Claudin gene expression profiles and clinical value in colorectal tumors classified according to their molecular subtype

Sara Cherradi
Pierre Martineau
Céline Gongora
Maguy Del Rio
Institut de Recherche en Cancérologie de Montpellier (IRCM), INSERM U1194, Université de Montpellier, Institut régional du Cancer de Montpellier, Montpellier F-34298, France

Purpose: Colorectal cancer (CRC) is a heterogeneous disease that can be classified into distinct molecular subtypes. The aims of this study were 1) to compare claudin (CLDN) gene expression in CRC samples and normal colon mucosa, and then in the different CRC molecular subtypes, and 2) to assess their prognostic value.

Patients and methods: CLDN expression in CRC samples was analyzed using gene expression data for a cohort of 143 primary CRC samples, and compared in the same CRC samples classified into different molecular subtypes (C1 to C6 according to the Marisa’s classification, and CMS1 to CMS4 of the consensus classification). Comparison of CLDN expression in normal and tumor colon samples was also made on a smaller number of samples. Then, the relationship between CLDN expression profiles and overall survival (OS) and progression-free survival was examined.

Results: Compared with normal mucosa, CLDN1 and CLDN2 were upregulated, whereas CLDN5, 7, 8, and 23 were downregulated in CRC samples. Variations in CLDN expression profiles were observed mainly in the CMS2/C1 and CMS4/C4 subtypes. Overall, expression of CLDN2 or CLDN4 alone had a strong prognostic value that increased when they were associated. In the CMS4/C4 subtypes, lower expressions of CLDN11, CLDN12, and CLDN23 were associated with longer OS. Conversely, in the CMS2 and C1 subtypes, low CLDN23 expression was associated with shorter OS and progression-free survival, suggesting a dual role for CLDN23 as a tumor suppressor/promoter in CRC. CLDN6 and CLDN11 had a prognostic value in the CMS2 and C4 subtypes, respectively.

Conclusion: This analysis of CLDN gene expression profiles and prognostic value in CRC samples classified according to their molecular subtype shows that CRC heterogeneity must be taken into account when assessing CLDN potential value as prognostic markers or therapeutic targets.

Keywords: colon, cancer, classification, prognosis, claudin, target, tight junction, heterogeneity

Introduction

Colorectal cancer (CRC) originates from the oncogenic transformation of the intestinal epithelium that physiologically acts as a functional barrier between the intestinal mucosa and the luminal environment. It is now becoming clear that epithelial cell polarity is a major gatekeeper against cancer initiation and metastasis formation. Epithelial cell polarity depends on the establishment of the apical junctional complex that includes tight junctions (TJs) and adherens junctions. TJs, the most apical of these intercellular junctions, play an essential role in maintaining cell polarity and in...
the regulation of paracellular permeability. Alterations in TJs by downregulation or upregulation of TJ proteins can trigger malignant transformation and influence cancer progression.

Claudins (CLDNs) are TJ core components that are essential for TJ formation and contribute to their selectivity. In mammals, the CLDN family includes 27 members divided into two groups: classic and nonclassic CLDNs. Almost all CLDNs have a short intracellular N-terminal domain, four transmembrane domains, two extracellular loops, and an intracellular C-terminal domain that contains a PDZ-domain-binding motif for linking to TJ-associated proteins, such as MUPP1, PATJ, ZO-1, ZO-2 and ZO-3, MAGUKs, PAR3, PAR6, and PALS. These proteins function as adaptors at the cytoplasmic surface of TJ strands and can directly or indirectly interact with cytosolic and nuclear proteins, for instance cytoskeletal molecules, regulatory proteins, tumor suppressors, and transcription factors. Finally, some CLDNs interact with cell adhesion proteins (eg, EPCAM) or receptors (eg, EPHA and EPHB). The cytoplasmic tail of most CLDNs contains a large number of predicted phosphorylation sites that could be involved in molecular interactions. Accumulated evidence indicates that CLDNs are associated with various pathways, including the WNT/β-catenin, JAK-STAT3, and Notch signaling cascades.

CLDNs are expressed in a cell- and tissue-specific manner. In the intestine, CLDNs display specific spatiotemporal expression profiles with variations along the crypt–lumen axis. Their expression can be regulated by various mechanisms at the transcriptional or posttranscriptional level, but also via mRNA stability modulation and through epigenetic mechanisms.

CLDN expression is altered in several cancer types in a tumor-specific manner, and can vary according to the tumor stage. CLDN aberrant expression in tumors may have opposite functions (promotion of tumorigenesis and metastasis formation, or suppressive effects). For example, CLDN-1 is a cancer invasion/metastasis suppressor in lung adenocarcinoma, while in CRC, CLDN1 expression enhances the invasive ability and metastatic properties. Moreover, some CLDNs have an important regulatory role in the epithelial–mesenchymal transition (EMT). CLDNs can also serve as a hub for different signaling proteins, and therefore could have a critical role in the regulation of carcinogenesis or cancer progression. Finally, CLDN expression has been associated with patient survival, suggesting that they could be prognostic markers and/or therapeutic targets.

In CRC, research has focused mainly on CLDN1, CLDN2, CLDN18, CLDN30, and CLDN7. Changes in the expression of TJ-related genes, including CLDNs, have been reported in CRC; however, these studies did not consider CRC heterogeneity, which can be described using molecular subtypes based on gene expression profiles.

In this study, CLDN gene expression profiles were investigated in a cohort of 143 primary CRC samples classified according to their molecular subtype and for which gene expression and clinical data were available. The expression of each CLDN gene was first compared in normal and tumor colon samples, and then among the different CRC molecular subtypes. Finally, the prognostic value of the different expression profiles was evaluated.

**Patients and methods**

**Gene expression analysis**

In this study, expression data for tumor samples from 143 patients coming from three cohorts (REG/P, COSIVAL, and BIOCOLON) were used. These three studies were approved by the relevant ethics committees and all participants were informed about the study, and they signed a written informed consent before enrolment. All patients selected for this study had metastatic colorectal cancer (mCRC), and did not receive any chemotherapy treatment before primary tumor resection. Colon samples (normal colon, primary tumor, and hepatic metastasis samples from the REG/P cohort, and only primary tumor specimens from the COSIVAL and BIOCOLON cohorts) were collected at the time of surgery, following a standardized procedure to obtain high-quality RNA. Samples were then hybridized to human genome U133 Plus 2.0 arrays (Affymetrix Inc., Santa Clara, CA, USA). The gene expression data can be found online at the Gene Expression Omnibus under the accession numbers GSE62080 and GSE72970.

All 143 CRC samples were classified using the molecular classifications based on gene expression profiles that have been proposed by Marisa et al and Guinney et al (Table 1), as described in each reference publication. Briefly, Marisa et al described six molecular subtypes (C1 to C6) with the following main features: C1 = CIN and immune pathway downregulation; C2 = MSI; C3 = mutated KRAS; C4 = stem cell phenotype-like; C5 = CIN and upregulation of the WNT pathways; and C6 = CIN and normal-like gene expression profile. The consensus classification includes four subtypes: CMS1 (microsatellite instability [MSI]-immune), CMS2 (epithelial and canonical), CMS3 (epithelial and metabolic), and CMS4 (mesenchymal). Most of the MSI-high tumors belong to the CMS1 subtype that has the best survival without recurrence. Chromosomal instability (CIN) tumors have a
more heterogeneous gene expression pattern and, therefore, can be classified from CMS2 to CMS4. CMS4 tumors have a significantly higher risk of distant relapse.38,39

The Tsuji cohort (GSE28702)44 includes 83 stage IV CRC samples from patients treated with the FOLFOX regimen. Dr Shingo Tsuji kindly provided the overall survival (OS) data for this series.

### Results

#### CLDN gene expression patterns in colon

Analysis of the CLDN gene expression levels from Affymetrix data for 17 normal colon mucosa, 20 primary CRC samples, and 19 hepatic metastases (REG/P cohort)40,41 showed that 8 CLDN genes (CLDN6, 9, 10, 11, 14, 16, 17, and 18) were weakly expressed in all samples. Among the other ten claudins, CLDN1, CLDN3, CLDN4, and CLDN7 were strongly expressed in primary CRC and metastatic samples, and CLDN3, CLDN4, CLDN7, CLDN8, and CLDN23 in normal mucosa (Figure 1A).

To determine whether CLDN gene expression levels change during tumorigenesis, CLDN expression profiles in normal mucosa, primary tumor, and hepatic metastases were compared. CLDN4, 6, 9, 10, 12, 14, 15, 16, 17, and 18 displayed similar expression levels in all tissue samples, whereas CLDN3 and CLDN11 showed a tendency to downregulation in primary tumor samples (Figure S1). On the other hand, CLDN1 and CLDN2 were significantly upregulated, and CLDN5, 7, 8, and 23 downregulated in primary tumor and hepatic metastasis samples compared with normal mucosa (Figure 1B). These results were validated in 15 matched samples from the study cohort (normal mucosa and primary tumor from the same patient; Figure S2). The largest variation of expression between paired samples was observed for CLDN2 and CLDN8.

#### Comparison of CLDN gene expression in the different CRC molecular subtypes

CLDN expression was then analyzed in all 143 primary CRC samples classified in molecular subtypes according to Marisa et al (C1 to C6) and the consensus classification (CMS1 to 4; see Table 1 and Figure 2 for the relationships between classifications).42

Among the CLDN genes that were upregulated in tumors compared with normal mucosa, CLDN1 expression was significantly higher in the CMS2 subtype and in the C1 and C5 subtypes, while CLDN2 expression was similar in all subtypes (Figure 3A and Figure S3A). Among the CLDN genes downregulated in tumors, CLDN8 expression reduction was less pronounced in the CMS3 and C6 subtypes, while CLDN7 and CLDN23 were more downregulated in the CMS4 and C4 subtypes (Figure 3B and Figure S3B). Among the CLDN genes with comparable expression in normal and tumor tissues, CLDN3 and CLDN4 showed lower expressions in the CMS4 subtype (not significant for CLDN4) and in the C4 subtype (Figure 3C and Figure S3C). Among the CLDN genes that were weakly expressed in primary tumors, CLDN6 expression level was higher in the CMS2 and C1 subtypes, and CLDN11 expression was strongest in the CMS4 and C4 subtypes (Figure 3C and Figure S3C). The other CLDN genes did not show any significant expression level difference among CRC subtypes.

<p>| Table 1 Distribution of patients with mCRC according to the tumor molecular subtype |</p>
<table>
<thead>
<tr>
<th>Subtypes</th>
<th>N=143</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marisa (n=138)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>27</td>
<td>19.6</td>
</tr>
<tr>
<td>C2</td>
<td>17</td>
<td>12.3</td>
</tr>
<tr>
<td>C3</td>
<td>19</td>
<td>13.8</td>
</tr>
<tr>
<td>C4</td>
<td>20</td>
<td>14.5</td>
</tr>
<tr>
<td>C5</td>
<td>36</td>
<td>26.1</td>
</tr>
<tr>
<td>C6</td>
<td>19</td>
<td>13.8</td>
</tr>
<tr>
<td>Missing</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Consensus (n=104)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMS1</td>
<td>15</td>
<td>14.4</td>
</tr>
<tr>
<td>CMS2</td>
<td>29</td>
<td>27.9</td>
</tr>
<tr>
<td>CMS3</td>
<td>24</td>
<td>23.1</td>
</tr>
<tr>
<td>CMS4</td>
<td>36</td>
<td>34.6</td>
</tr>
<tr>
<td>Missing</td>
<td>39</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: mCRC, metastatic colorectal cancer.

For personal use only.
Association between CLDN expression level and OS and PFS

Then, the correlation between CLDN gene expression level and PFS and OS rates was evaluated in the 143 patients divided into two groups (high/low expression relative to the median CLDN gene expression) of the same size. This analysis found a significant correlation only between CLDN2 gene expression (low/high) and OS ($P=0.03$) and PFS ($P=0.005$), and between CLDN4 gene expression and OS ($P=0.01$; Figure 4A and B). These results were confirmed in an independent cohort of 80 patients (Tsuji’s cohort;44 Figure S4A).

As low CLDN2 expression and high CLDN4 tumor expression were good prognostic factors in patients with CRC, the next step was to assess whether OS was correlated with the expression of both CLDNs. OS rate was significantly higher in patients with CRC displaying concomitant low CLDN2 and high CLDN4 expressions, particularly compared with patients with high CLDN2 and low CLDN4 tumor expressions ($P=0.006$; Figure 4C). OS time was more than twofold longer in patients with low CLDN2/high CLDN4 tumors than in those with high CLDN2/low CLDN4 tumors (38 months vs 14.4 months). Moreover, analysis of the correlation between OS and CLDN2 and CLDN4 gene expressions according to the CRC molecular subtype confirmed the significant associations between CLDN2 and CLDN4 expressions and OS for the CMS3 and C3 and the CMS2 and C1 subtypes, respectively (Figure S4B).

Similarly, high expressions of CLDN3 and CLDN23 and low expression of CLDN6 in the CMS2 subtype were associated with longer OS (Figure 5). Low CLDN6 expression...
in the CMS2 subtype and high CLDN23 expression in the CMS2 and C1 subtypes were correlated with better PFS rates (Figure S5). In the CMS4 and C4 subtypes, low expression of CLDN12 and CLDN23 was associated with longer OS (Figure 5). In the C4 subtype, low CLDN11 expression was correlated with longer OS (Figure 5) and also PFS (Figure S5). Finally, high CLDN8 expression was associated with longer OS in the C1 and C3 subtypes (Figure 5).

Discussion
In this study, using Affymetrix gene expression data from 143 patients with CRC categorized according the CRC molecular subtypes,12 we showed that CLDN gene expression profiles vary according to the tumor stage (normal tissue, primary tumor, and metastasis) and also the molecular subtype. We assessed CLDN gene expression in the four subtypes proposed by the consensus molecular classification,39 and in the six subtypes defined by Marisa et al36 to refine the tumor type identification. Our analysis (summary in Figure 6) indicated that eight CLDNs were differentially expressed in CRC tumors compared with normal mucosa. CLDN expression alterations were mainly seen in the CMS2/C1 and CMS4/C4 subtypes. The changes in the expression of these eight CLDN genes were associated with a prognostic value in the whole cohort of CRC samples and also in specific molecular subtypes. Conversely, we did not detect any differential expression or prognostic value of CLDN expression in the CMS1 and C2 subtypes. This could be due to the low number of tumors in these two groups (n=15 and 17, respectively).

CLDN gene expression in normal colon tissues is consistent with a previous study where the expression patterns were analyzed along the proximal–distal axis of the human intestine by real-time PCR.45 In our study, CLDN3, 4, 7, and 23 displayed the highest expressions in normal colon epithelial cells with a predominant expression of CLDN7. CLDN3, 4, and 7 are strongly expressed in several normal tissues, including colon, and their similar expression and localization profiles suggest a coordinated regulation.46 These CLDNs have an important role in the maintenance of homeostasis of colon epithelium. CLDNs 3 and 4 have been classified as colon barrier proteins.47,48 CLDN 7 maintains the barrier function of the intestinal epithelium and regulates...
epithelial cell renewal.\textsuperscript{49} CLDN 7 is localized at apical TJs and also in basolateral membranes, and CLDN 7-mediated cell–matrix interaction is indispensable in the intestine.\textsuperscript{50} Except for CLDN4, their expression was downregulated in CRC samples. CLDN3 and CLDN7 downregulation was more pronounced in the mesenchymal CMS4 and in the stem cell phenotype-like C4 subtypes. CMS4 and C4 tumors are characterized by activation of pathways related to EMT and stemness.\textsuperscript{36,38,39} Bhat et al\textsuperscript{24} reported that CLDN7 has a key role in EMT regulation in colon epithelial cells, and that low CLDN7 expression promotes EMT and tumor progression. CLDN7 is also frequently associated with the stem cell marker EPCAM,\textsuperscript{51} and this association could contribute to EMT.\textsuperscript{52} CLDN8 displayed the strongest downregulation in CRC samples, but little is known about its function in the colon. CLDN8 was identified as a critical downstream component of the IL9 inflammatory cascade in inflammatory bowel disease.\textsuperscript{53} Here, we found that in the C1 and C3 subtypes, CLDN8 expression has a prognostic value. The common features of these two subtypes are KRAS mutations and the suppression of pathways associated with activation of the immune system and EMT.\textsuperscript{36} Moreover, higher expressions of CLDN3 and CLDN4 were associated with a better outcome in patients with CMS2 and C1 cancers. These findings are in agreement with CLDN3 and CLDN7 tumor-suppressive functions;\textsuperscript{33,54,55} conversely, the possible tumor suppressor role of CLDN4 and CLDN8 remains to be demonstrated.

CLDN23 expression was significantly reduced in CRC samples, as previously reported in different intestinal cancer types.\textsuperscript{56} including CRC.\textsuperscript{17,35,57} Like CLDN3, CLDN4, and CLDN7, CLDN23 was only slightly downregulated in the CMS4 and C4 subtypes, and this was correlated with longer OS. We can hypothesize that in these subtypes, CLDN23 expression is regulated by stromal suppressor
genes. Conversely, in the canonical CMS2 and C1 subtypes, patients with low CLDN23 tumor expression had shorter OS and PFS, in agreement with a previous analysis in 53 patients with CRC.57 CLDN23 seems to have dual role as a tumor suppressor and a tumor promoter, depending on the CRC subtype where it is expressed. This hypothesis requires additional investigations because very few studies have assessed CLDN23’s role in CRC since its first characterization following the observation that CLDN23 is downregulated in gastric cancer.56

On the other hand, the expressions of CLDN1 and CLDN2 were significantly increased in tumor tissues, compared with normal mucosa. This confirms previous reports on their upregulation in CRC and their involvement in CRC cell tumorigenicity.22,28,58,59 Induction of CLDN1 and CLDN2 expressions has been related to overactivation

Figure 4 Clinical value of CLDN gene expression in the 143 CRC samples.
Notes: Association between CLDN2 and CLDN4 gene expression levels and (A) OS and (B) PFS (log-rank test). High and low: higher or lower expressions relative to the median expression for that CLDN gene. (C) Association of combined CLDN2 and CLDN4 gene expressions with OS.
Abbreviations: CRC, colorectal cancer; OS, overall survival; PFS, progression-free survival.
of WNT/β-catenin signaling in CRC cells. Moreover, expressions of CLDN1 and CLDN2 have been associated with EMT and cancer progression. We previously showed that CLDN1 expression is higher in the canonical CMS2 subtype, which includes the C1 and C5 subtypes, and represents epithelial tumors with marked upregulation of WNT and MYC downstream targets, and that PFS is significantly longer in patients with C3 and C5 tumors with low CLDN1 expression. Here, we found that CLDN2 level is not different among CRC subtypes, but has a strong prognostic value in all patients and also in the CMS3 and C3 subtypes, which often harbor KRAS-activating mutations. Our findings are in agreement with recent data showing that high CLDN2 expression is linked to posttreatment recurrence in patients with stage II/III CRC. This makes CLDN2 a good candidate for therapeutic target in CRC.

Three other claudin genes (CLDN6, CLDN11, and CLDN12) also displayed a prognostic value. CLDN6 was more expressed in the CMS2 and C1 subtypes, and low expression in CMS2 was associated with longer OS and PFS. CLDN6 expression and function have never been studied in CRC. It was described as a cancer-promoting factor in gastric cancer, and as a tumor suppressor in breast cancer. Low CLDN11 expression correlation with longer OS and PFS in the C4 subtype (at high risk of relapse), and CLDN11 upregulation in the CMS4 and C4 subtypes suggest a pro-tumor function in CRC. CLDN11 could play a role in the TGFβ1-OCLN/CLDN11 paracrine axis between cancer cells and cancer-associated fibroblasts in CRC. CLDN12 was strongly expressed in both normal and tumor colon tissues. In the CMS4 and C4 subtypes, low expression was associated with better prognosis. CLDN12 contributes to Ca²⁺ absorption in intestinal epithelial cells, and it is one of the few claudins that do not possess a PDZ binding motif.

The specific biological features of the CRC subtypes could explain the differences in CLDN gene expression.
Figure 6 Summary of CLDN gene expression and prognostic value in CRC.

Notes: (1) Expression in CRC samples (n=20) compared with normal mucosa samples (n=17); red = upregulated; green = downregulated; white, no expression difference with normal mucosa. (2) Comparison of CLDN expression in the different CRC subtypes; red = subtype where CLDN expression is strongest; green = subtype where the expression of that CLDN gene is lowest compared with the other subtypes; white = no expression difference among subtypes. YES indicates that the expression level of that CLDN gene shows prognostic value, with the correlation between CLDN expression level (high/low relative to the median expression of that gene) and improved prognosis.

Abbreviation: CRC, colorectal cancer.
Indeed, the gene expression profiles of claudins involved in EMT (CLDN7), acting as tumor suppressors (CLDN3, 7, 23) or linked to the stem cell phenotype (CLDN7), were markedly different in CMS4/C4 tumors (Figure 6). In CMS2/C1 and C5 tumors, CLDN gene expression seemed to be more dependent on WNT signaling, as already described for CLDN1.

CLDN gene downregulation could be explained by alterations in their sequences or in epigenetic regulation mechanisms, such as histone modifications, DNA methylation, and chromatin remodeling. This was demonstrated for CLDN1,16 CLDN2,18 CLDN7,19 and CLDN23.17

Changes in CLDN gene expression were essentially seen in the CMS2 and CMS4 subtypes for which a clear distinction in sensitivity to chemotherapy-induced apoptosis has been reported.67 This suggests that in CRC, claudins could play a role in chemoresistance, as previously shown for CLDN4 and CLDN7 in ovarian cancer,68,69 and more recently, for CLDN6 in triple-negative breast cancer cells.70

Altogether, we showed that CLDN gene expression has a significant clinical relevance. First, in the whole sample, the expressions of CLDN2 and CLDN4 alone showed a strong prognostic value that was increased when both genes were associated (median survival: 14.41 months for patients with high CLDN2/low CLDN4 expression and 38 months for patients with low CLDN2/high CLDN4 expression; HR =0.38). If these expression data are confirmed at the protein level, a prognostic gene/protein expression score can be developed for CRC patients. Second, for almost all CRC subtypes, we identified at least one CLDN gene, the expression of which was correlated with survival (Figure 6). In addition, we defined a CMS4/C4 tumor subtype signature characterized by low CLDN3, 4, 7, and 23 expressions associated with high CLDN11 expression. Moreover, in the C1 subtype, high CLDN3, 4, 8, and 23 expressions could be associated with good prognosis. In a breast cancer subtype, a signature characterized by low CLDN gene expression was associated with an aggressive phenotype.71 Finally, the best candidates as therapeutic targets seem to be CLDN6 in CMS2 tumors and CLDN11 in C4 tumors because they are overexpressed in these subtypes and low expression was associated with better prognosis. Future studies should thoroughly investigate their value as prognostic markers or therapeutic targets.

Conclusion

CLDN gene expression differences in CRC reflect CRC heterogeneity, and indicate that it is not enough simply to examine CLDN expression level globally. Moreover, claudin functions are not limited to TJs. Indeed, growing evidence shows that claudins are localized to sites outside the TJ complex,72 or even delocalized from the membrane to the nucleus, thus acquiring an important role in tumorigenesis. Finally, our findings demonstrated that CLDN expression is regulated in a CRC molecular subtype manner, and highlighted that to evaluate claudin function and potential value as prognostic markers or therapeutic targets, it is essential to take into account CRC heterogeneity.

Data sharing statement

Transcript profiles: GSE62080, GSE72970, GSE28702.

Acknowledgments

This work was supported by grants from the French National Research Agency under the program “Investissementsd’avenir”, Grant Agreement LabExMABImprove: ANR-10-LABX-53, and SIRIC Montpellier Cancer Grant: “INCa-DGOS-Inserm 6045”.

Disclosure

The authors report no conflicts of interest in this work.

References

34. Wang X, Tully O, Ngo B, Zitin M, Mullin JM. Epithelial tight junc-
31. Buchert M, Papin M, Bonnans C. Symplekin promotes tumorigenic-
29. Ouban A. Claudin-1 role in colon cancer: an update and a review.
26. Singh AB, Dhawan P. Claudins and cancer: fall of the soldiers entrusted
23. Joyce T, Cantarella D, Isella C, Medico E, Pintzas A. A molecular


