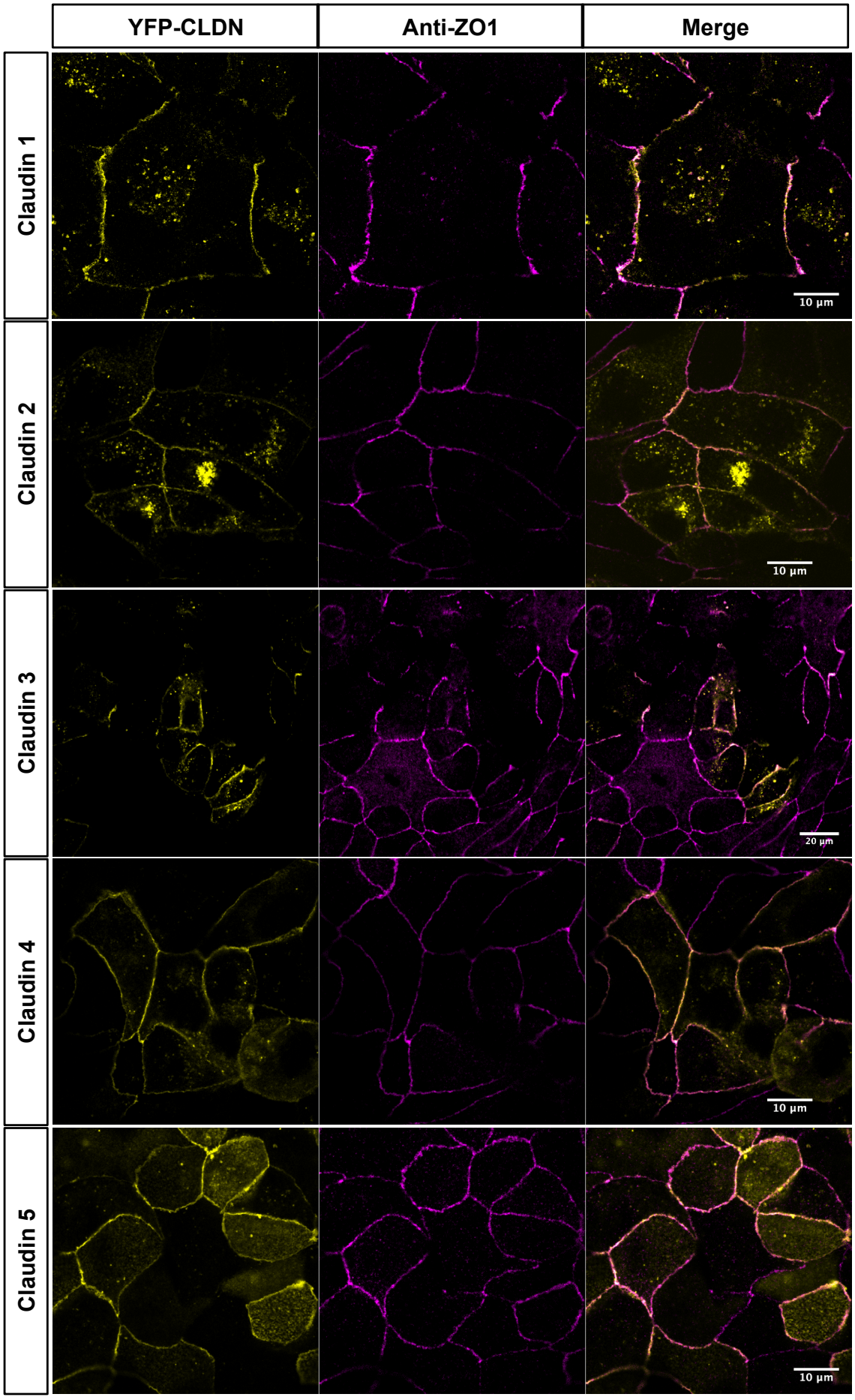
Claudin	Peptide sequence	Amino acid position	PTM position
Claudin-1 (O95832)	SCPRKTTSYPTPRPY	185 - 199	
	TTSYPTPRPYPKPAP	190 - 204	
	TPRPYPKPAPSSGKD	195 - 209	
	PKPAPSSGKDYV	200 - 211	
	PKPAPSSGKDy(p)V	200 - 211	Y210
Claudin-2 (P57739)	SCSSQRNRSNYYDAY	184 - 198	
	RNRSNYYDAYQAQPL	189 - 203	
	YYDAYQAQPLATRSSP	194 - 208	
	QAQPLATRSSPRPGQP	199 - 213	
	ATRSSPRPGQPPKVKS	204 - 218	
	PRPGQPPKVKSEFNSY	209 - 224	
	PRPGQPPKVKSEFNSy(p)	209 - 224	Y224
	PPKVKSEFNSYSLTGYV	214 - 230	
	PPKVKSEFNSy(p)SLTGYV	214 - 230	Y224
Claudin-3 (O15551)	CCSCPPREKKYTATK	181 - 195	
	PREKKYTATKVVYSA	186 - 200	
	YTATKVVYSAPRSTG	191 - 205	
	VVYSAPRSTGPGASL	196 - 210	
	PRSTGPGASLGTGYD	201 - 215	
	PRSTGPGASLGTGy(p)D	201 - 215	Y214
	PGASLGTGYDRKDYV	206 - 220	
	PGASLGTGy(p)DRKDYV	206 - 220	Y214
	PGASLGTGYDRKDy(p)V	206 - 220	Y219
Claudin-4 (O14493)	CCNCPPRTDKPYSAK	182 - 196	
	PRTDKPYSAKYSAAR	187 - 201	
	PYSAKYSAARSAAASNYV	192 - 209	
	PYSAKYSAARSAAASNy(p)V	192 - 209	Y208
Claudin-5 (O00501)	LCCGAWVCTGRPDLS	181 - 195	
	WVCTGRPDLSFPVKY	186 - 200	
	RPDLSFPVKYSAPRR	191 - 205	
	FPVKYSAPRRPTATG	196 - 210	
	SAPRRPTATGDYDKKNYV	201 - 218	
	SAPRRPTATGDy(p)DKKNYV	201 - 217	Y212
	SAPRRPTATGDYDKKNy(p)V	201 - 217	Y217
Claudin-6 (P56747)	CCTCPSGGSQGPSHY	182 -196	
	SGGSQGPSHYMARYS	187 - 201	
	GPSHYMARYSTSAPA	192 - 206	
	MARYSTSAPAISRGP	197 - 211	
	TSAPAISRGPSEYPTKNYV	202 - 220	
	TSAPAISRGPSEy(p)PTKNYV	202 - 220	Y214
	TSAPAISRGPSEYPTKNy(p)V	202 - 220	Y219
Claudin-7 (O95471)	LSCSCPGNESKAGYR	182 - 196	
	PGNESKAGYRVPRSY	187 - 201	
	KAGYRVPRSYPKSNS	192 - 206	

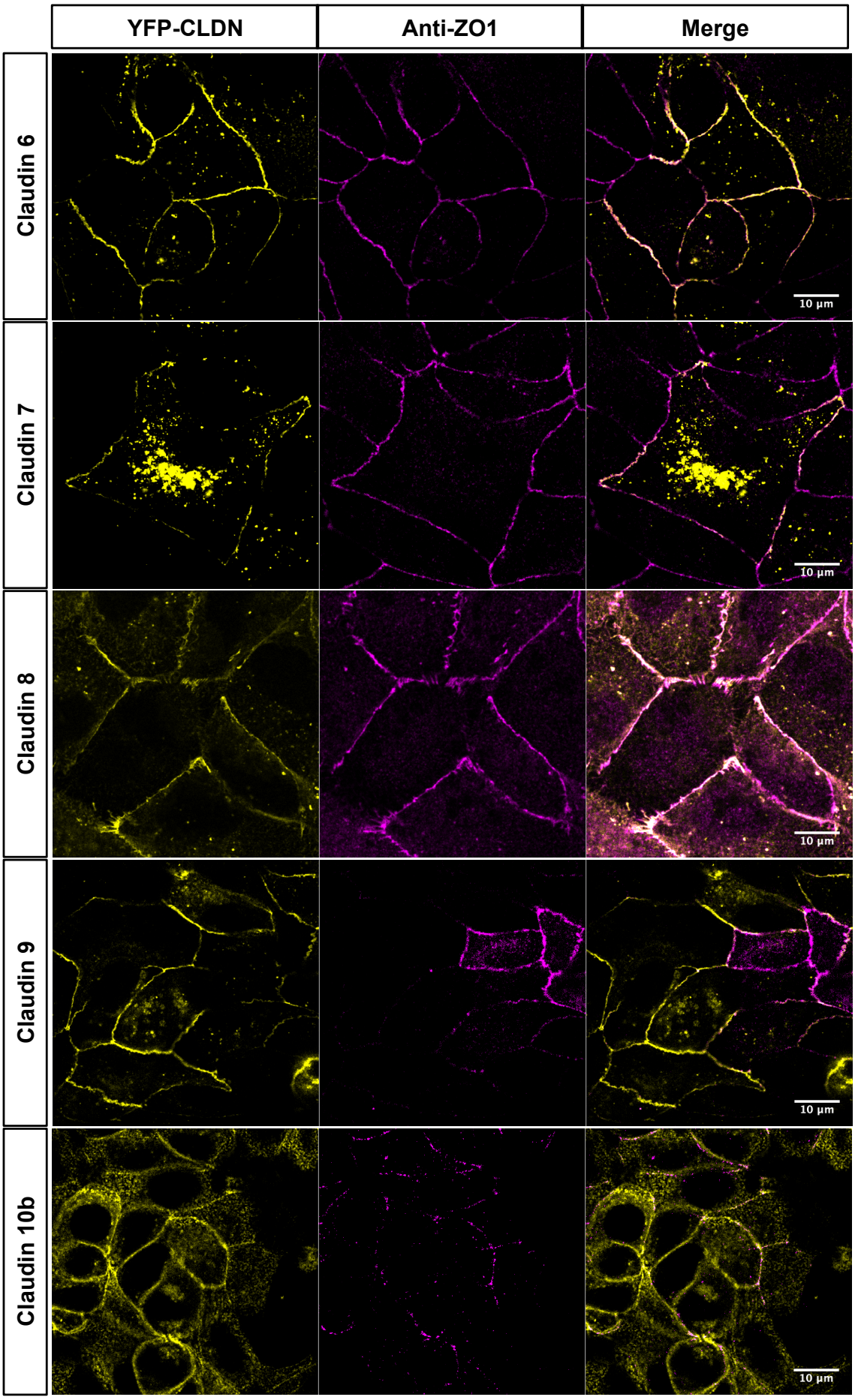
	VPRSYPKSNSSKEYV	197 - 211	
	VPRSYPKSNSSKEy(p)V	197 - 211	Y210
Claudin-8 (P56748)	CNEKSSSYRYSIPSH	188 - 202	
	SSYRYSIPSHRTTQK	193 - 207	
	SIPSHRTTQKSYHTG	198 - 212	
	RTTQKSYHTGKKSPS	203 - 217	
	SYHTGKKSPSVYSRSQYV	208 - 225	
Claudin-9 (O95484)	LCCTCPPPQVERPRG	181 - 195	
	PPPQVERPRGPRLGY	186 - 200	
	ERPRGPRLGYSIPSR	191 - 205	
	PRLGYSIPSRSGASG	196 - 210	
	SIPSRSGASGLDKRDYV	201 - 217	
	SIPSRSGASGLDKRDy(p)V	201 - 217	Y216
Claudin-10 (P78369)	SISDNNKTPRYTYNG	182 - 196	
	NKTPRYTYNGATSVM	187 - 201	
	YTYNGATSVMSSRTK	192 - 206	
	ATSVMSSRTKYHGGE	197 - 211	
	SSRTKYHGGEDFKTT	202 - 216	
	YHGGEDFKTTNPSKQ	207 - 221	
	DFKTTNPSKQFDKNAYV	212 - 228	
	DFKTTNPSKQFDKNAy(p)V	212 - 228	Y227
Claudin-11 (O75508)	AGDAQAFGENRFYYT	179 - 193	
	AFGENRFYYTAGSSS	184 - 198	
	RFYYTAGSSSPTHAKSAHV	189 - 207	
Claudin-12 (P56749)	YCTCKSLPSPFWQPL	196 - 210	
	SLPSPFWQPLYSHPP	201 - 215	
	FWQPLYSHPPSMHTY	206 - 220	
	YSHPPSMHTYSQPYS	211 - 225	
	SMHTYSQPYSARSRL	216 - 230	
	SQPYSARSRLSAIEI	221 - 235	
	ARSRLSAIEIDIPVVSHTT	226 - 244	
Claudin-14 (O95500)	SCQDEAPYRPYQAPP	184 - 198	
	APYRPYQAPPRATTT	189 - 203	
	YQAPPRATTTTANTA	194 - 208	
	RATTTTANTAPAYQP	199 - 213	
	TANTAPAYQPPAAYK	204 - 218	
	PAYQPPAAYKDNRAP	209 - 223	
	PAAYKDNRAPSVTSA	214 - 228	
	DNRAPSVTSATHSGY	219 - 233	
	DNRAPSVTSATHSGy(p)	219 - 233	Y233
	SVTSATHSGYRLNDYV	224 - 239	
	SVTSATHSGy(p)RLNDYV	224 - 239	Y233
Claudin-15 (P56746)	CCCGSDEDPAASARR	183 - 197	
	DEDPAASARRPYQAP	188 - 202	
	ASARRPYQAPVSVMP	193 - 207	

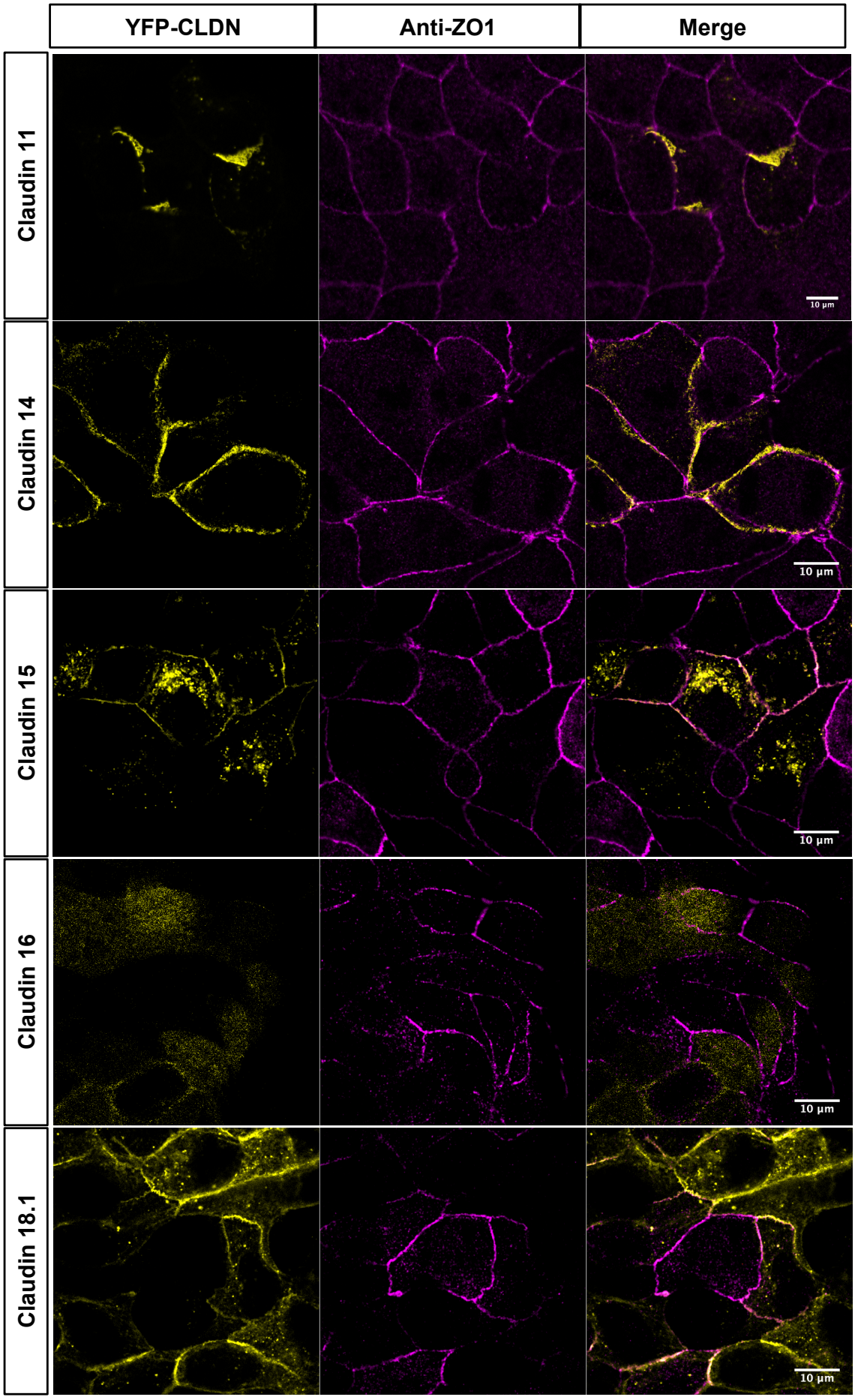
	PYQAPVSVMPVATSD	198 - 212	
	VSVMPVATSDQEGDS	203 - 217	
	VATSDQEGDSSFGKY	208 - 222	
	QEGDSSFGKYGRNAYV	213 - 228	
Claudin-16 (Q9Y5I7)	KDVGPERNYPYSLRK	261 - 275	
	ERNYPYSLRKAYSAA	266 - 280	
	YSLRKAYSAAGVSMA	271 - 285	
	AYSAAGVSMAKSYSA	276 - 290	
	GVSMAKSYSAPRTET	281 - 295	
	KSYSAPRTETAKMYA	286 - 300	
	KSYSAPRTETAKMy(p)A	286 - 300	Y299
	PRTETAKMYAVDTRV	291 - 305	
	PRTETAKMy(p)AVDTRV	291 - 305	Y299
Claudin-17 (P56750)	CCCNRKKQGYRYPVP	186 - 200	
	KKQGYRYPVPGYRVP	191 - 205	
	RYPVPGYRVPHTDKR	196 - 210	
	GYRVPHTDKRRNTTM	201 - 215	
	HTDKRRNTTMLSKTSTSYV	206 - 224	
Claudin-18 (P56856)	CRGLAPEETNYKAVS	196 - 210	
	PEETNYKAVSYHASG	201 - 215	
	YKAVSYHASGHSVAY	206 - 220	
	YHASGHSVAYKPGGF	211 - 225	
	HSVAYKPGGFKASTG	216 - 230	
	KPGGFKASTGFGSNT	221 - 235	
	KASTGFGSNTKNKKI	226 - 240	
	FGSNTKNKKIYDGGA	231 - 245	
	KNKKIYDGGARTEDE	236 - 250	
	YDGGARTEDEVQSYP	241 - 255	
	RTEDEVQSYPSKHDYV	246- 261	
	RTEDEVQSYPSKHDy(p)V	246- 261	Y260
Claudin-19A (Q8N6F1-1)	LCCTCPEPERPNSSP	182 - 196	
	PEPERPNSSPQPYRP	187 - 201	
	PNSSPQPYRPGPSAA	192 - 206	
	QPYRPGPSAAAREPV	197 - 211	
	GPSAAAREPVVKLPA	202 - 216	
	AREPVVKLPASAKGPLGV	207 - 224	
Claudin-19B (Q8N6F1-2)	QPYRPGPSAAAREYV	197 - 211	
Claudin-20 (P56880)	CTSCIKRNPEARLDP	182 - 196	
	KRNPEARLDPPTQQP	187 - 201	
	ARLDPPTQQPISNTQ	192 - 206	
	PTQQPISNTQLENNS	197 - 211	
	ISNTQLENNSTHNLKDYV	202 - 219	
Claudin-22 (Q8N7P3)	AACSSHAPLASGHYA	186 - 200	
	HAPLASGHYAVAQTQ	191 - 205	
	SGHYAVAQTQDHHQE	196 - 210	

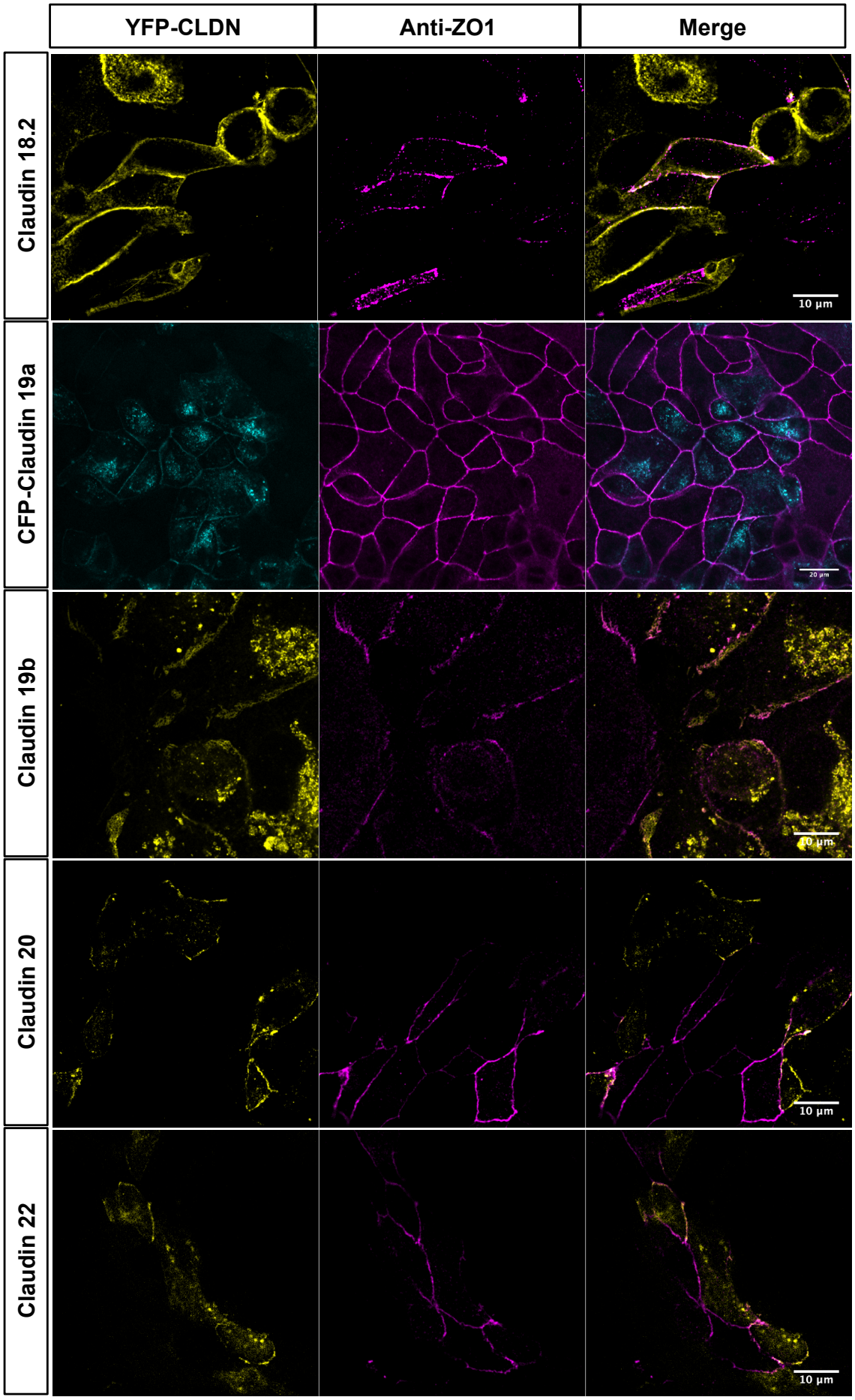
	VAQTQDHHQELETRNTNLK	201 - 220	
Claudin-23 (Q96B33)	APWCDERCRRRRKGP	182 - 196	
	ERCRRRRKGPSAGPR	187 - 201	
	RRKGPSAGPRRSSVS	192 - 206	
	SAGPRRSSVSTIQVE	197 - 211	
	RSSVSTIQVEWPEPD	202 - 216	
	TIQVEWPEPDLAPAI	207 - 221	
	WPEPDLAPAIKYYSD	212 - 226	
	LAPAIKYYSDGQHRP	217 - 231	
	KYYSDGQHRPPPAQH	222 - 236	
	GQHRPPPAQHRKPKP	227 - 241	
	PPAQHRKPKPKPKVG	232 - 246	
	RKPKPKPKVGFPMPR	237 - 251	
	KPKVGFPMPRPRPKA	242 - 256	
	FPMPRPRPKAYTNSV	247 - 261	
	PRPKAYTNSVDVLDG	252 - 266	
	YTNSVDVLDGEGWES	257 - 271	
	DVLDGEGWESQDAPS	262 - 276	
	EGWESQDAPSCSTHP	267 - 281	
	QDAPSCSTHPCDSSL	272 - 286	
	CSTHPCDSSLPCDSDL	277 - 292	
Claudin-24 (A6NM45)	LNCAACSSHAPLALG	183 - 197	
	CSSHAPLALGHYAVA	188 - 202	
	PLALGHYAVAQMQTQ	193 - 207	
	HYAVAQMQTQCPYLE	198 - 212	
	QMQTQCPYLEDGTADPQV	203 - 220	
Claudin-25 (C9JDP6)	SACLGKEDVPFPLMA	186 - 200	
	KEDVPFPLMAGPTVP	191 - 205	
	FPLMAGPTVPLSCAP	196 - 210	
	GPTVPLSCAPVEESD	201 - 215	
	LSCAPVEESDGSFHL	206 - 220	
	VEESDGSFHLMLRPRNLVI	211 - 229	

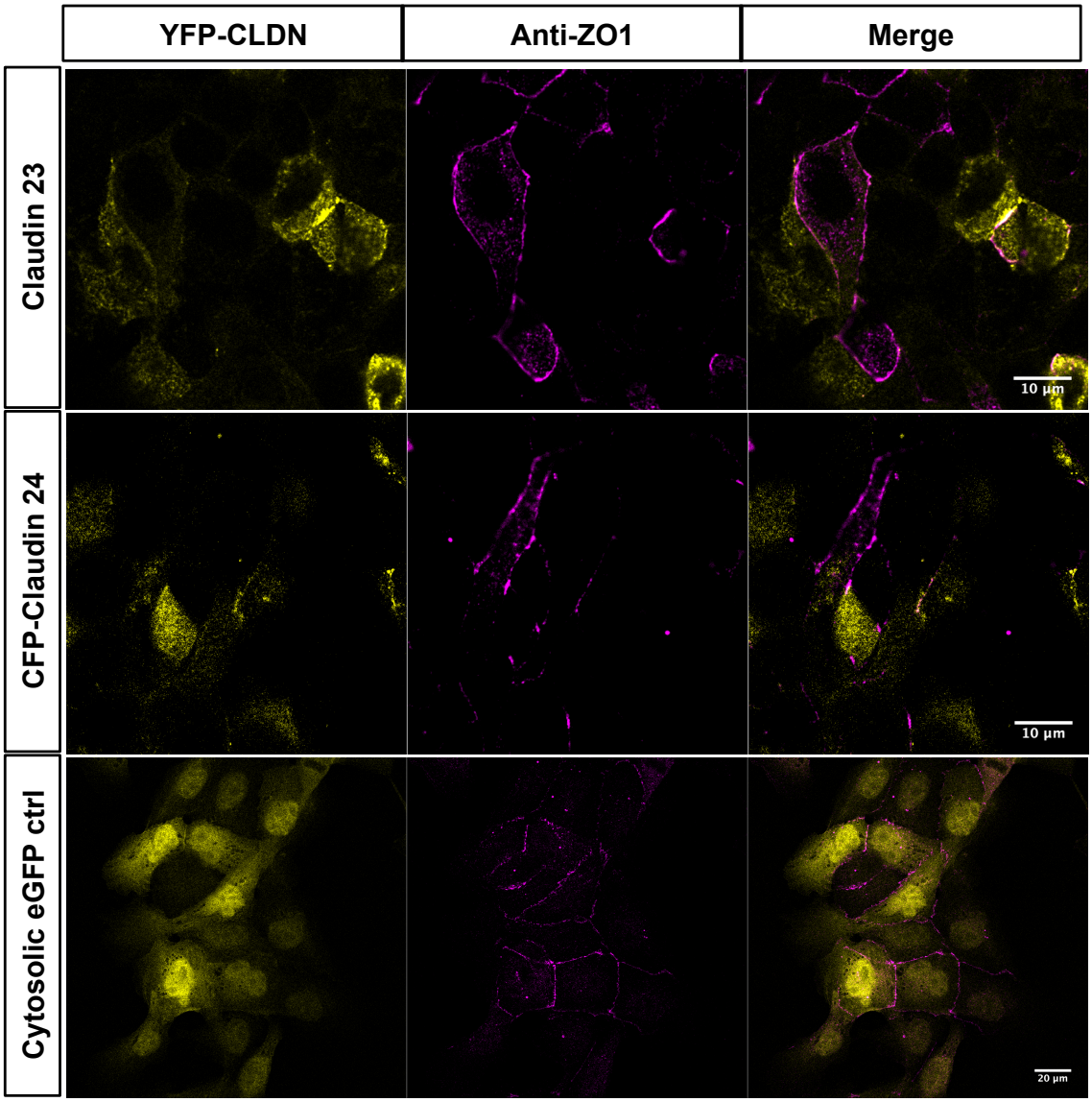
Supplementary Table 1. Claudin peptides used for the PRISMA screening. Peptide sequences derived from the human claudin protein family. Uniprot entries in brackets. Phosphorylated tyrosine residues were selected from the PhosphoSitePlus® database.


YFP-CLDN
Anti-ZO1
Merge
Claudin 1
Claudin 2
Claudin 3
Claudin 4
Claudin 5
10 µm
10 µm
20 µm
10 µm
10 µm

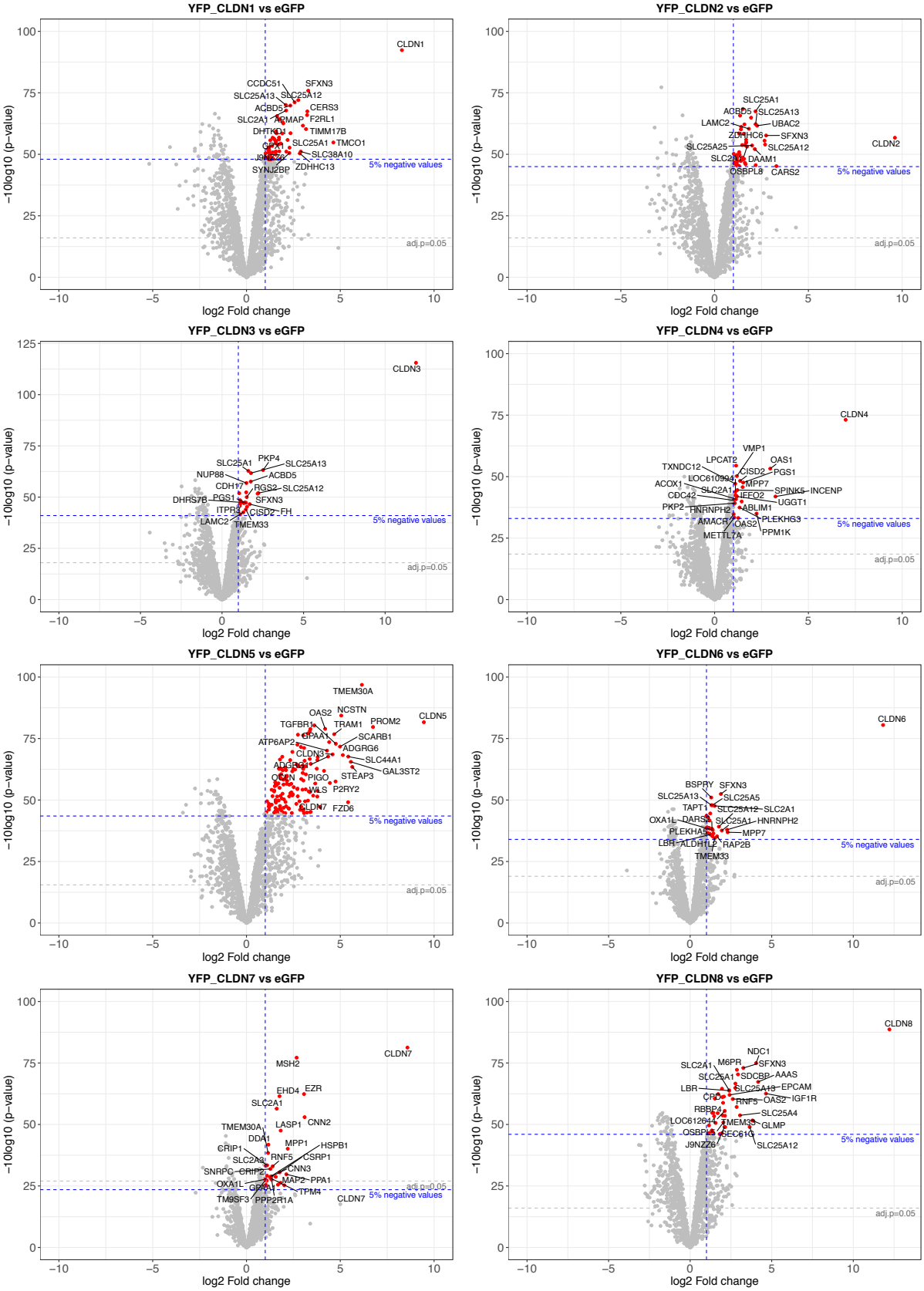

YFP-CLDN
Anti-ZO1
Merge
Claudin 6
Claudin 7
Claudin 8
Claudin 9
Claudin 10b
10 µm

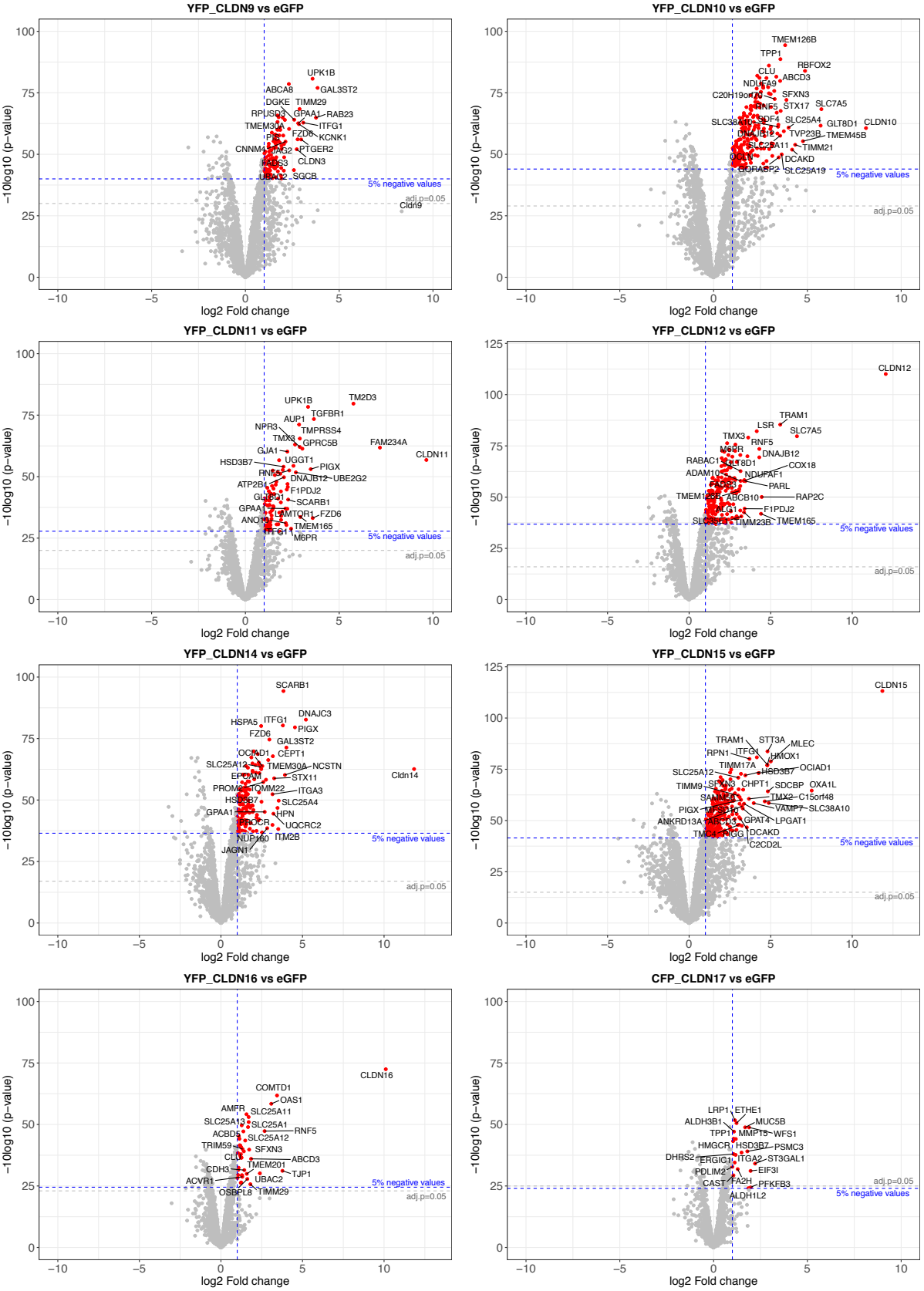

YFP-CLDN
Anti-ZO1
Merge
Claudin 11
Claudin 14
Claudin 15
Claudin 16
Claudin 18.1
10 µm

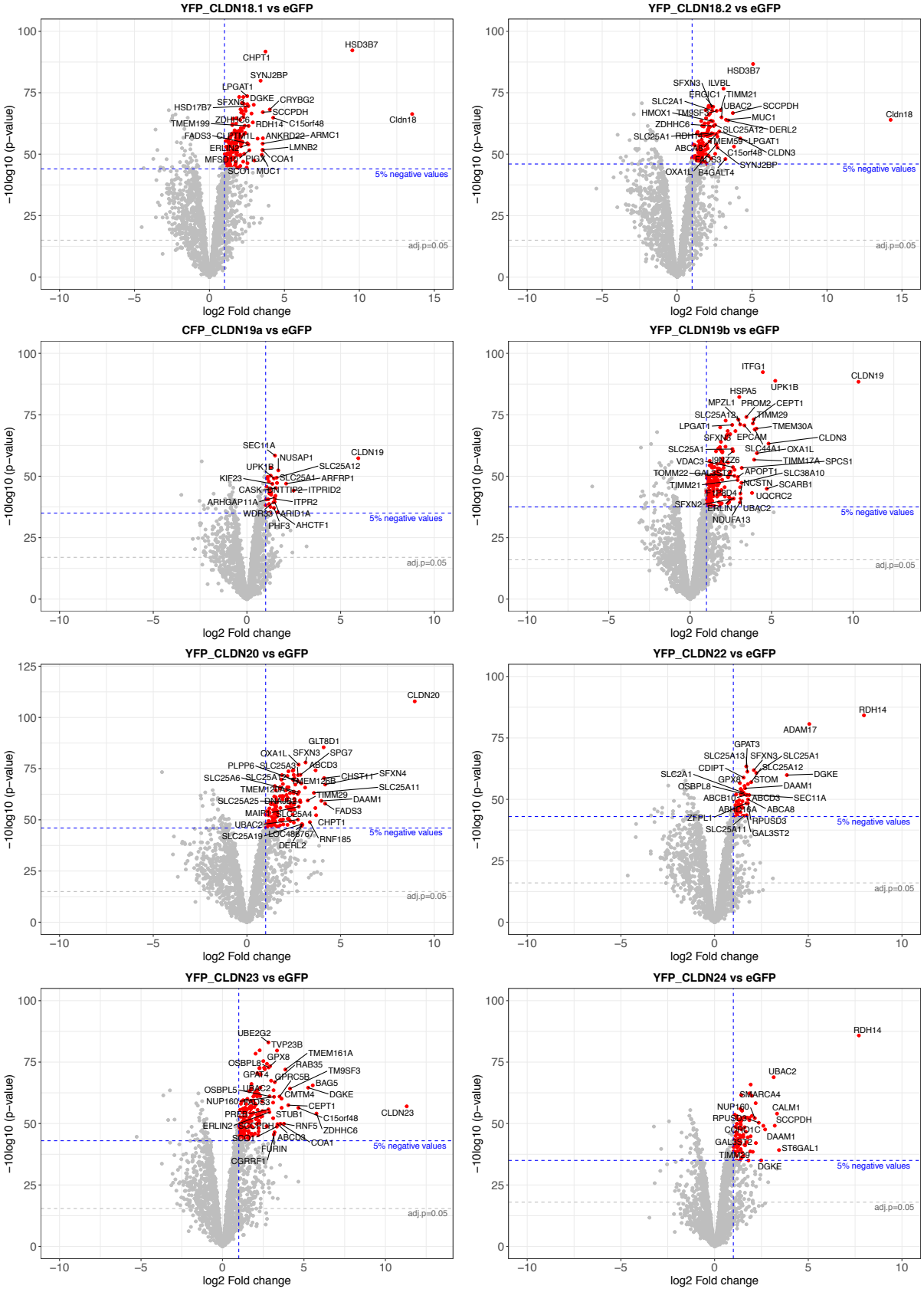

YFP-CLDN
Anti-ZO1
Merge
Claudin 18.2
CFP-Claudin 19a
Claudin 19b
Claudin 20
Claudin 22
10 µm
20 µm
10 µm
10 µm
10 µm



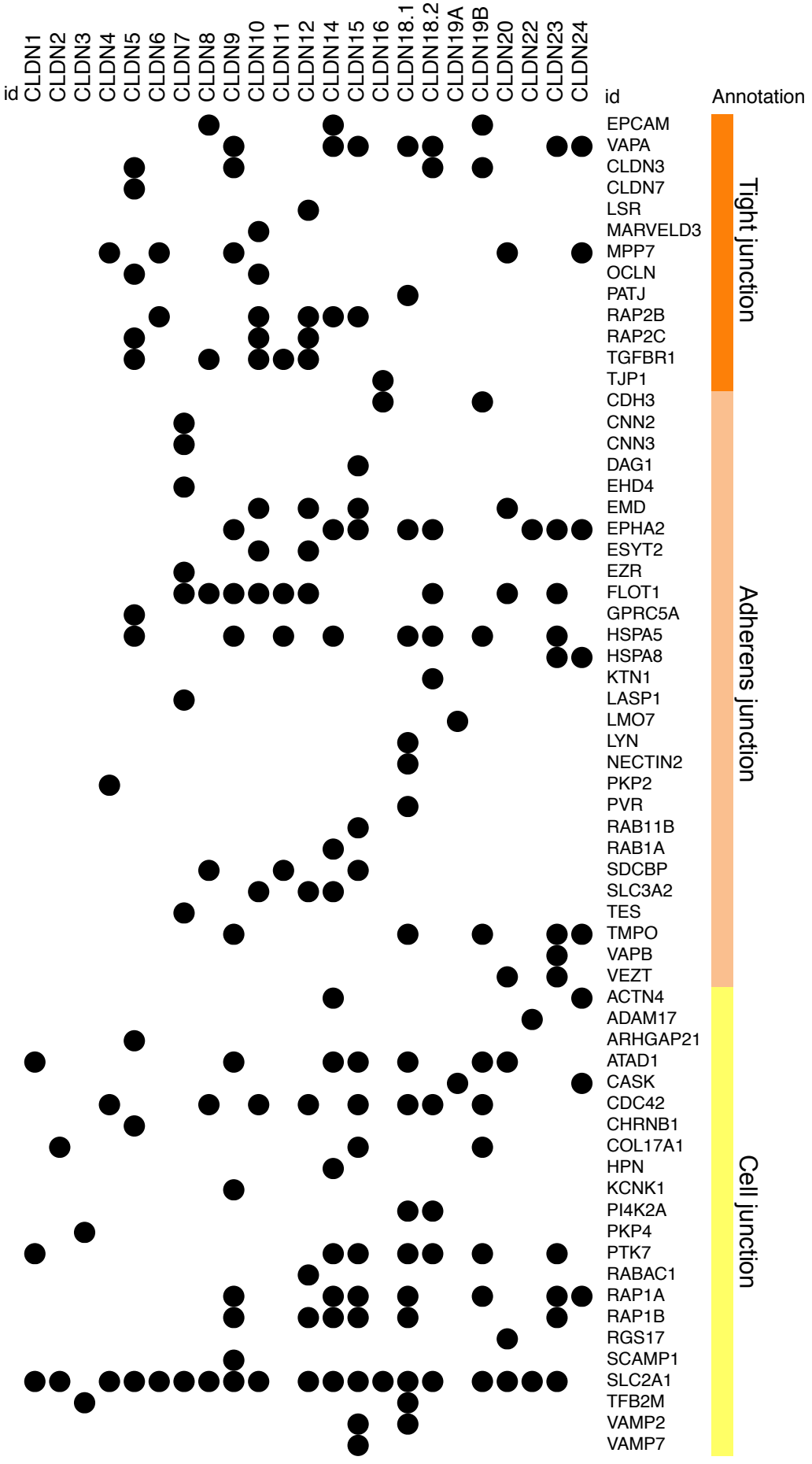
Suppl. Figure 1. Confocal microscopy images of MDCK-C7 stable cell lines overexpressing recombinant YFP-/CFP claudins. Anti-ZO-1 (61-7300 Invitrogen) and Alexa Fluor® 647 wese used for TJ staining.

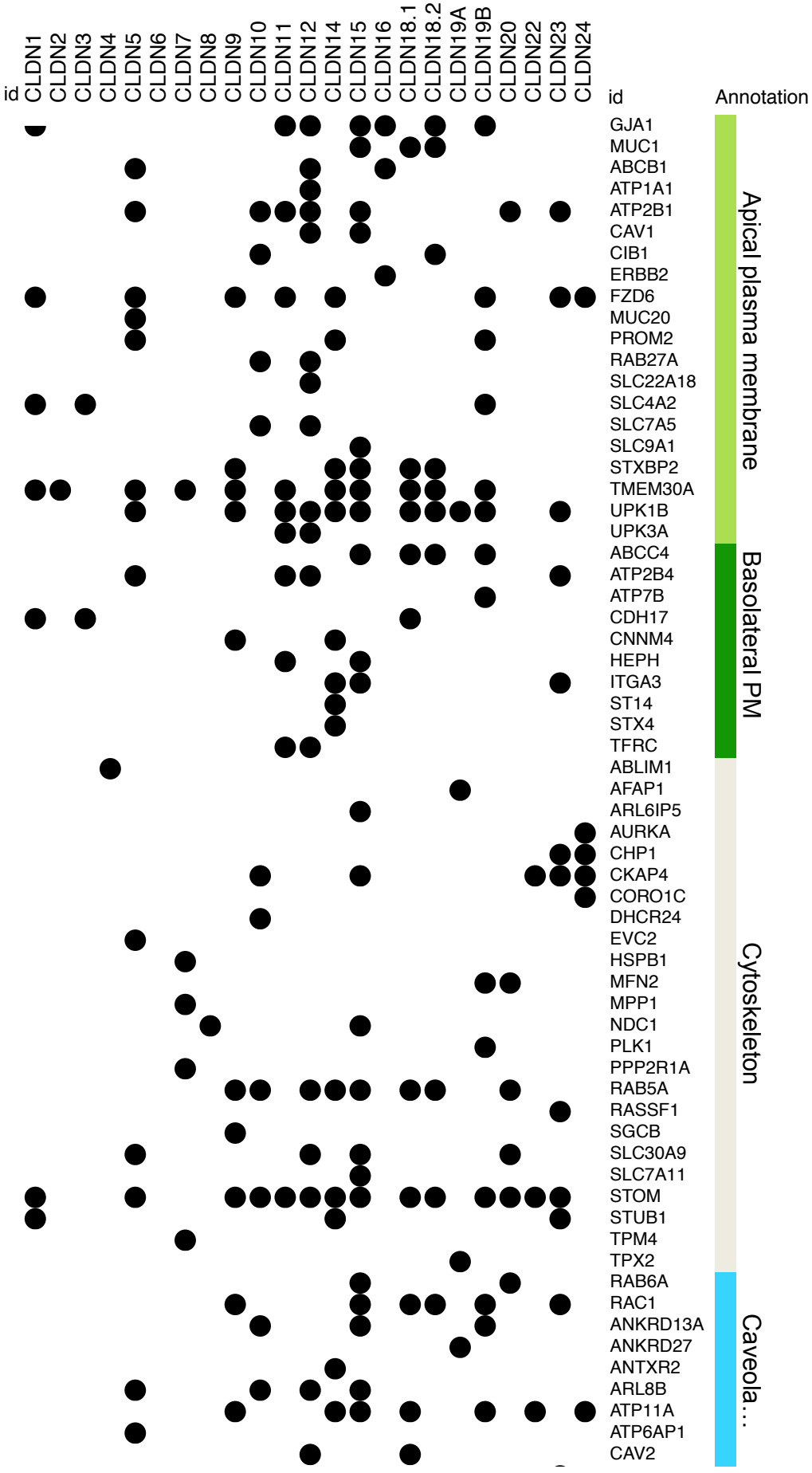

YFP_CLDN1 vs eGFP
YFP_CLDN2 vs eGFP
YFP_CLDN3 vs eGFP
YFP_CLDN4 vs eGFP
YFP_CLDN5 vs eGFP
YFP_CLDN6 vs eGFP
YFP_CLDN7 vs eGFP
YFP_CLDN8 vs eGFP
−10log10 (p−value)
log2 Fold change
5% negative values
adj.p=0.05

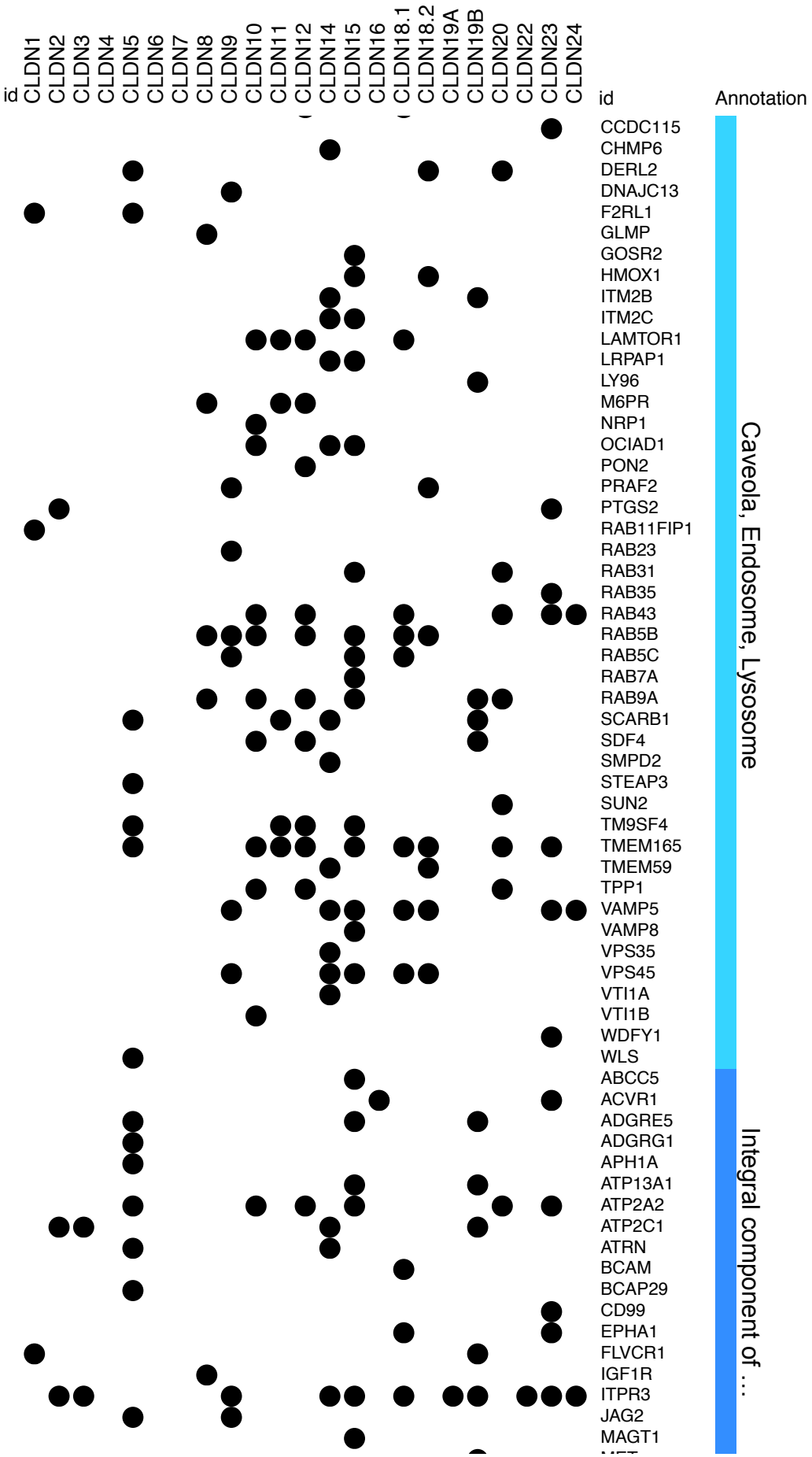

YFP_CLDN9 vs eGFP
YFP_CLDN10 vs eGFP
YFP_CLDN11 vs eGFP
YFP_CLDN12 vs eGFP
YFP_CLDN14 vs eGFP
YFP_CLDN15 vs eGFP
YFP_CLDN16 vs eGFP
CFP_CLDN17 vs eGFP
−10log10 (p−value)
log2 Fold change
5% negative values
adj.p=0.05
Cldn9
CLDN10
CLDN11
CLDN12
Cldn14
CLDN15
CLDN16

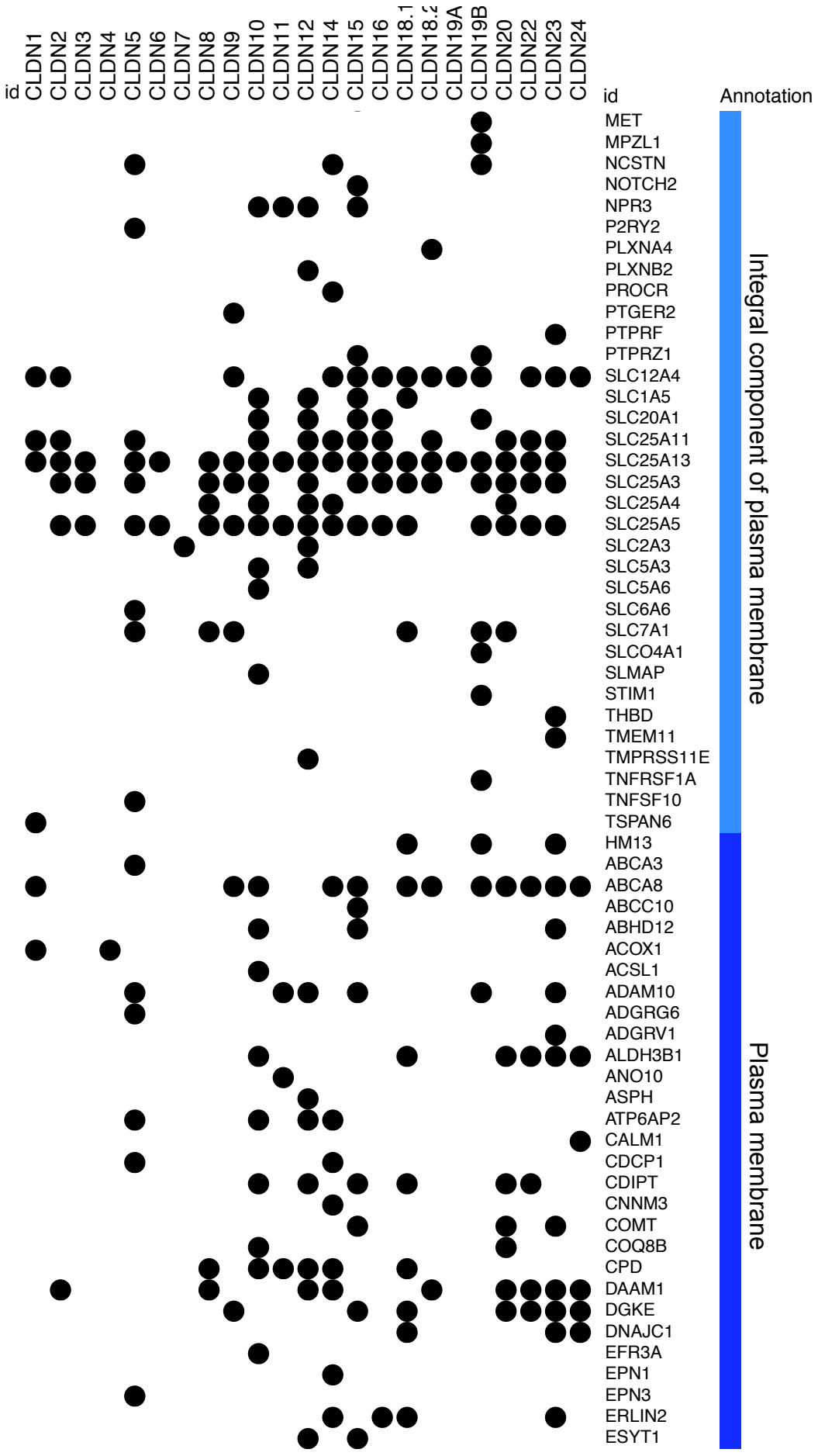

YFP_CLDN18.1 vs eGFP
YFP_CLDN18.2 vs eGFP
CFP_CLDN19a vs eGFP
YFP_CLDN19b vs eGFP
YFP_CLDN20 vs eGFP
YFP_CLDN22 vs eGFP
YFP_CLDN23 vs eGFP
YFP_CLDN24 vs eGFP
−10log10 (p−value)
log2 Fold change
5% negative values
adj.p=0.05
Cldn18
CLDN19
CLDN20
CLDN23
RDH14
HSD3B7
CHPT1
SYNJ2BP
LPGAT1
SFXN4
DGKE
CRYBG2
SCCPDH
HSD17B7
TMEM199
ZDHHC6
RDH14
C15orf48
FADS3
ANKRD22
ARMC1
LMNB2
COA1
MUC1
SFXN3
ILVBL
ERGIC1
TIMM21
SLC2A1
UBAC2
HMOX1
MUC1
SLC25A12
DERL2
SLC25A1
LPGAT1
ABCB10
C15orf48
CLDN3
SYNJ2BP
OXA1L
B4GALT4
SEC11A
NUSAP1
UPK1B
SLC25A12
KIF23
SLC25A1
ARFRP1
CASK
ITPRID2
ARHGAP11A
ITPR2
WDR33
ARID1A
PHF3
AHCTF1
ITFG1
UPK1B
HSPA5
MPZL1
PROM2
CEPT1
SLC25A13
TIMM29
LPGAT1
TMEM30A
SFXN3
EPCAM
CLDN3
SLC25A1
SLC44A1
OXA1L
VDAC3
TIMM17A
SPCS1
TOMM22
APOPT1
SLC38A10
TIMM21
NCSTN
SCARB1
UQCRC2
SFXN2
ERLIN1
UBAC2
NDUFA13
GLT8D1
OXA1L
SFXN3
SPG7
PLPP6
SLC25A3
ABCD3
SLC25A6
SLC25A1
CHST11
SFXN4
TMEM126B
SLC25A11
TIMM29
SLC25A25
DAAM1
FADS3
MAIP1
SLC25A4
CHPT1
UBAC2
SLC25A19
LOC488767
RNF185
DERL2
ADAM17
GPAT3
SLC25A13
SFXN3
SLC25A1
CDIPT
SLC25A12
SLC2A1
GPX8
STOM
DGKE
OSBPL8
DAAM1
ABCB10
ABCD3
SEC11A
ABHD16A
ABCA8
ZFPL1
RPUSD3
SLC25A11
GAL3ST2
UBE2G2
TVP23B
TMEM161A
OSBPL8
GPX8
RAB35
GPAT4
GPRC5B
TM9SF3
BAG5
OSBPL5
UBAC2
CMTM4
DGKE
NUP160
CEPT1
PRDX4
STUB1
C15orf48
ERLIN2
RNF5
ZDHHC6
SDC1
ABCD3
COA1
FURIN
CORO1C
UBAC2
SMARCA4
NUP160
CALM1
RPUSD3
SCCPDH
CORO1C
DAAM1
GAL3ST2
ST6GAL1
TIMM29
DGKE

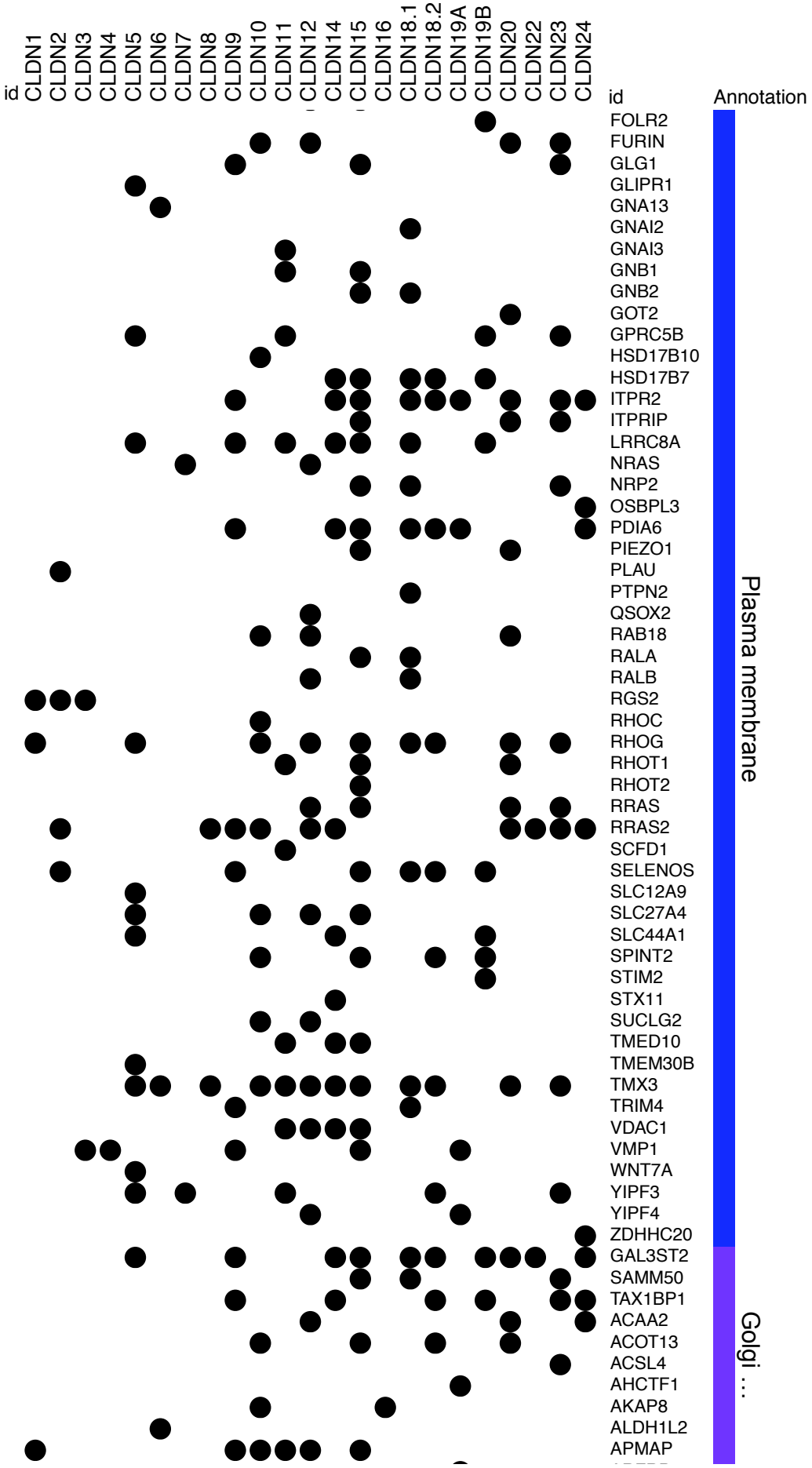
Suppl. Figure 2. Volcano plots showing the results from all CoIP experiments. LFQ intensity values of YFP/CFP-Claudin pull-downs were compared against the GFP control using moderated t-test. First significance level adj-pval<0.05 (grey line), 2nd level cutoff of adj-pval that leaves only 5% of the interactions identified in the GFP control (blue line).

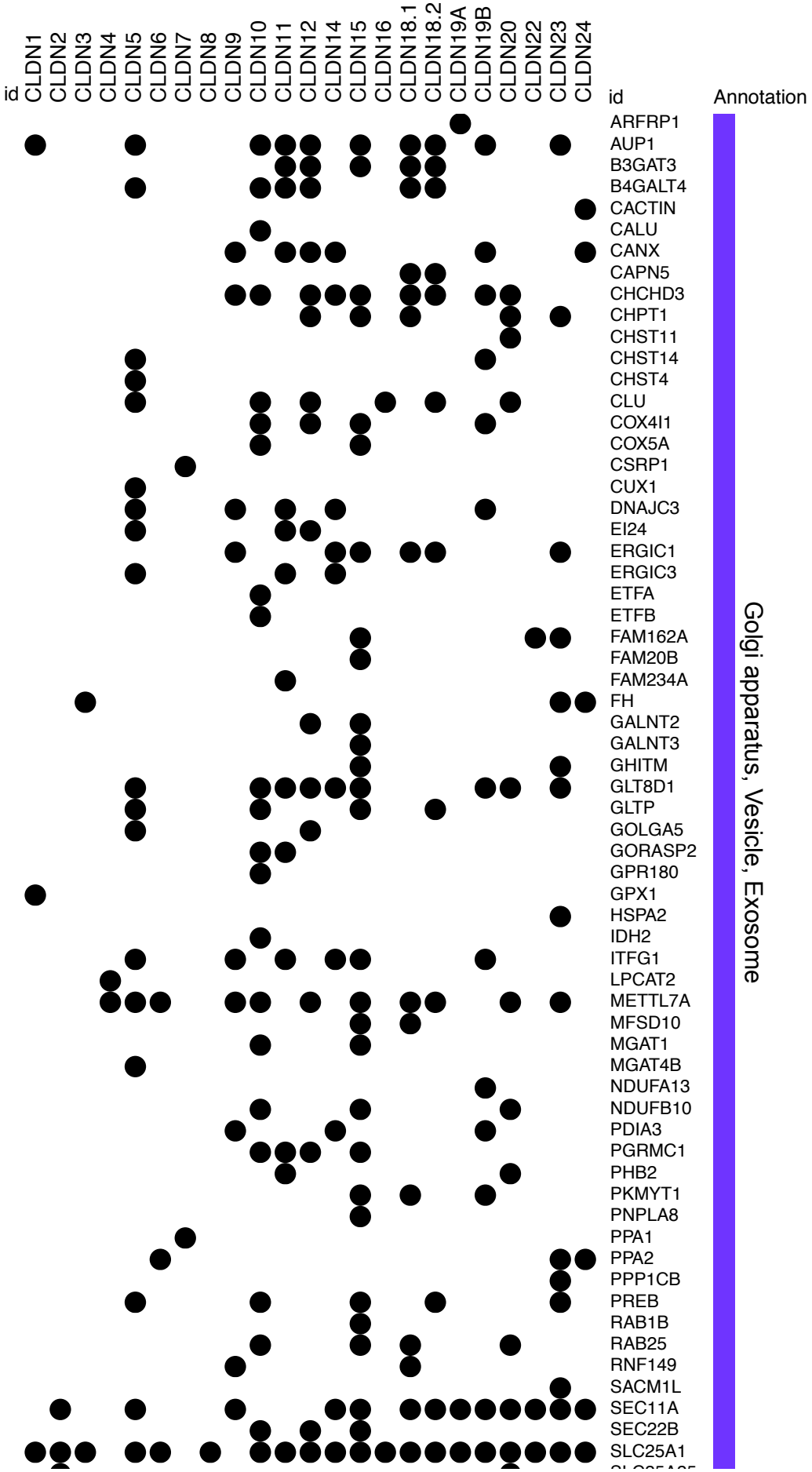

id
CLDN1
CLDN2
CLDN3
CLDN4
CLDN5
CLDN6
CLDN7
CLDN8
CLDN9
CLDN10
CLDN11
CLDN12
CLDN14
CLDN15
CLDN16
CLDN18.1
CLDN18.2
CLDN19A
CLDN19B
CLDN20
CLDN22
CLDN23
CLDN24
id
Annotation
EPCAM
VAPA
CLDN3
CLDN7
LSR
MARVELD3
MPP7
OCLN
PATJ
RAP2B
RAP2C
TGFBR1
TJP1
CDH3
CNN2
CNN3
DAG1
EHD4
EMD
EPHA2
ESYT2
EZR
FLOT1
GPRC5A
HSPA5
HSPA8
KTN1
LASP1
LMO7
LYN
NECTIN2
PKP2
PVR
RAB11B
RAB1A
SDCBP
SLC3A2
TES
TMPO
VAPB
VEZT
ACTN4
ADAM17
ARHGAP21
ATAD1
CASK
CDC42
CHRNB1
COL17A1
HPN
KCNK1
PI4K2A
PKP4
PTK7
RABAC1
RAP1A
RAP1B
RGS17
SCAMP1
SLC2A1
TFB2M
VAMP2
VAMP7
Tight junction
Adherens junction
Cell junction

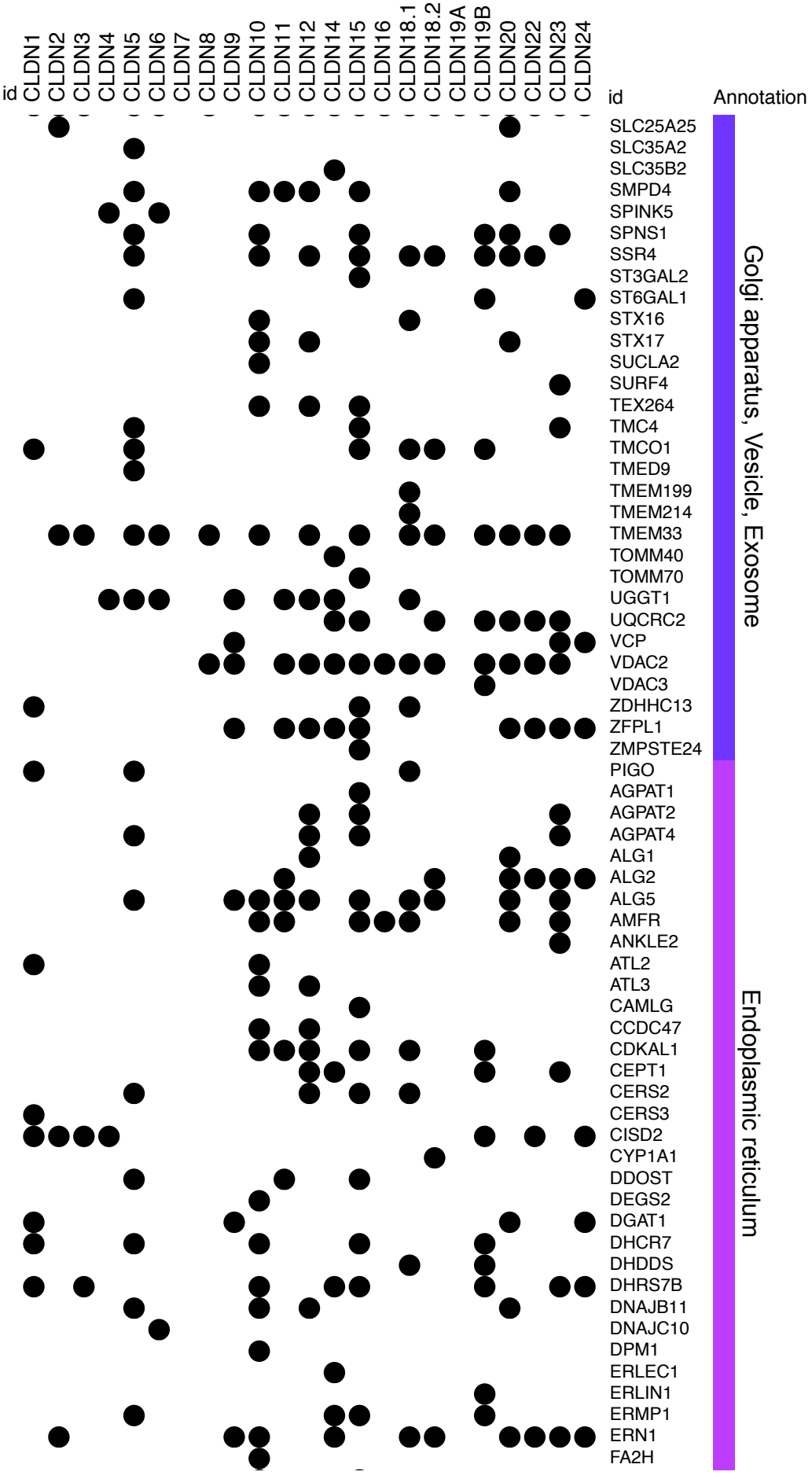

id
CLDN1
CLDN2
CLDN3
CLDN4
CLDN5
CLDN6
CLDN7
CLDN8
CLDN9
CLDN10
CLDN11
CLDN12
CLDN14
CLDN15
CLDN16
CLDN18.1
CLDN18.2
CLDN19A
CLDN19B
CLDN20
CLDN22
CLDN23
CLDN24
id
Annotation
GJA1
MUC1
ABCB1
ATP1A1
ATP2B1
CAV1
CIB1
ERBB2
FZD6
MUC20
PROM2
RAB27A
SLC22A18
SLC4A2
SLC7A5
SLC9A1
STXBP2
TMEM30A
UPK1B
UPK3A
ABCC4
ATP2B4
ATP7B
CDH17
CNNM4
HEPH
ITGA3
ST14
STX4
TFRC
ABLIM1
AFAP1
ARL6IP5
AURKA
CHP1
CKAP4
CORO1C
DHCR24
EVC2
HSPB1
MFN2
MPP1
NDC1
PLK1
PPP2R1A
RAB5A
RASSF1
SGCB
SLC30A9
SLC7A11
STOM
STUB1
TPM4
TPX2
RAB6A
RAC1
ANKRD13A
ANKRD27
ANTXR2
ARL8B
ATP11A
ATP6AP1
CAV2
Apical plasma membrane
Basolateral PM
Cytoskeleton
Caveola...

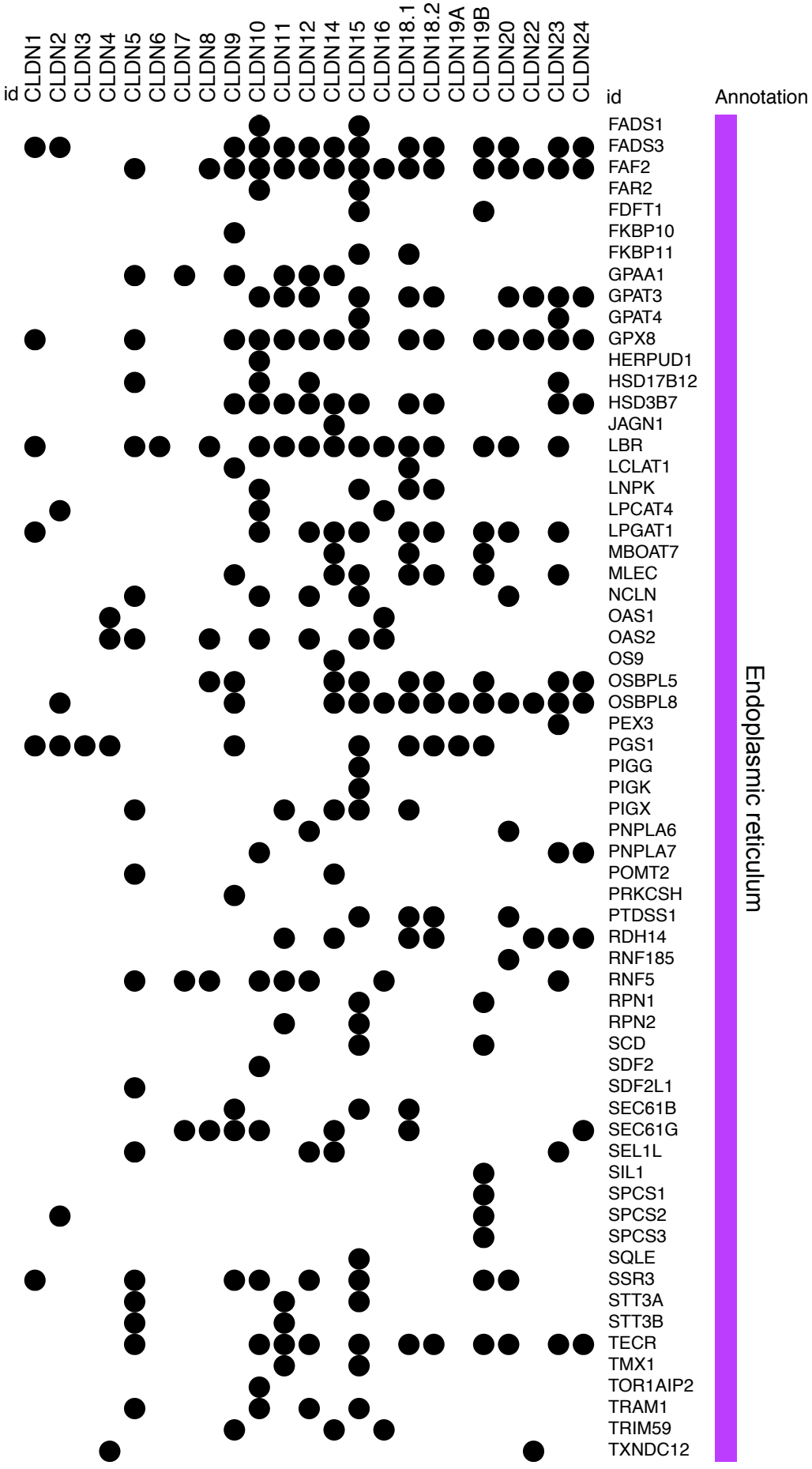

id
CLDN1
CLDN2
CLDN3
CLDN4
CLDN5
CLDN6
CLDN7
CLDN8
CLDN9
CLDN10
CLDN11
CLDN12
CLDN14
CLDN15
CLDN16
CLDN18.1
CLDN18.2
CLDN19A
CLDN19B
CLDN20
CLDN22
CLDN23
CLDN24
id
Annotation
CCDC115
CHMP6
DERL2
DNAJC13
F2RL1
GLMP
GOSR2
HMOX1
ITM2B
ITM2C
LAMTOR1
LRPAP1
LY96
M6PR
NRP1
OCIAD1
PON2
PRAF2
PTGS2
RAB11FIP1
RAB23
RAB31
RAB35
RAB43
RAB5B
RAB5C
RAB7A
RAB9A
SCARB1
SDF4
SMPD2
STEAP3
SUN2
TM9SF4
TMEM165
TMEM59
TPP1
VAMP5
VAMP8
VPS35
VPS45
VTI1A
VTI1B
WDFY1
WLS
ABCC5
ACVR1
ADGRE5
ADGRG1
APH1A
ATP13A1
ATP2A2
ATP2C1
ATRN
BCAM
BCAP29
CD99
EPHA1
FLVCR1
IGF1R
ITPR3
JAG2
MAGT1
Caveola, Endosome, Lysosome
Integral component of …

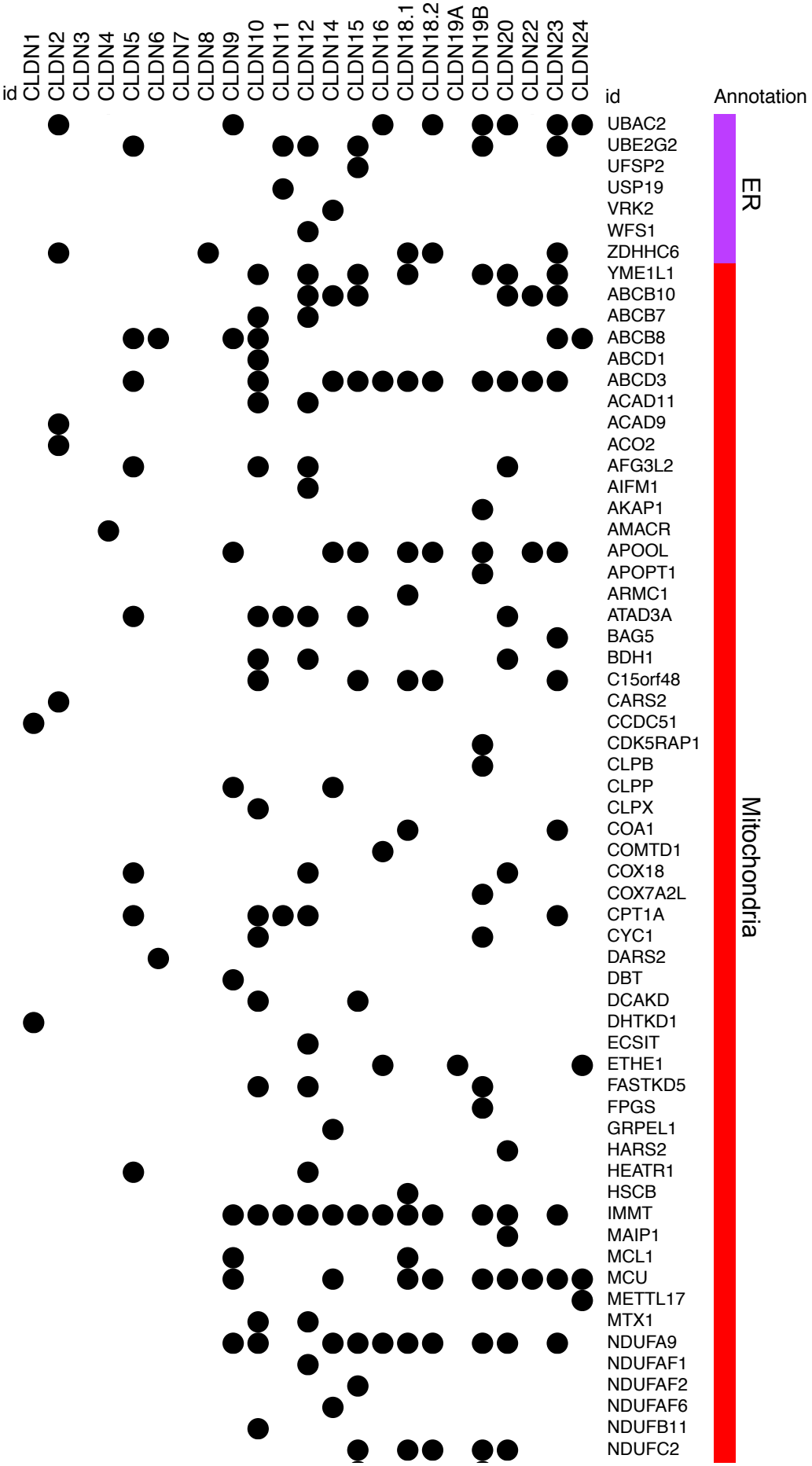

id
CLDN1
CLDN2
CLDN3
CLDN4
CLDN5
CLDN6
CLDN7
CLDN8
CLDN9
CLDN10
CLDN11
CLDN12
CLDN14
CLDN15
CLDN16
CLDN18.1
CLDN18.2
CLDN19A
CLDN19B
CLDN20
CLDN22
CLDN23
CLDN24
id
Annotation
MET
MPZL1
NCSTN
NOTCH2
NPR3
P2RY2
PLXNA4
PLXNB2
PROCR
PTGER2
PTPRF
PTPRZ1
SLC12A4
SLC1A5
SLC20A1
SLC25A11
SLC25A13
SLC25A3
SLC25A4
SLC25A5
SLC2A3
SLC5A3
SLC5A6
SLC6A6
SLC7A1
SLCO4A1
SLMAP
STIM1
THBD
TMEM11
TMPRSS11E
TNFRSF1A
TNFSF10
TSPAN6
HM13
ABCA3
ABCA8
ABCC10
ABHD12
ACOX1
ACSL1
ADAM10
ADGRG6
ADGRV1
ALDH3B1
ANO10
ASPH
ATP6AP2
CALM1
CDCP1
CDIPT
CNNM3
COMT
COQ8B
CPD
DAAM1
DGKE
DNAJC1
EFR3A
EPN1
EPN3
ERLIN2
ESYT1
Integral component of plasma membrane
Plasma membrane

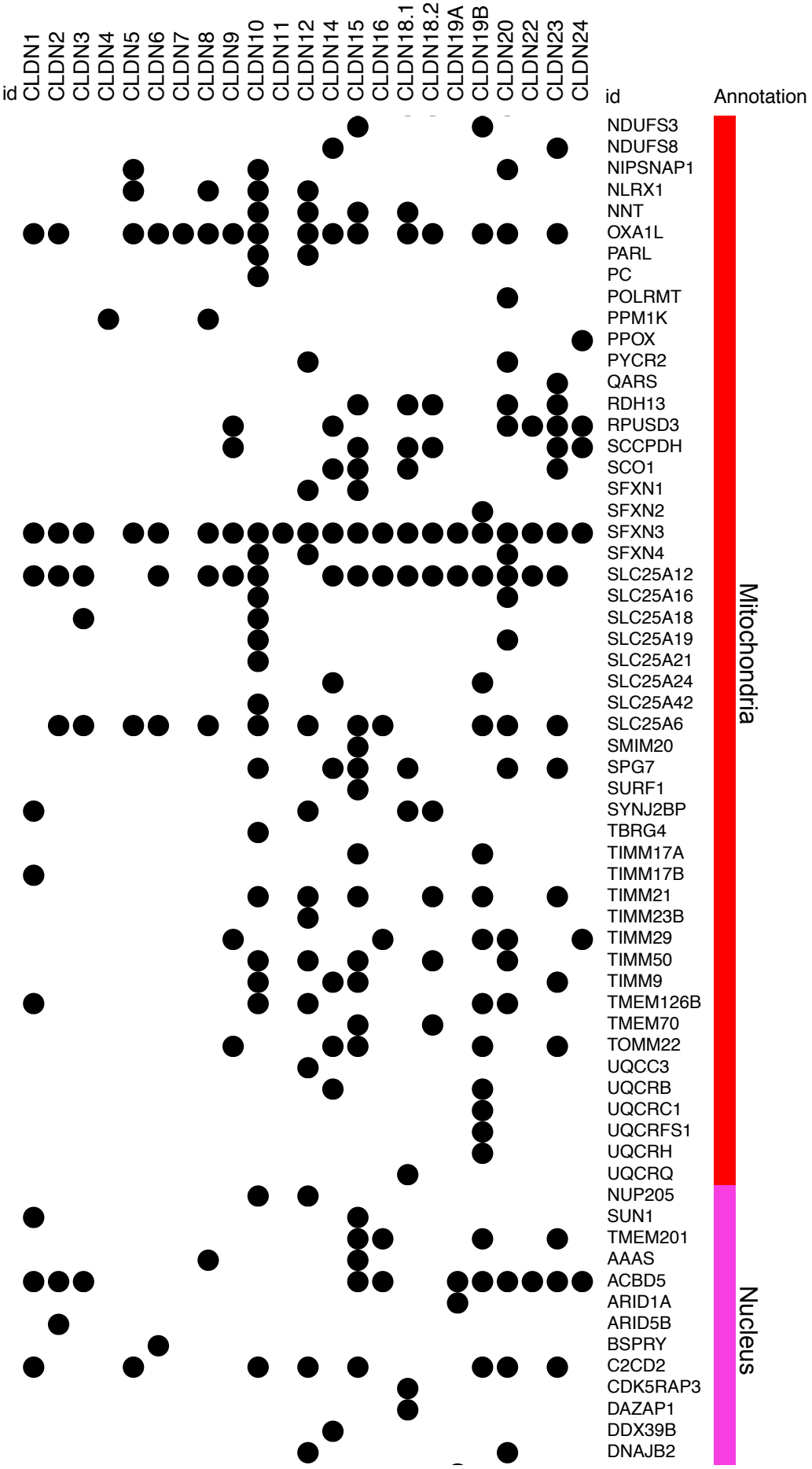

id
CLDN1
CLDN2
CLDN3
CLDN4
CLDN5
CLDN6
CLDN7
CLDN8
CLDN9
CLDN10
CLDN11
CLDN12
CLDN14
CLDN15
CLDN16
CLDN18.1
CLDN18.2
CLDN19A
CLDN19B
CLDN20
CLDN22
CLDN23
CLDN24
id
Annotation
FOLR2
FURIN
GLG1
GLIPR1
GNA13
GNAI2
GNAI3
GNB1
GNB2
GOT2
GPRC5B
HSD17B10
HSD17B7
ITPR2
ITPRIP
LRRC8A
NRAS
NRP2
OSBPL3
PDIA6
PIEZO1
PLAU
PTPN2
QSOX2
RAB18
RALA
RALB
RGS2
RHOC
RHOG
RHOT1
RHOT2
RRAS
RRAS2
SCFD1
SELENOS
SLC12A9
SLC27A4
SLC44A1
SPINT2
STIM2
STX11
SUCLG2
TMED10
TMEM30B
TMX3
TRIM4
VDAC1
VMP1
WNT7A
YIPF3
YIPF4
ZDHHC20
GAL3ST2
SAMM50
TAX1BP1
ACAA2
ACOT13
ACSL4
AHCTF1
AKAP8
ALDH1L2
APMAP
Plasma membrane
Golgi ...

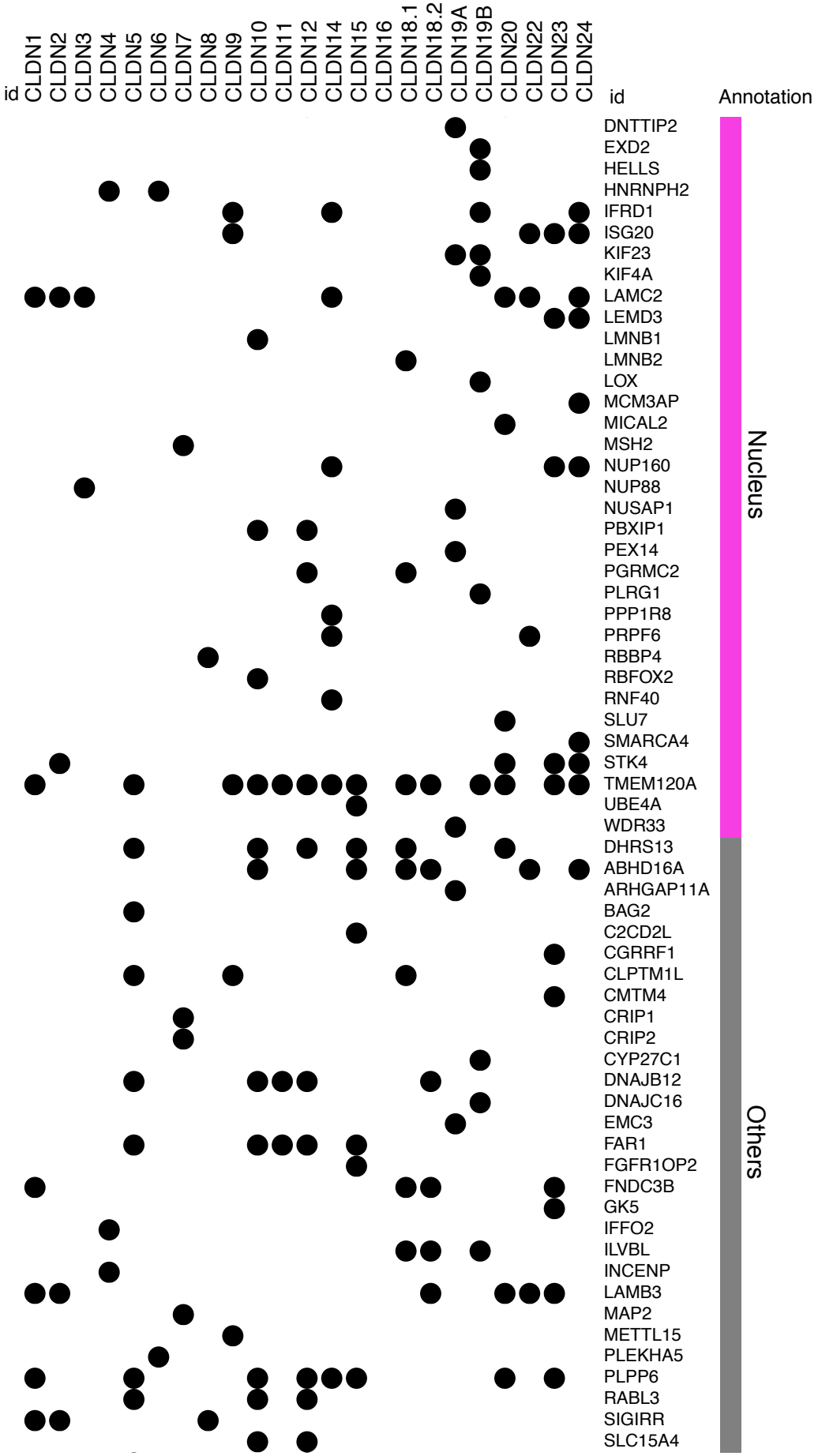

id
CLDN1
CLDN2
CLDN3
CLDN4
CLDN5
CLDN6
CLDN7
CLDN8
CLDN9
CLDN10
CLDN11
CLDN12
CLDN14
CLDN15
CLDN16
CLDN18.1
CLDN18.2
CLDN19A
CLDN19B
CLDN20
CLDN22
CLDN23
CLDN24
id
Annotation
ARFRP1
AUP1
B3GAT3
B4GALT4
CACTIN
CALU
CANX
CAPN5
CHCHD3
CHPT1
CHST11
CHST14
CHST4
CLU
COX4I1
COX5A
CSRP1
CUX1
DNAJC3
EI24
ERGIC1
ERGIC3
ETFA
ETFB
FAM162A
FAM20B
FAM234A
FH
GALNT2
GALNT3
GHITM
GLT8D1
GLTP
GOLGA5
GORASP2
GPR180
GPX1
HSPA2
IDH2
ITFG1
LPCAT2
METTL7A
MFSD10
MGAT1
MGAT4B
NDUFA13
NDUFB10
PDIA3
PGRMC1
PHB2
PKMYT1
PNPLA8
PPA1
PPA2
PPP1CB
PREB
RAB1B
RAB25
RNF149
SACM1L
SEC11A
SEC22B
SLC25A1
Golgi apparatus, Vesicle, Exosome

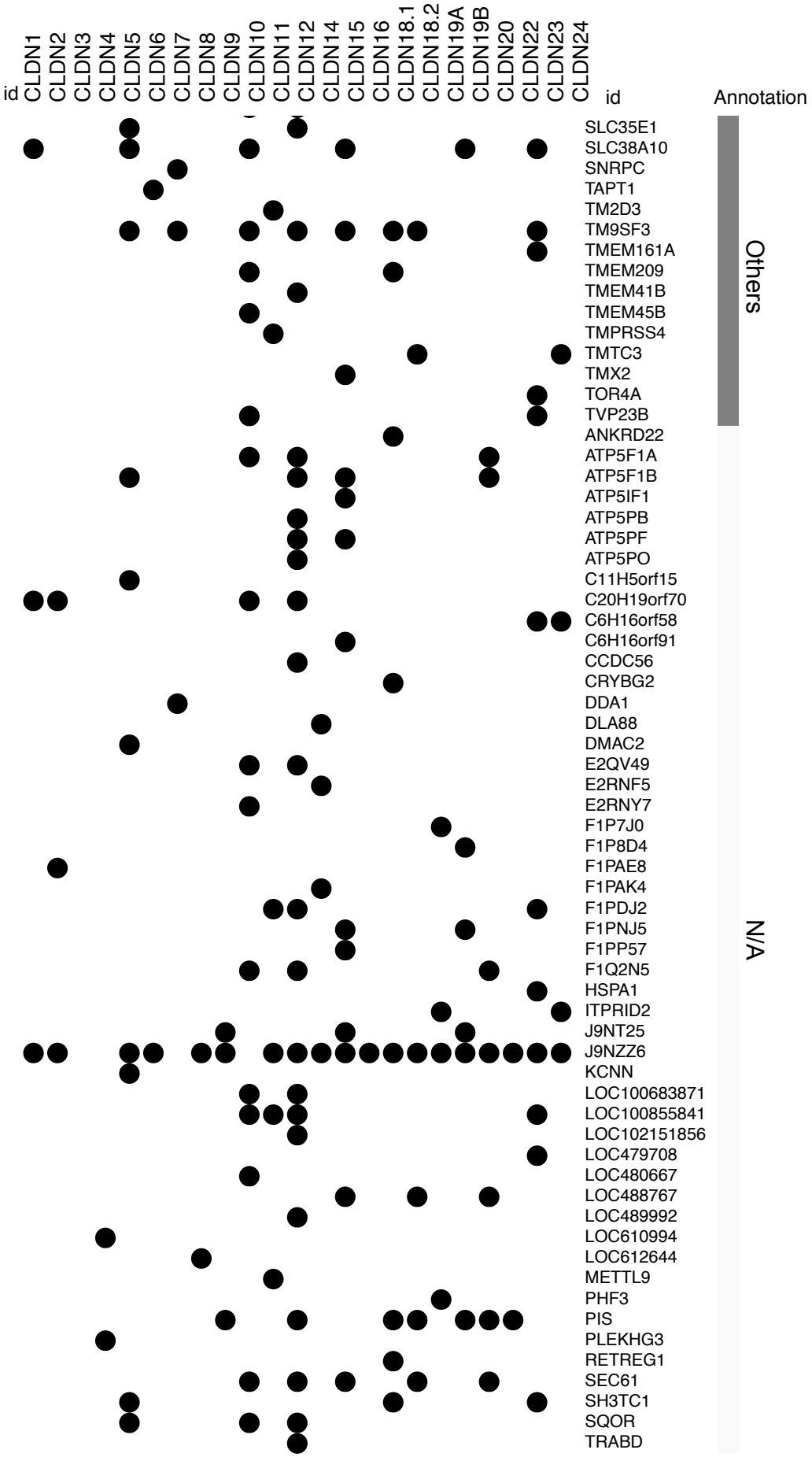

id
CLDN1
CLDN2
CLDN3
CLDN4
CLDN5
CLDN6
CLDN7
CLDN8
CLDN9
CLDN10
CLDN11
CLDN12
CLDN14
CLDN15
CLDN16
CLDN18.1
CLDN18.2
CLDN19A
CLDN19B
CLDN20
CLDN22
CLDN23
CLDN24
id
Annotation
SLC25A25
SLC35A2
SLC35B2
SMPD4
SPINK5
SPNS1
SSR4
ST3GAL2
ST6GAL1
STX16
STX17
SUCLA2
SURF4
TEX264
TMC4
TMCO1
TMED9
TMEM199
TMEM214
TMEM33
TOMM40
TOMM70
UGGT1
UQCRC2
VCP
VDAC2
VDAC3
ZDHHC13
ZFPL1
ZMPSTE24
PIGO
AGPAT1
AGPAT2
AGPAT4
ALG1
ALG2
ALG5
AMFR
ANKLE2
ATL2
ATL3
CAMLG
CCDC47
CDKAL1
CEPT1
CERS2
CERS3
CISD2
CYP1A1
DDOST
DEGS2
DGAT1
DHCR7
DHDDS
DHRS7B
DNAJB11
DNAJC10
DPM1
ERLEC1
ERLIN1
ERMP1
ERN1
FA2H
Golgi apparatus, Vesicle, Exosome
Endoplasmic reticulum


id
CLDN1
CLDN2
CLDN3
CLDN4
CLDN5
CLDN6
CLDN7
CLDN8
CLDN9
CLDN10
CLDN11
CLDN12
CLDN14
CLDN15
CLDN16
CLDN18.1
CLDN18.2
CLDN19A
CLDN19B
CLDN20
CLDN22
CLDN23
CLDN24
id
Annotation
FADS1
FADS3
FAF2
FAR2
FDFT1
FKBP10
FKBP11
GPAA1
GPAT3
GPAT4
GPX8
HERPUD1
HSD17B12
HSD3B7
JAGN1
LBR
LCLAT1
LNPK
LPCAT4
LPGAT1
MBOAT7
MLEC
NCLN
OAS1
OAS2
OS9
OSBPL5
OSBPL8
PEX3
PGS1
PIGG
PIGK
PIGX
PNPLA6
PNPLA7
POMT2
PRKCSH
PTDSS1
RDH14
RNF185
RNF5
RPN1
RPN2
SCD
SDF2
SDF2L1
SEC61B
SEC61G
SEL1L
SIL1
SPCS1
SPCS2
SPCS3
SQLE
SSR3
STT3A
STT3B
TECR
TMX1
TOR1AIP2
TRAM1
TRIM59
TXNDC12
Endoplasmic reticulum

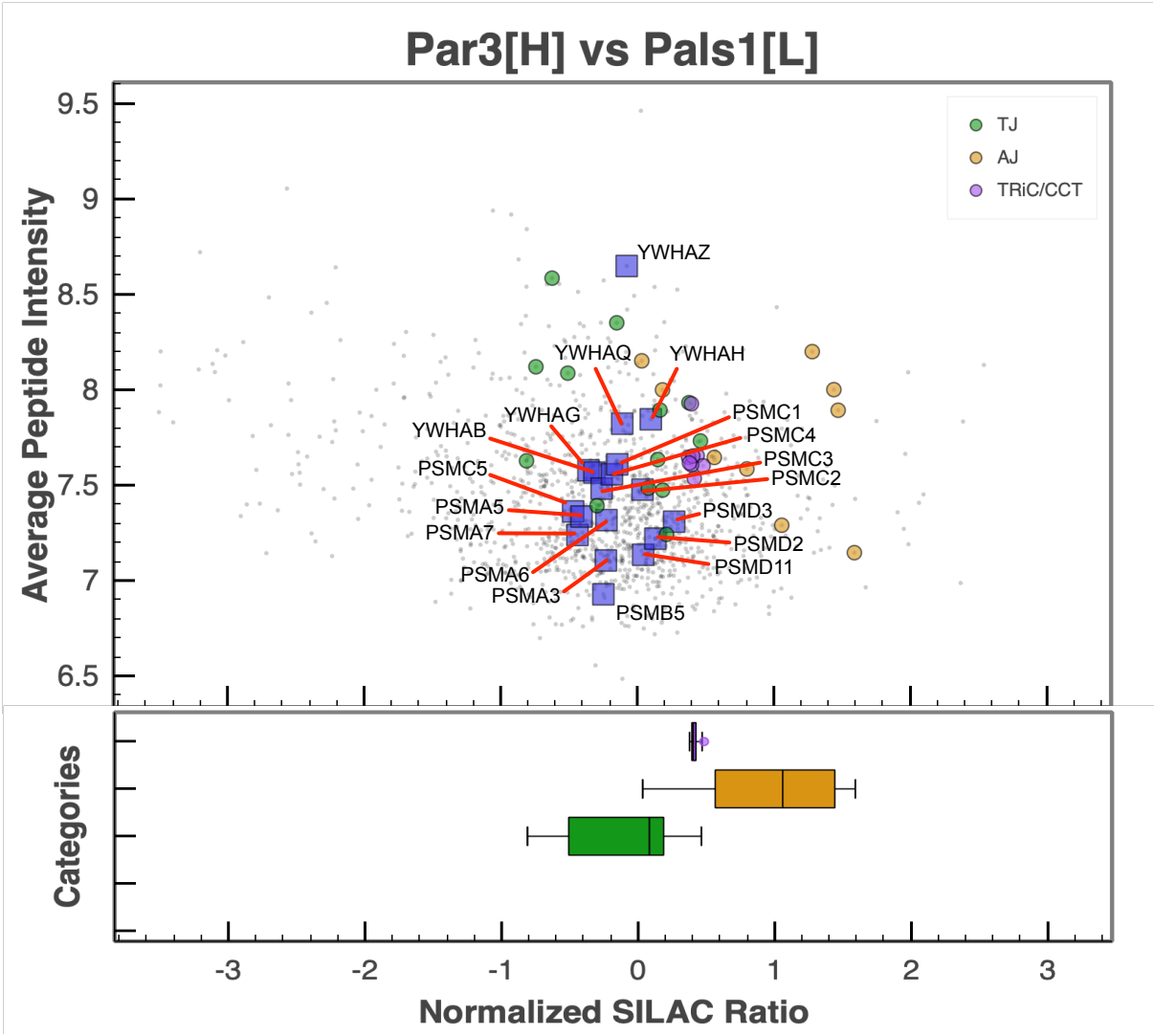

id
CLDN1
CLDN2
CLDN3
CLDN4
CLDN5
CLDN6
CLDN7
CLDN8
CLDN9
CLDN10
CLDN11
CLDN12
CLDN14
CLDN15
CLDN16
CLDN18.1
CLDN18.2
CLDN19A
CLDN19B
CLDN20
CLDN22
CLDN23
CLDN24
id
Annotation
UBAC2
UBE2G2
UFSP2
USP19
VRK2
WFS1
ZDHHC6
YME1L1
ABCB10
ABCB7
ABCB8
ABCD1
ABCD3
ACAD11
ACAD9
ACO2
AFG3L2
AIFM1
AKAP1
AMACR
APOOL
APOPT1
ARMC1
ATAD3A
BAG5
BDH1
C15orf48
CARS2
CCDC51
CDK5RAP1
CLPB
CLPP
CLPX
COA1
COMTD1
COX18
COX7A2L
CPT1A
CYC1
DARS2
DBT
DCAKD
DHTKD1
ECSIT
ETHE1
FASTKD5
FPGS
GRPEL1
HARS2
HEATR1
HSCB
IMMT
MAIP1
MCL1
MCU
METTL17
MTX1
NDUFA9
NDUFAF1
NDUFAF2
NDUFAF6
NDUFB11
NDUFC2
ER
Mitochondria


id
CLDN1
CLDN2
CLDN3
CLDN4
CLDN5
CLDN6
CLDN7
CLDN8
CLDN9
CLDN10
CLDN11
CLDN12
CLDN14
CLDN15
CLDN16
CLDN18.1
CLDN18.2
CLDN19A
CLDN19B
CLDN20
CLDN22
CLDN23
CLDN24
id
Annotation
NDUFS3
NDUFS8
NIPSNAP1
NLRX1
NNT
OXA1L
PARL
PC
POLRMT
PPM1K
PPOX
PYCR2
QARS
RDH13
RPUSD3
SCCPDH
SCO1
SFXN1
SFXN2
SFXN3
SFXN4
SLC25A12
SLC25A16
SLC25A18
SLC25A19
SLC25A21
SLC25A24
SLC25A42
SLC25A6
SMIM20
SPG7
SURF1
SYNJ2BP
TBRG4
TIMM17A
TIMM17B
TIMM21
TIMM23B
TIMM29
TIMM50
TIMM9
TMEM126B
TMEM70
TOMM22
UQCC3
UQCRB
UQCRC1
UQCRFS1
UQCRH
UQCRQ
NUP205
SUN1
TMEM201
AAAS
ACBD5
ARID1A
ARID5B
BSPRY
C2CD2
CDK5RAP3
DAZAP1
DDX39B
DNAJB2
Mitochondria
Nucleus


id
CLDN1
CLDN2
CLDN3
CLDN4
CLDN5
CLDN6
CLDN7
CLDN8
CLDN9
CLDN10
CLDN11
CLDN12
CLDN14
CLDN15
CLDN16
CLDN18.1
CLDN18.2
CLDN19A
CLDN19B
CLDN20
CLDN22
CLDN23
CLDN24
id
Annotation
DNTTIP2
EXD2
HELLS
HNRNPH2
IFRD1
ISG20
KIF23
KIF4A
LAMC2
LEMD3
LMNB1
LMNB2
LOX
MCM3AP
MICAL2
MSH2
NUP160
NUP88
NUSAP1
PBXIP1
PEX14
PGRMC2
PLRG1
PPP1R8
PRPF6
RBBP4
RBFOX2
RNF40
SLU7
SMARCA4
STK4
TMEM120A
UBE4A
WDR33
DHRS13
ABHD16A
ARHGAP11A
BAG2
C2CD2L
CGRRF1
CLPTM1L
CMTM4
CRIP1
CRIP2
CYP27C1
DNAJB12
DNAJC16
EMC3
FAR1
FGFR1OP2
FNDC3B
GK5
IFFO2
ILVBL
INCENP
LAMB3
MAP2
METTL15
PLEKHA5
PLPP6
RABL3
SIGIRR
SLC15A4
Nucleus
Others


id
CLDN1
CLDN2
CLDN3
CLDN4
CLDN5
CLDN6
CLDN7
CLDN8
CLDN9
CLDN10
CLDN11
CLDN12
CLDN14
CLDN15
CLDN16
CLDN18.1
CLDN18.2
CLDN19A
CLDN19B
CLDN20
CLDN22
CLDN23
CLDN24
id
Annotation
SLC35E1
SLC38A10
SNRPC
TAPT1
TM2D3
TM9SF3
TMEM161A
TMEM209
TMEM41B
TMEM45B
TMPRSS4
TMTC3
TMX2
TOR4A
TVP23B
ANKRD22
ATP5F1A
ATP5F1B
ATP5IF1
ATP5PB
ATP5PF
ATP5PO
C11H5orf15
C20H19orf70
C6H16orf58
C6H16orf91
CCDC56
CRYBG2
DDA1
DLA88
DMAC2
E2QV49
E2RNF5
E2RNY7
F1P7J0
F1P8D4
F1PAE8
F1PAK4
F1PDJ2
F1PNJ5
F1PP57
F1Q2N5
HSPA1
ITPRID2
J9NT25
J9NZZ6
KCNN
LOC100683871
LOC100855841
LOC102151856
LOC479708
LOC480667
LOC488767
LOC489992
LOC610994
LOC612644
METTL9
PHF3
PIS
PLEKHG3
RETREG1
SEC61
SH3TC1
SQOR
TRABD
Others
N/A

Suppl. Figure 3. . Complete list of claudin family interactors identified by CoIP classified by GOCC annotation. Significant interactors were annotated with DAVID using the database from human.



Suppl. Figure 4. Distribution of proteasome complex and CCT/TRIC complex subunits, and 14-3-3 proteins determined by quantitative proximity proteomics. The scatter plot shows the average log10 peptide intensities over the average log2(H/L) normalized SILAC ratios for all proteins identified by Tan et al. 2020. The box plot shows the median distribution of proteins from the categories Adherens Junction (AJ) and Tight Junction (TJ). Plot obtained from the interactive online portal of the Par3-Pals1 HCP. (Tan et al. (2020), Current Biology, Suppl. File S3)

List of publications

Perdomo-Ramirez, A., de Armas-Ortiz, M., Ramos-Trujillo, E., Suarez-Artiles, L., Claverie-Martin, F., 2019. **Exonic CLDN16 mutations associated with familial hypomagnesemia with hypercalciuria and nephrocalcinosis can induce deleterious mRNA alterations.** BMC Med Genet 20, 6. https://doi.org/10.1186/s12881-018-0713-7

Guneykaya, D., Ivanov, A., Hernandez, D.P., Haage, V., Wojtas, B., Meyer, N., Maricos, M., Jordan, P., Buonfiglioli, A., Gielniewski, B., Ochocka, N., Cömert, C., Friedrich, C., Artiles, L.S., Kaminska, B., Mertins, P., Beule, D., Kettenmann, H., Wolf, S.A., 2018. **Transcriptional and Translational Differences of Microglia from Male and Female Brains**. Cell Rep 24, 2773-2783.e6. https://doi.org/10.1016/j.celrep.2018.08.001

Suarez-Artiles, L., Perdomo-Ramirez, A., Ramos-Trujillo, E., Claverie-Martin, F., 2018. **Splicing Analysis of Exonic OCRL Mutations Causing Lowe Syndrome or Dent-2 Disease.** Genes (Basel) 9. https://doi.org/10.3390/genes9010015

Ramberger, E.*, Suarez-Artiles, L.*, Perez-Hernandez, D.*, Haji, M., Popp, O., Reimer,U., Leutz, A., Dittmar, G., Mertins, P., **A universal peptide matrix interactomics approach to disclose motif dependent protein binding.** *Submitted.* (* These authors contributed equally)

Acknowledgements

I want to start by thanking my supervisors Prof. Dr. Dominik Müller and Prof. Dr. Gunnar Dittmar, for giving me the excellent opportunity to work on this project and for their support. I would also like to express my gratitude towards Dr. Philipp Mertins, Dr. Tilman Breiderhoff, and Prof. Dr. Silke Rickert-Sperling for their support and feedback during the process of completing this project.

I would like to thank the Berlin Institute of Health and the Max-Delbrück Center for Molecular Medicine for providing funding and a stimulating environment for the development of my Ph.D. project.

I am grateful for all the help I got during this time, especially from Dr. Evelyn Ramberger, Dr. Daniel Pérez Hernádez, Dr. Dorothee Günzel, Hannes Gonschior, Dr. Antoine Lesur, Dr. Elena Martínez García, Dr. Petr Nazarov, and Dr. Rosanna Girardello. I also want to thank all my other former and present colleagues in the proteomics platform at the MDC for great discussions, helping me whenever needed, and fun times outside of the lab: Corinna Friedrich, Mohamad Haji, Sylvia Niquet, Dr. Marie-Luise Kirchner, Dr. Oliver Popp, Hans Werner, Merve Alp, Ann-Katrin Wagner, Valeria Sapozhnikova, Dr. Matthias Ziehm, Dr. Patrick Beaudette, Gülkiz Baytek, and Dr. Tamara Kanashova. I also want to thank Prof. Dr. Matthias Selbach and his group for the prolific discussions and input for this work. I also want to thank all members from the Advanced Light Microscopy platform, Flow Cytometry facility, and Leutz lab for all the technical support and all the members from the Dittmar lab at the Luxembourg Institute of Health.

I want to thank Dr. Rita Rosenthal and all members and students in the DFG Graduate School TJ-Train for kindly invite me to participate in their lectures and various scientific events, and give me the opportunity to present and discuss my project numerous times.

To my former colleagues and friends at the Hospital Universitario Nuestra Señora de la Candelaria, especially Ana Perdomo Ramírez, Manuel Rodriguez del Rio, Dr. Ernesto Martín Nuñez, and Dr. Javier Donate Correa, thank you for showing me how good teamwork looks like and for the great fun moments. I am forever grateful to my first mentors, Dr. Félix Claverie Martín and Dr. Hilaria González Acosta, who helped me set the foundations of my career in science, encouraged me to do my Ph.D. abroad, and always gave me the best advice.

Special thanks to my friends and former MDC colleagues Miguel Rodriguez de los Santos and Dr. Francesca Imbastari for being there in the good, the bad, and the crazy. Lastly, I want to thank my parents, big family, and friends for their unconditional and very much needed support and encouragement during the last five years.