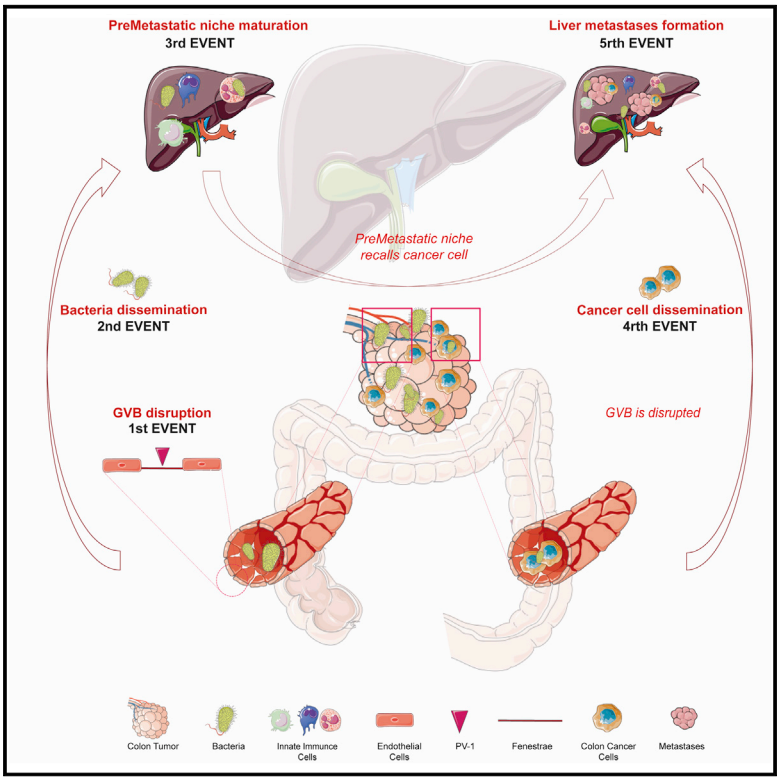
Cancer Cell

Article

# Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver

## Graphical abstract




## Highlights

- The endothelial marker PV-1 is an independent marker of CRC recurrence
- Specific tumor-resident bacteria modulate PV-1 via a *virf1*-dependent mechanism
- Increased PV-1 detection correlates with bacteria translocation and liver metastases
- Migrated bacteria induce the premetastatic niche in the liver

## Authors

Alice Bertocchi, Sara Carloni, Paola Simona Ravenda, ..., Chiara Pozzi, Giuseppe Penna, Maria Rescigno

## Correspondence

maria.rescigno@hunimed.eu

## In brief

Bertocchi et al. demonstrate the critical role of the gut vascular barrier (GVB) in the hematogenous route of liver CRC metastases. They link GVB impairment with bacteria translocation into the liver, the formation of a premetastatic niche, and tumor cell seeding. Targeting tumor-associated bacteria might be a new strategy to counteract liver metastases.

Bertocchi et al., 2021, Cancer Cell *39*, 1–17
May 10, 2021 
https://doi.org/10.1016/j.ccell.2021.03.004

CellPress

Cancer Cell

CellPress

Article

# Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver

Alice Bertocchi,[1] Sara Carloni,[1,2,9] Paola Simona Ravenda,[3,9] Giovanni Bertalot,[3] Ilaria Spadoni,[1,2] Antonino Lo Cascio,[1,2] Sara Gandini,[3] Michela Lizier,[1] Daniele Braga,[1] Francesco Asnicar,[4] Nicola Segata,[4] Chris Klaver,[3,10] Paola Brescia,[1] Elio Rossi,[5] Achille Anselmo,[1] Silvia Guglietta,[3,11] Annalisa Maroli,[1] Paola Spaggiari,[1] Noelia Tarazona,[6] Andres Cervantes,[6] Silvia Marsoni,[7] Luca Lazzari,[7] Maria Giovanna Jodice,[3] Chiara Luise,[3,12] Marco Erreni,[1] Salvatore Pece,[3,8] Pier Paolo Di Fiore,[3,8] Giuseppe Viale,[3,8] Antonino Spinelli,[1,2] Chiara Pozzi,[1] Giuseppe Penna,[1] and Maria Rescigno[1,2,13,*]

[1]IRCCS Humanitas Research Hospital, via Manzoni 56, Rozzano, Milan 20089, Italy
[2]Department of Biomedical Sciences, Humanitas University, Via Rita Levi Montalcini, Pieve Emanuele, MI 20072, Italy
[3]IEO European Institute of Oncology IRCCS, Milan 20141, Italy
[4]CIBIO Department, University of Trento, Trento, Italy
[5]Department of Biosciences, Università degli Studi di Milano, Via Celoria 26, Milan 20133, Italy
[6]Biomedical Research Institute INCLIVA, Hospital Clínico Universitario, Department Medical Oncology, University of Valencia, Valencia, Spain
[7]IFOM – the FIRC Institute of Molecular Oncology, via Adamello 16, Milano, MI 20139, Italy
[8]Department of Oncology and Hemato-Oncology, Università degli Studi di Milano, Milan 20142, Italy
[9]These authors contributed equally
[10]Present address: Division of Tumor Biology & Immunology, Netherlands Cancer Institute, Amsterdam, the Netherlands
[11]Present address: Medical University of South Carolina and Hollings Cancer Center, Department of Microbiology and Immunology, 173 Ashley Ave. 29,425, Charleston, SC, USA
[12]Present address: Tethis SPA, Milan, Italy
[13]Lead contact
*Correspondence: maria.rescigno@hunimed.eu
https://doi.org/10.1016/j.ccell.2021.03.004

## SUMMARY

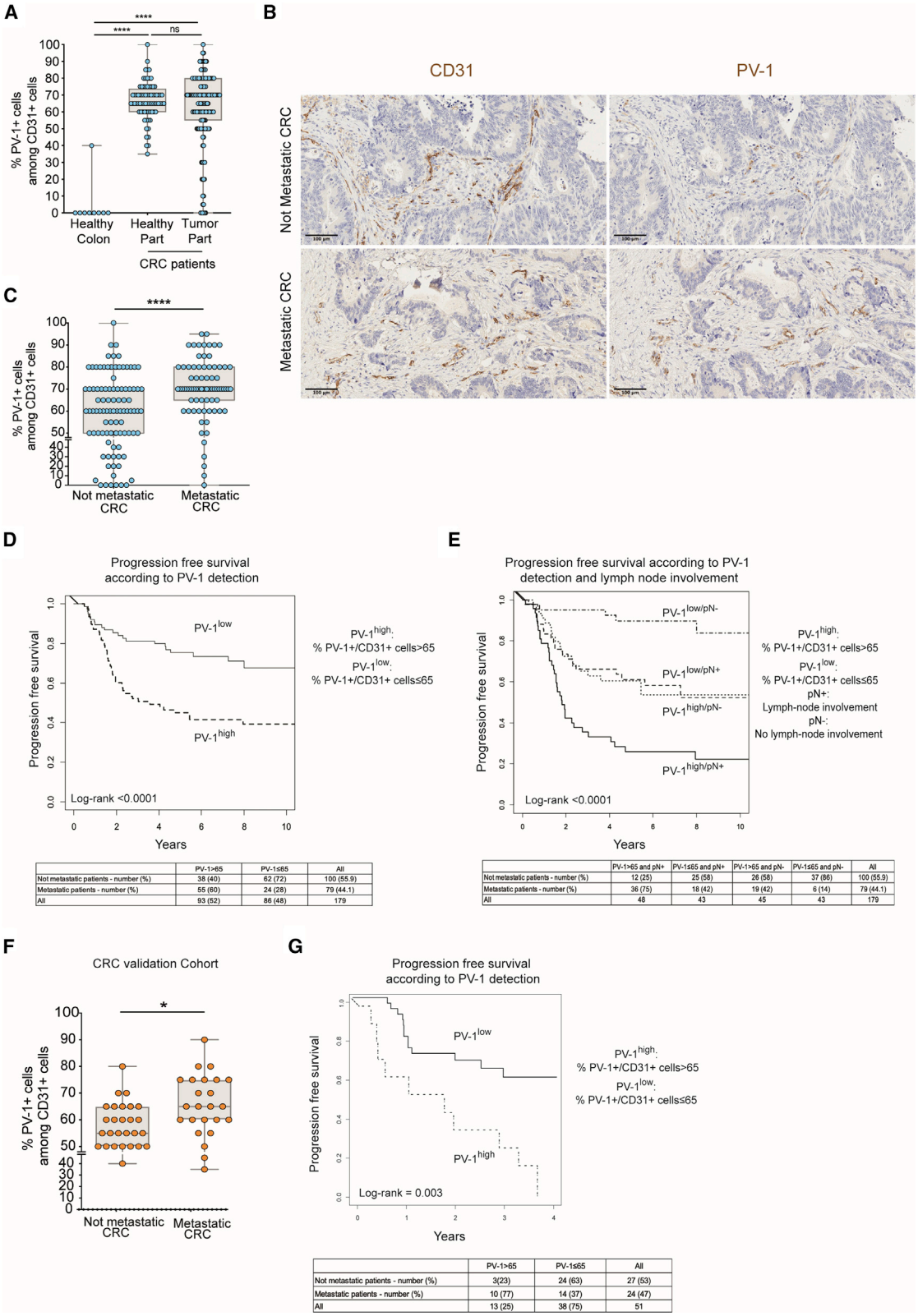
Metastasis is facilitated by the formation of a "premetastatic niche," which is fostered by primary tumor-derived factors. Colorectal cancer (CRC) metastasizes mainly to the liver. We show that the premetastatic niche in the liver is induced by bacteria dissemination from primary CRC. We report that tumor-resident bacteria *Escherichia coli* disrupt the gut vascular barrier (GVB), an anatomical structure controlling bacterial dissemination along the gut-liver axis, depending on the virulence regulator VirF. Upon GVB impairment, bacteria disseminate to the liver, boost the formation of a premetastatic niche, and favor the recruitment of metastatic cells. In training and validation cohorts of CRC patients, we find that the increased levels of PV-1, a marker of impaired GVB, is associated with liver bacteria dissemination and metachronous distant metastases. Thus, PV-1 is a prognostic marker for CRC distant recurrence and vascular impairment, leading to liver metastases.

## INTRODUCTION

Beyond the mucus and intestinal epithelial barrier, the gut is equipped with a gut vascular barrier (GVB), which acts as a gatekeeper to control the access of molecules and microorganisms in the systemic blood circulation (Spadoni et al., 2015, 2016). Under hazardous conditions, harmful enteric pathogens can cross the epithelial barrier and cause GVB damage, which eventually allows the dissemination of bacteria or their constituents into the blood circulation and the liver (Spadoni et al., 2015, 2016). GVB damage occurs also under particular dietary regimens, such as a high-fat or a methionine- and choline-deficient diet, and that the alteration is dependent on diet-induced modified microbiota (Mouries et al., 2019), indicating that also endogenous microbiota members can alter the GVB. Defective GVB and subsequent increased blood vessel permeability can be assessed by increased plasmalemma vesicle-associated protein-1 (PV-1) detection (Spadoni et al., 2015, 2016), a blood vessel endothelial-specific transmembrane protein that is associated to the diaphragms of the fenestrated endothelium (Rantakari et al., 2016; Stan et al., 2012). Here, we hypothesize that GVB disruption may be exploited by colorectal cancer (CRC) cells to metastasize to the liver. CRC is the second leading cause of cancer-related mortality (Bray et al., 2018). Most CRC patients eventually develop distant metastasis within 5 years from primary tumor diagnosis (metachronous metastatic patients) (Pita-Fernández et al., 2015). Distant metastases are facilitated by formation of a "premetastatic niche (PMN)," which is a site of immune deregulation, owing to the accumulation of innate immune cells and the presence of a pro-tumorigenic inflammatory milieu

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004






| | PV-1>65 | PV-1≤65 | All |
|---|---|---|---|
| Not metastatic patients - number (%) | 38 (40) | 62 (72) | 100 (55.9) |
| Metastatic patients - number (%) | 55 (60) | 24 (28) | 79 (44.1) |
| All | 93 (52) | 86 (48) | 179 |

| | PV-1>65 and pN+ | PV-1≤65 and pN+ | PV-1>65 and pN- | PV-1≤65 and pN- | All |
|---|---|---|---|---|---|
| Not metastatic patients - number (%) | 12 (25) | 25 (58) | 26 (58) | 37 (86) | 100 (55.9) |
| Metastatic patients - number (%) | 36 (75) | 18 (42) | 19 (42) | 6 (14) | 79 (44.1) |
| All | 48 | 43 | 45 | 43 | 179 |

| | PV-1>65 | PV-1≤65 | All |
|---|---|---|---|
| Not metastatic patients - number (%) | 3(23) | 24 (63) | 27 (53) |
| Metastatic patients - number (%) | 10 (77) | 14 (37) | 24 (47) |
| All | 13 (25) | 38 (75) | 51 |

**Figure 1. PV-1 is a prognostic biomarker of distant recurrence**

(A) Percentage of PV-1+ cells among CD31+ cells in tumor (n = 179), paired healthy mucosae (n = 76) of CRC patients and healthy colons (n = 10).

(B) Representative images of CD31 and PV-1 staining on FFPE specimens from CRC patients with/without distant metastases. Scale bars, 100 μm.

*(legend continued on next page)*

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004



before the arrival of metastatic cells (Costa-Silva et al., 2015; Erler et al., 2009; Hiratsuka et al., 2006, 2008; Hoshino et al., 2015; Qian et al., 2011; Seubert et al., 2015). CRC is also associated to the presence of an intratumoral microbiota known to influence cancer development and progression (Garrett, 2015; Bullman et al., 2017). Bullman et al. (2017) have shown that primary colon cancer and paired liver metastases are colonized by identical bacteria, such as *Fusobacterium nucleatum*, *Bacteroides fragilis*, and *Prevotella*. It is still unknown whether bacteria may participate also in metastasis formation and may colonize distant organs before, after, or concomitantly to cancer cell spreading. Here, we evaluated whether CRC liver metastases formation correlated with GVB opening, microbiota translocation into the liver, and formation of a PMN.

## RESULTS

### PV-1 is a prognostic biomarker for predicting CRC recurrence

To investigate the state of the GVB in colon vessels, we performed a retrospective analysis (with a follow-up of at least 5 years) on formalin-fixed paraffin-embedded (FFPE) resected human colons from CRC patients at the time of diagnosis (without distant metastases, n = 179) and healthy subjects (n = 10) of two cohorts from two different institutes in Milan (Table S1). First, we observed that the frequency of PV-1+ cells among CD31+ endothelial cells was higher in CRC patients as compared with healthy control samples. Notably, PV-1 detection was similar between paired cancerous and non-cancerous tissues in CRC patients, suggesting that the GVB is altered not only inside the tumor tissue but also in the adjacent healthy mucosa (Figure 1A). To assess if higher PV-1 levels in CRC patients correlated with the development of distant metastases over time, we separated samples from CRC patients who developed (n = 79) or not (n = 100) metachronous distant metastases (Table S1). This analysis showed that primary tumors of CRC patients who developed metachronous distant metastases exhibited a higher frequency of PV-1+ cells compared with patients who did not develop distant metastases in the follow-up time (Figures 1B and 1C). We stratified the patients into "PV-1$^{high}$" (%PV-1+/CD31+ cells >65) and "PV-1$^{low}$" (%PV-1+/CD31+ cells $\leq$65) according to the best cutoff value of PV-1 identified by the Youden index (Luo and Xiong, 2013). An example of the scoring attribution is provided in Figure S1A. To evaluate the impact of increased PV-1 detection on patient clinical course, we assessed whether PV-1 was associated with progression-free survival (PFS), disease-free survival (DFS), and overall survival (OS) of CRC patients enrolled in the study (Figures 1D and S1B). Kaplan-Meier curves show that PV-1$^{high}$ was associated with poor prognosis and with a lower rate of 10-year PFS (40% versus 72% for PV-1$^{high}$ and PV-1$^{low}$, respectively $p < 0.0001$) (Figure 1D). These correlations remained highly significant in multivariate Cox proportional hazard models adjusted for significant prognostic and confounding variables (age, lymph node [LN] involvement, perineural invasion, and neutrophil-lymphocyte ratio). Indeed, we found that PV-1$^{high}$ patients had double the risk of relapse compared with PV-1$^{low}$ patients (PFS: hazard ratio [HR] = 2.34; 95% confidence interval [CI], 1.38–3.98; $p = 0.0017$; DFS: HR = 1.91; 95% CI, 1.18–3.11; $p = 0.009$) (Figure S1B). Unlike the PFS and DFS, no differences in the OS were detected (OS: HR = 1.28; 95% CI, 0.49–3.32; $p = 0.6125$) (Figure S1B). Except for perineural invasion, the univariate analyses indicated that PV-1 status is independent of all the clinical prognostic factors that we analyzed (Figure S2A). Furthermore, as it has been shown that in CRC the proportion of T cells inversely correlates with the prognosis (Fridman et al., 2012; Galon et al., 2006), we analyzed T cell infiltration in the primary tumor and correlated it with PV-1 detection. In line with the literature (Fridman et al., 2012; Galon et al., 2006), we observed that metastatic patients had less T cell infiltration in their primary tumor compared with non-recurrent patients (Figures S2B and S2C) and, unexpectedly, this phenomenon was inversely correlated with PV-1 detection (Figure S2D). Indeed, we thought that increased blood vessel permeability would allow cells and molecules to translocate both ways: in and out of the tumor tissue. One possible explanation could be that T cells are recruited into the tumor tissue by an active mechanism and/or a gradient of attracting chemokines, and not simply by a passive passage through a disrupted GVB.

Another important marker linked to poor CRC prognosis is the presence of LN metastases at the time of primary tumor diagnosis (Laubert et al., 2012). In line with current knowledge (Laubert et al., 2012), the univariate analyses of our training cohort confirmed that the presence of LN metastases was linked to relapse (60% versus 28% in LN positive (pN+) and LN negative (pN−), respectively, $p < 0.0001$) (Figures S2E and S1B). However, in our training cohort of CRC patients, increased PV-1 did not correlate with the number of metastatic regional LNs at the time of surgery (Figure S2A). This suggests that the number of metastatic regional LNs and the frequency of PV-1+ cells are two independent markers indicating the presence of two different paths for cancer cell dissemination: the vascular and the lymphatic routes. We assessed whether the combination of PV-1 positivity and LN involvement could improve the prediction of CRC recurrence, by evaluating the PFS, DFS, and OS. CRC patients usually relapse within 5 years (Pita-Fernández et al., 2015). Seventy-five percent of patients with PV-1$^{high/pN+}$ had a 2-fold higher risk of relapse compared with 42% of PV-1$^{low/pN+}$ patients within follow-up (HR = 2.15; 95% CI, 1.16–3.97; $p = 0.015$) (Figures 1E and S3A). Furthermore, we observed that 42% of patients with PV-1$^{high/pN-}$ developed distant metastases, which is comparable with that of PV-1$^{low/pN+}$ patients (42%) and

(C, D, and G) (C) Percentage of PV-1+ cells among CD31+ cells in CRC patients with (n = 100) or without metachronous distant metastases (n = 79). Kaplan-Meier curves of progression-free survival (PFS) according to PV-1 detection in (D) training cohort and (G) validation cohort.

(E) Kaplan-Meier curves of PFS according to PV-1 detection and LN involvement.

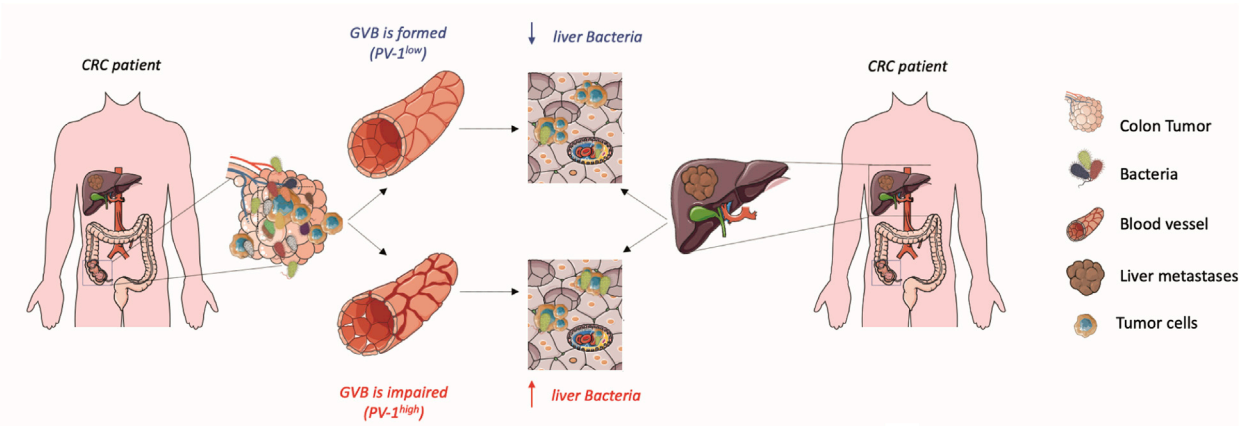
(F) Percentage of PV-1+ cells among CD31+ cells in CRC patients with (n = 24) or without distant metastases (n = 27). (A, C, and F) Data are represented using box and whisker plots. Boxplots display values of minimum, first quartile, median, third quartile, and maximum. Each data point represents one sample. Statistical significance was evaluated using two-sided Mann-Whitney unpaired test. *$p < 0.05$, ****$p < 0.0001$.

See also Figures S1–S3 and Tables S1 and S2.

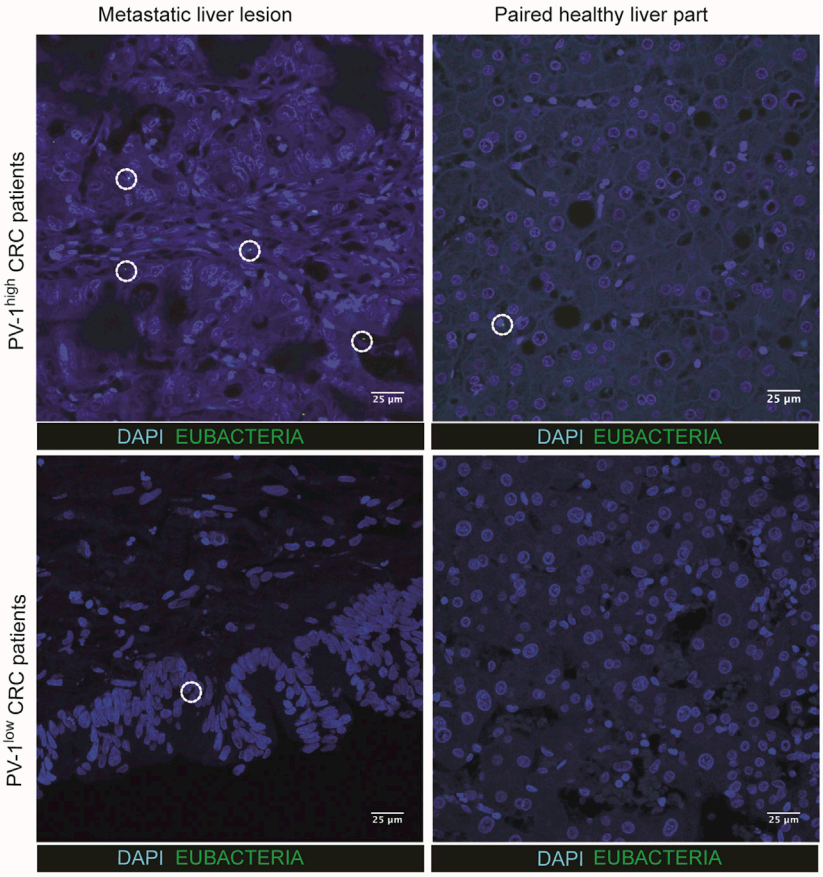
Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004



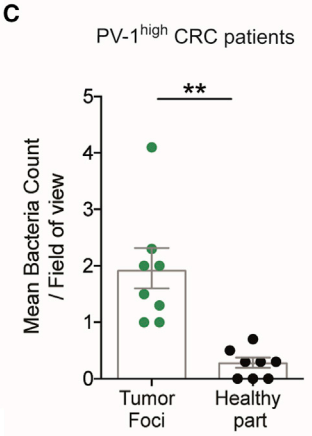
A



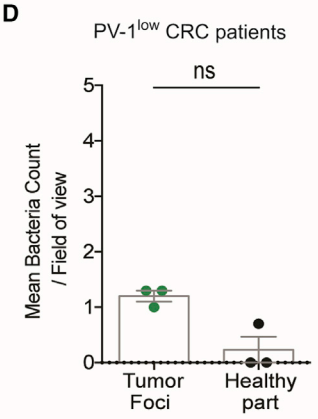

B




C




D




E

| | Liver metastases | Lung metastases | Not lung/liver metastases | Total |
|---|---|---|---|---|
| PV-1$^{high}$ (PV-1>65) - number (%) | 28 (76) | 17 (63) | 8 (62) | 53 (69) |
| PV-1$^{low}$ (PV-1≤65) - number (%) | 9 (24) | 10 (37) | 5 (38) | 24 (31) |
| All | 37 (48) | 27 (35) | 13 (17) | 77 |
| ** for two metastatic patients with PV-1 $^{high}$ we don't know the site of metastases* | | | | |

*(legend on next page)*

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004



much higher than PV-1$^{low/pN-}$ patients (14%, Figure 1E). We found similar associations with DFS (HR = 1.93; 95% CI, 1.08–3.45; p = 0.026) but not with OS (HR = 1.53; 95% CI, 0.52–4.57; p = 0.443) (Figure S3A). These data suggest that predicting tumor recurrence using these two markers in combination is more efficient than analyzing PV-1 and LN involvement independently, and that PV-1$^{high/pN-}$ patients should be considered at high risk similarly to PV-1$^{low/pN+}$ patients.

## Confirmation of PV-1 as a CRC prognostic biomarker in a validation cohort

These results prompted us to validate our findings in an independent cohort of 51 CRC patients coming from another two hospitals (Table S2). Consistent with the training cohort, we observed a higher frequency of PV-1+ cells in the primary tumors of metachronous CRC patients as compared with CRC patients who did not develop distant metastases in the follow-up time (Figure 1F). Kaplan-Meier curves showed that PV-1 $^{high}$ CRC patients were associated with poor prognosis and with a lower rate of 4-year PFS (p = 0.003; Figure 1G). These correlations remained significant in multivariate Cox proportional hazard models adjusted for significant prognostic and confounding variables (LN involvement and neutrophil-lymphocyte ratio) (Figure S3B). In line with the training cohort, we found that PV-1$^{high}$ patients had a 2-fold higher risk of relapse compared with PV-1$^{low}$ (PFS: HR = 2.94; 95% CI, 1.20–7.23; p = 0.02) (Figure S3B). Taken together, these data indicate that PV-1 is a prognostic biomarker of distant recurrence for CRC via the hematic route.

## CRC patients with PV-1$^{high}$ cells in the primary tumor have more bacteria in their metastatic liver lesions

We previously demonstrated that a disrupted GVB results in bacterial systemic dissemination and liver colonization (Spadoni et al., 2015). As it has been shown that cancer-associated commensal bacteria, such as *Fusobacterium nucleatum*, can colonize both primary colon tumor and the corresponding distant metastasis (Bullman et al., 2017), we hypothesized that bacteria migration could be secondary to GVB disruption (Figure 2A). We performed fluorescence *in situ* 16S rRNA-hybridized (FISH) microscopy to visualize bacteria in metastatic livers of PV-1$^{high}$ and PV-1$^{low}$ metachronous CRC patients (Figure 2B). We observed that only PV-1$^{high}$ but not PV-1$^{low}$ CRC patients had a statistically higher number of bacteria in the liver metastatic lesions than in paired healthy hepatic parenchyma, suggesting that bacteria translocate mainly when PV-1 is high (Figures 2B–2D). Unfortunately, we had the possibility to analyze the metastatic liver from only three PV-1$^{low}$ CRC patients. This was due to the low number, only 24% (n = 9), of PV-1$^{low}$ CRC patients having developed liver metastases and the limited availability of liver specimens (only three of nine, Figure 2E). Indeed, PV-1$^{high}$ CRC patients metastasize primarily to the liver (76%) than the lung (Figure 2E), whereas CRC patients with PV-1$^{low}$ metastasize equally to both organs (Figure 2E). These data indicate that PV-1 positivity in primary colon tumors correlates with increased bacteria detection in liver metastatic foci.

## Bacteria localize near SOX9+ highly proliferative cancer cells in metastatic livers

To evaluate where bacteria localize within the metastatic lesions, we used imaging mass cytometry (Hyperion) (Figure 3A). This innovative technique allowed us to simultaneously analyze seven different markers and couple it to FISH staining of bacteria on serial sections (Figures 3A and 3B). We found that highly proliferative cells (Ki67+) and SOX9+ cancer cells were surrounded by a remodeled extracellular matrix stained with $\alpha$-Sma and collagen type III-positive cells (Figure 3A). Liver lesions were enclosed by a vascular network of CD31+ blood vessels, and bacteria were located within areas of SOX9+ cancer cells in CRC metastatic regions (Figures 3A and 3B). To corroborate this finding, we evaluated the colocalization of bacteria stained by 16S rRNA-FISH and SOX9+ cancer cells in other metastatic livers and primary colon tumors. Bacteria colocalized with SOX9+ cancer cells both in metastatic lesions (Figures 3C and 3D) and in primary tumors (Figure 3E). Our observations are supported by a recent study showing that bacteria localize mainly in cancer cells in different tumor types (Nejman et al., 2020). Thus, besides extending previous results (Bullman et al., 2017), we demonstrate that bacteria are present in liver metastases of metachronous metastatic CRC patients with altered GVB. Whether the two events are causally and temporally linked is unclear.

## Bacteria disseminate to the liver and trigger a premetastatic niche formation

To investigate if bacteria were preceding metastases formation, we switched to a spontaneous mouse model of CRC, which we developed in our laboratory, that combines the Apc$^{Min/+}$ mutation with the deficiency in the C3a complement anaphylatoxin receptor (C3ar) (Apc$^{Min/+}$C3arKo) (Guglietta et al., 2016). Similar to human CRC, in Apc$^{Min/+}$C3arKO mice we observed increased PV-1 detection in colon tumors as compared with healthy colons of control mice (Figure 4A), and this occurred even before any sign of distant metastases. We previously showed that increased PV-1 detection highlights GVB disruption and increased blood vessel permeability (Mouries et al., 2019; Spadoni et al., 2015), thus we functionally assessed whether this was true also in the context of CRC. We performed a FITC dextran permeability assay in Apc$^{Min/+}$C3arKO, C3arKO, and wild-type (WT) mice, at age 25 weeks, a time point in which PV-1 becomes detectable. We challenged mice with 4 kDa FITC dextran by oral gavage

**Figure 2. PV-1$^{high}$ CRC patients show more bacteria in their metastatic liver lesions**

(A) Experimental outline.

(B) Confocal images show fluorescence *in situ* hybridization of prokaryotic 16S rRNA signal. Analysis of two representative human metastatic livers and paired healthy parenchyma (n = 11). Tissue sections were stained with DAPI (blue), eubacteria probes (green), and scramble probes (not shown). Bacteria are encircled by white dashed circles. Scale bars, 25 μm.

(C and D) Mean bacteria count inside metastatic lesions and paired healthy parenchyma of (C) PV-1$^{high}$ and (D) PV-1$^{low}$ CRC patients (n = 8 and 3, respectively). Each data point represents one sample. Data are represented as mean ± SEM using scatter dot plots. Statistical significance was evaluated using two-sided Wilcoxon matched-pairs signed rank test. ns p > 0.5, **p < 0.01.

(E) Percentage of PV-1$^{high}$ and PV-1$^{low}$ CRC patients who develop or not liver metastases.

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004



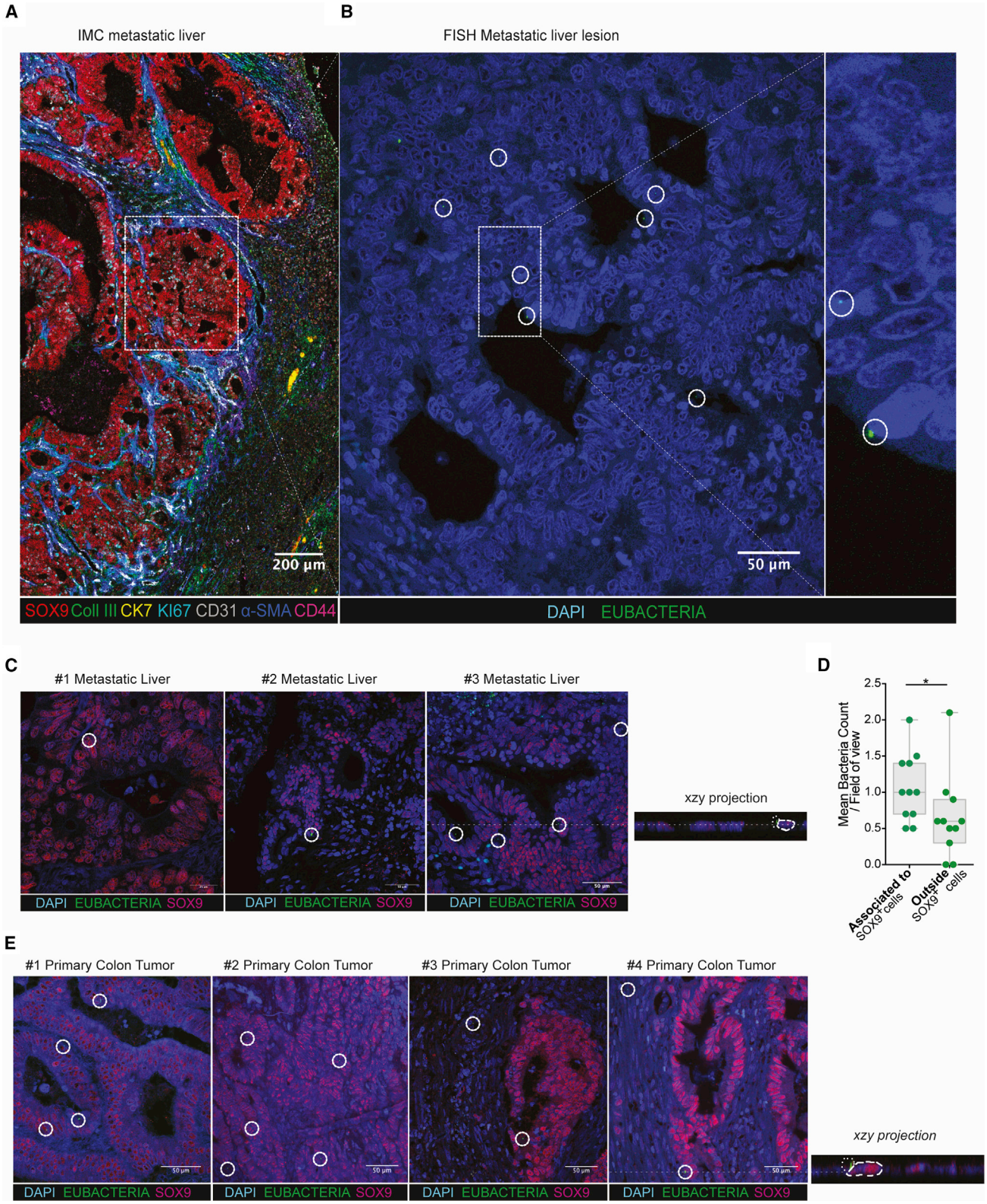



**Figure 3. Bacteria are localized near SOX9+ cancer cells both in the primary tumor and liver metastatic lesions**

(A) FFPE section of human metastatic liver was stained with seven different metal-tag antibodies. Imaging mass cytometry (IMC) shows the expression of SOX9, Coll III, CK7, Ki67, CD31, α-SMA, and CD44+ cells.

*(legend continued on next page)*

and, after 4 h, we quantified the amount of dextran translocated into the blood of these mice. As shown in Figures 4B and 4C, we detected significantly more dextran in the blood of tumor-bearing mice compared with control mice, which correlated with the presence of both intracellular aerobic and anaerobic bacteria in the livers of Apc$^{Min/+}$C3arKO mice. This suggests that increased PV-1 detection reflects increased blood permeability also in a tumor context, likely allowing tumor-associated bacteria to reach the liver. Primary tumor-derived factors can instruct distant target organs leading to the formation of the PMN (Peinado et al., 2017). Thus, we evaluated whether the presence of bacteria correlated and was responsible for PMN maturation in the liver of tumor-bearing mice. We treated mice at age 25 weeks when the primary tumors are already established for 1 week with a broad-spectrum antibiotic mix (ABX) (Figure 4D). The antibacterial efficacy of ABX was confirmed by measuring 16S rRNA in the fecal content (Figure 4E). Liver bacteria were also depleted by antibiotic treatment, confirming their bacterial nature and excluding a technical artifact (Figure 4C). We analyzed innate immune cell recruitment, as a proxy of PMN formation, in the livers of ABX-treated versus untreated mice and observed that gut and liver bacteria depletion directly impaired innate immune cell recruitment within the livers of Apc$^{Min/+}$C3arKO (Figures 4F–4G). In addition, neutrophils were mainly localized in the portal space, whereas macrophages were present inside sinusoids and hepatic parenchyma (Figure 4G). We then evaluated if inflammatory cytokines and chemokines, which are involved in PMN formation (Costa-Silva et al., 2015; Erler et al., 2009; Hiratsuka et al., 2006, 2008; Hoshino et al., 2015; Qian et al., 2011; Seubert et al., 2015), were also modulated. Our results show that ABX treatment caused a significant reduction of *Tnf-α*, *Tgf-β*, and *Ccl-2* gene expression in the liver of Apc$^{Min/+}$C3arKO mice (Figure 4H). Differently from Apc$^{Min/+}$C3arKO mice, the number of bacteria (Figures S4A and S4C) and infiltrating immune cells (Figures S4B and S4D) was extremely low in the livers of control mice. Inflammatory cytokine and chemokine gene expression was also very low and differently affected by ABX treatment in control mice (Figures S4E and S4F). Of note, we observed an opposite regulation of *Ccl-2* gene expression in the C3arKO mice upon antibiotic treatment compared with tumor-bearing mice (Figure S4E). We then deeply characterized the PMN by mRNA sequencing of RNA isolated from the livers of Apc$^{Min/+}$C3arKO mice at age 25 weeks, treated or not with antibiotics for 1 week. We identified 779 differentially expressed genes that were upregulated in the liver of tumor-bearing mice not treated with ABX compared with the treated ones (Figure 4I). Among these, we found genes coding for proteins acting specifically as chemoattractants for myeloid cells (i.e., macrophages and neutrophils), such as *Saa 1/2/3/4*, *Lcn2*, and *Tlr5* (Hiratsuka et al., 2008; Ichikawa et al., 2011; Lee et al., 2019; Vijay-Kumar et al., 2010), and extracellular matrix deposition (*Mmp15*, *Fgfr-l1*, and *Col15a1*) (Gonzalez-Avila et al., 2019; Lemoinne et al., 2015) (Figure 4J), which are two key components of the PMN. These results indicate that GVB disruption, in tumor-bearing mice, correlates with bacteria dissemination to the liver where they promote the recruitment of innate immune cells and the formation of an inflammatory environment consistent with a PMN before liver metastases establishment.

## A specific microbiota landscape correlated with GVB impairment in murine colon tumors

We observed that PV-1 was less detectable in colon tumors of Apc$^{Min/+}$C3arKO mice treated with antibiotics (Figure 5A), suggesting that tumor-resident bacteria may contribute to the deregulation of the GVB. We previously demonstrated that *Salmonella typhimurium* is capable of controlling PV-1 expression (Spadoni et al., 2015), thus we supposed that intratumoral bacteria might be involved in increased PV-1 detection. We compared the microbiome of primary colon tumors/healthy colons, fecal pellets, and liver tissues of Apc$^{Min/+}$C3arKO and control mice, via taxonomic profiling of 16S rRNA (Figure 5B). We calculated Chao1 alpha diversity index as an estimator of richness among different communities. Our data show that the liver microbiome across different genotypes was characterized by a richer community of bacteria as compared with colon tissues and fecal pellets (Figure S4H). Then, we applied principal coordinate analysis to Bray-Curtis distances to evaluate dissimilarities among experimental groups. The bacterial communities in the liver were separated from those of the colon and feces in the PC1 in WT mice as compared with tumor-bearing mice (Figure S4G). This suggested that the liver and colon microbiota of tumor-bearing mice were more similar than those of control mice. Similarly, at the phylum level we observed that the microbiome was qualitatively similar between the three anatomical sites in Apc$^{Min/+}$C3arKO mice as compared with control mice (Figure 5C). At species level, we observed a gradient of *Escherichia coli* representation among the detected bacterial species both in the colon and in paired liver tissues of Apc$^{Min/+}$C3arKO followed by C3arKO and finally WT mice (Figures 5D and S4I). These data show that the colon and liver microbiome of tumor-bearing mice with disrupted GVB are similar and mainly colonized by *E. coli*, suggesting that there is a spatial connection between these two sites.

## *E. coli* C17 triggers GVB impairment and reaches the liver fostering PMN maturation

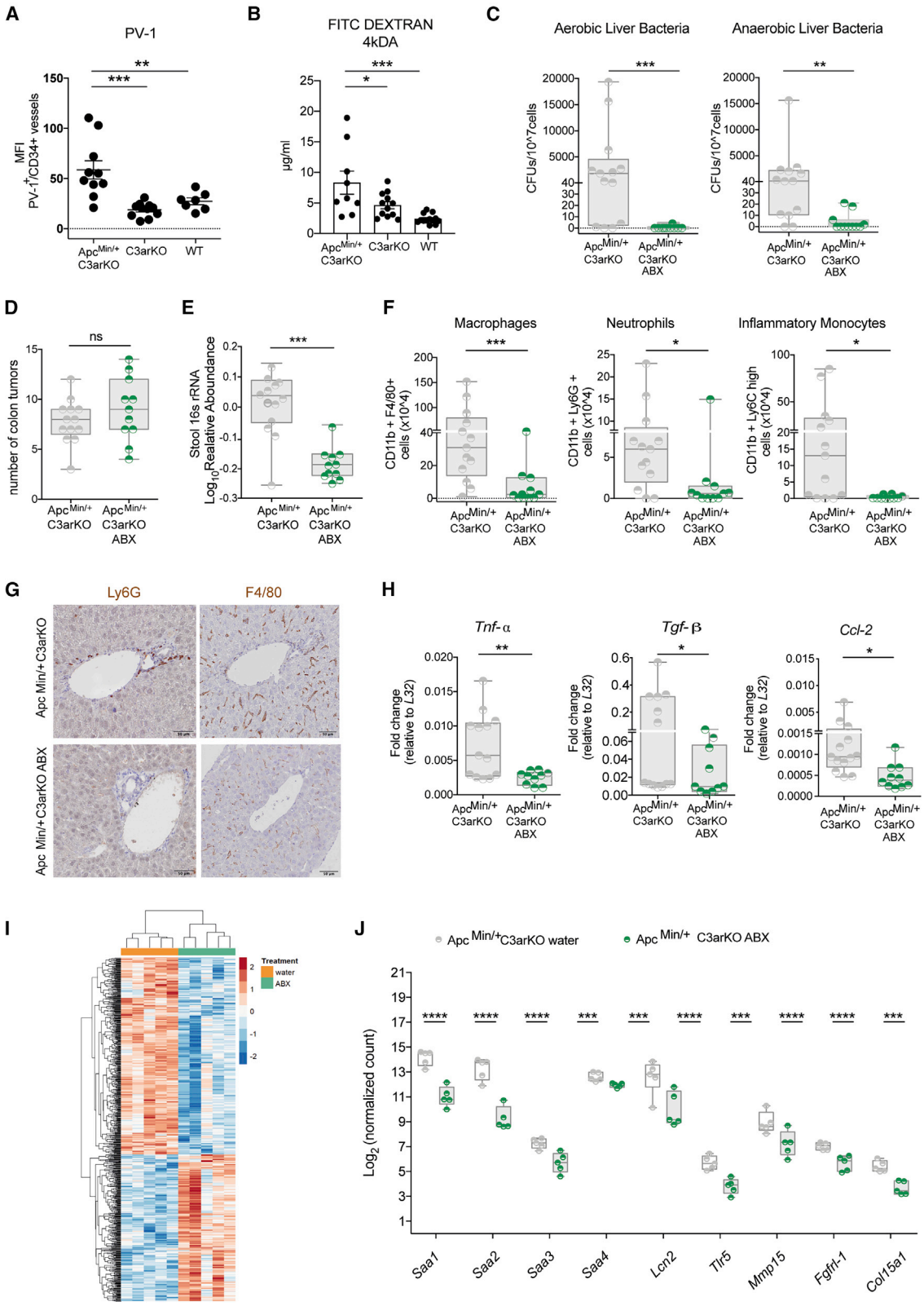
Next, we isolated bacterial strains from colon and liver tissues of Apc$^{Min/+}$C3arKO mice to identify tumor-resident bacteria actively involved in controlling GVB permeability and bacterial translocation to the liver. As expected, the most frequent isolates were strains of *E. coli*. We carried out a whole-genome

(B, C and E) (B) Confocal images show FISH analysis performed on a serial section of the liver analyzed for IMC. Confocal images show bacteria (FISH staining) and SOX9+ cancer cells in (C) liver metastases and (E) primary tumors. Analysis of three representative human metastatic livers (C, n = 11) and (E) four CRC. All tissue sections were stained with DAPI (blue), eubacteria probes (green), and SOX9 (red). (B, C, and E) Bacteria are encircled by white dashed circles. Scale bars, 200 μm (A),and 50 μm (B, C, and E).

(D) Mean bacteria count associated or not to SOX9+ cancer cells inside metastatic lesions (n = 11). Data are represented using box and whisker plots. Boxplots display values of minimum, first quartile, median, third quartile, and maximum. Statistical analysis was evaluated using two-sided Wilcoxon matched-pairs signed rank test, *p < 0.05. Each data point represents one sample.

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004






*(legend on next page)*

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004



sequencing of four *E. coli* strains isolated from two primary colon tumors and found that these strains were almost identical (99.9% average nucleotide identity) (Figure S5A). We chose one of these isolates (*E. coli* C17) to assess its ability to directly influence PV-1 detectability. C57BL/6 mice were orally gavaged with *E. coli* C17 ($1 \times 10^7$ colony-forming units [CFUs]) every other day for 1 week (Figure 5E). We observed increased PV-1 detection on colonic blood vessels 24 h after the last *E. coli* C17 administration (Figures 5F and 5G) and increased bacteria dissemination to the liver (Figure S6A). In addition, the livers of treated mice were preferentially infiltrated by macrophages and inflammatory monocytes (Figure S6B), two main components of the PMN (Costa-Silva et al., 2015; Qian et al., 2011). To evaluate where this bacterium localized *in vivo* we labeled live *E. coli* C17 with a plasmid harboring a reporter gene expressing the red fluorescent protein mCherry under a constitutive promoter and orally challenged C57BL/6 mice with *E. coli* C17 mCherry ($1 \times 10^7$ CFUs) every other day for 1 week. We tracked *E. coli* C17 in the liver 24 h after the last *E. coli* C17 mCherry administration and observed that it mainly localized inside macrophages and at a lower extent in other non-immune cells, likely hepatocytes (Figures 5H and S6C). Together, these data indicate that *E. coli* C17 colonizes the tumor tissue, disrupts the GVB, and migrates to the liver. We then compared the ability of *E. coli* C17 versus that of a beneficial bacterium, such as *Lactobacillus paracasei CNCM I-5220* (Roggero et al., 2020; Zagato et al., 2014), to ameliorate/worsen GVB abnormalities. $Apc^{Min/+}$C3arKO mice were treated or not every other day for 3 weeks with either *L. paracasei CNM I-5220* or *E. coli* C17 (Figure 5I). We observed that *L. paracasei CNCM I-5220* treatment reduced PV-1 detection to a level significantly lower than that observed in untreated $Apc^{Min/+}$C3arKO tumor-bearing mice, by contrast *E. coli* C17 boosted even further PV-1 upregulation (Figure 5J). While we did not observe an increase in liver CFUs, we found that *E. coli* C17 drastically fostered the recruitment of macrophages, neutrophils, and inflammatory monocytes as compared with $Apc^{Min/+}$C3arKO-untreated mice (Figures S6D and S6E). In contrast, *L. paracasei CNCM I-5220* had a beneficial role also on the liver microenvironment, by reducing the seeding of innate immune cells (Figure S6E). These data suggest that specific bacteria are actively involved in modulating GVB permeability and controlling leukocyte recruitment in the liver.

We previously showed that *Salmonella* uses the pathogenicity island 2 (Spi2), which encodes for the type III secretion system (TTSS) machinery (Ohl and Miller, 2001; Spadoni et al., 2015) to cause GVB impairment (Spadoni et al., 2015). We hypothesized that a similar mechanism may be used by *E. coli* C17, but this may be peculiar to C17, as a commensal strain of *E. coli* was unable to cross the GVB, disseminate systemically and colonize the liver (Spadoni et al., 2015). Thus, we compared the genome of *E. coli* C17 to an internal database of 32 *E. coli* strains. Among the genes unique to *E. coli* C17, we found *Virf1* and *2* that encode for proteins specifically involved in the formation of the TTSS machinery (Table S3) (Koppolu et al., 2013; Di Martino et al., 2016). To assess whether *Virf* was responsible for conferring GVB-disrupting properties to *E. coli* C17 we proceeded in generating knockout strains. While we were unable to delete *Virf2*, we succeeded in knocking out the *Virf1* gene. *E. coli* C17 Virf1:cat mutant derivative was constructed through homologous recombination using the λ-red technique (Datsenko and Wanner, 2000). PCR products for mutant construction were obtained by amplifying the resistance cassette flanked by the FRT sequences from the genome of the *E. coli* MG1655ΔcysH strain (Rossi et al., 2014) (Figure S6F; Table S5). To assess the contribution of *Virf1* in the GVB-disrupting properties of *E. coli* C17 *in vivo*, C57BL/6 mice were orally gavaged with either WT or the *Virf1* mutant strain (*E. coli* C17 Δ*virf1*) ($1 \times 10^7$ CFUs) for 2 consecutive days (Figure 6A). Even though the two strains colonized the gut similarly (Figure S6G), *E. coli* C17 Δ*virf1* mutant reduced PV-1 upregulation in the colon (Figures 6B and 6C). Hence, *E. coli* C17 uses a TTSS *Virf1*-mediated mechanism to dismantle the GVB. Since we were unable to inactivate *Virf2*, we cannot not exclude that both *Virf1* and *Virf2* are involved in this process. It is also possible that a mutant lacking both *Virf1* and *2* may be even more defective in controlling GVB disruption.

Finally, we addressed the relevance of *E. coli* C17 in human CRC by evaluating whether it was detectable in our cohort of

**Figure 4. GVB disruption in tumor-bearing mice allows bacteria to colonize the liver supporting PMN maturation before the development of distant metastases**

(A) Mean fluorescence intensity (MFI) of PV-1 among CD34 intestinal blood vessels of $Apc^{Min/+}$C3arKO (n = 10), C3arKO (n = 11), and WT (n = 7) mice at 25 weeks of age. Data are represented as mean ± SEM using scatter dot plots.

(B) Serum level of 4kDa FITC dextran collected from $Apc^{Min/+}$C3arKO (n = 9), C3arKO (n = 12) and WT (n = 13) mice at age 25 weeks. Data of two pooled experiments are represented as mean ± SEM using a scatter dot plots. (A and B) Statistical analysis was evaluated using ordinary one-way ANOVA with Bonferroni post-test; *p < 0.05, **p < 0.01, ***p < 0.001.

(C) Colony-forming units (CFUs) of culturable intracellular aerobic and anaerobic bacteria in the liver of $Apc^{Min/+}$C3arKO mice treated with (n = 11) or without (w/o) ABX (n = 13).

(D) Number of colon tumors in $Apc^{Min/+}$C3arKO treated with (n = 11) or w/o ABX (n = 13).

(E) Bacterial load quantification using real-time qPCR to detect 16s rRNA gene in the feces of $Apc^{Min/+}$C3arKO mice treated with (n = 11) or w/o ABX (n = 13).

(F) Flow cytometric analysis of CD11b+ F4/80+ (macrophages-like), CD11b+ Ly6G+ (neutrophils), and CD11b+ Ly6c+ (inflammatory monocytes) cells in the liver of $Apc^{Min/+}$C3arKO mice treated with (n = 11) or w/o ABX (n = 13).

(G) Images of Ly6G+ and F4/80+ cells in the livers of $Apc^{Min/+}$C3arKO mice treated with (bottom) or w/o ABX (top).

(H) mRNA expressions of *Tnf-α*, *Tgf-β*, and *Ccl-2* normalized to L32 gene transcription, in livers of $Apc^{Min/+}$C3arKO mice treated with (n = 10) or w/o ABX (n = 13).

(C–F, and H) Data of three pooled experiments are represented using box and whisker plots. Boxplots display values of minimum, first quartile, median, third quartile, and maximum.

(I) Heatmap showing differentially expressed genes in the liver of $Apc^{Min/+}$C3arKO mice treated with (n = 5) or w/o ABX (n = 5).

(J) Normalized count values for specific genes in the liver of $Apc^{Min/+}$C3arKO mice treated with (n = 5) or w/o ABX (n = 5). Data are represented using box and whisker plots. Boxplots display values of minimum, first quartile, median, third quartile, and maximum. Statistical analysis was evaluated using (C–F, and H) two-sided Mann-Whitney test or (J) Kruskal-Wallis Dunn's multiple comparison test; ns p > 0.05, *p < 0.05, **p < 0.01, ***p < 0.001. Each data point represents one mouse.

See also Figure S4.

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004



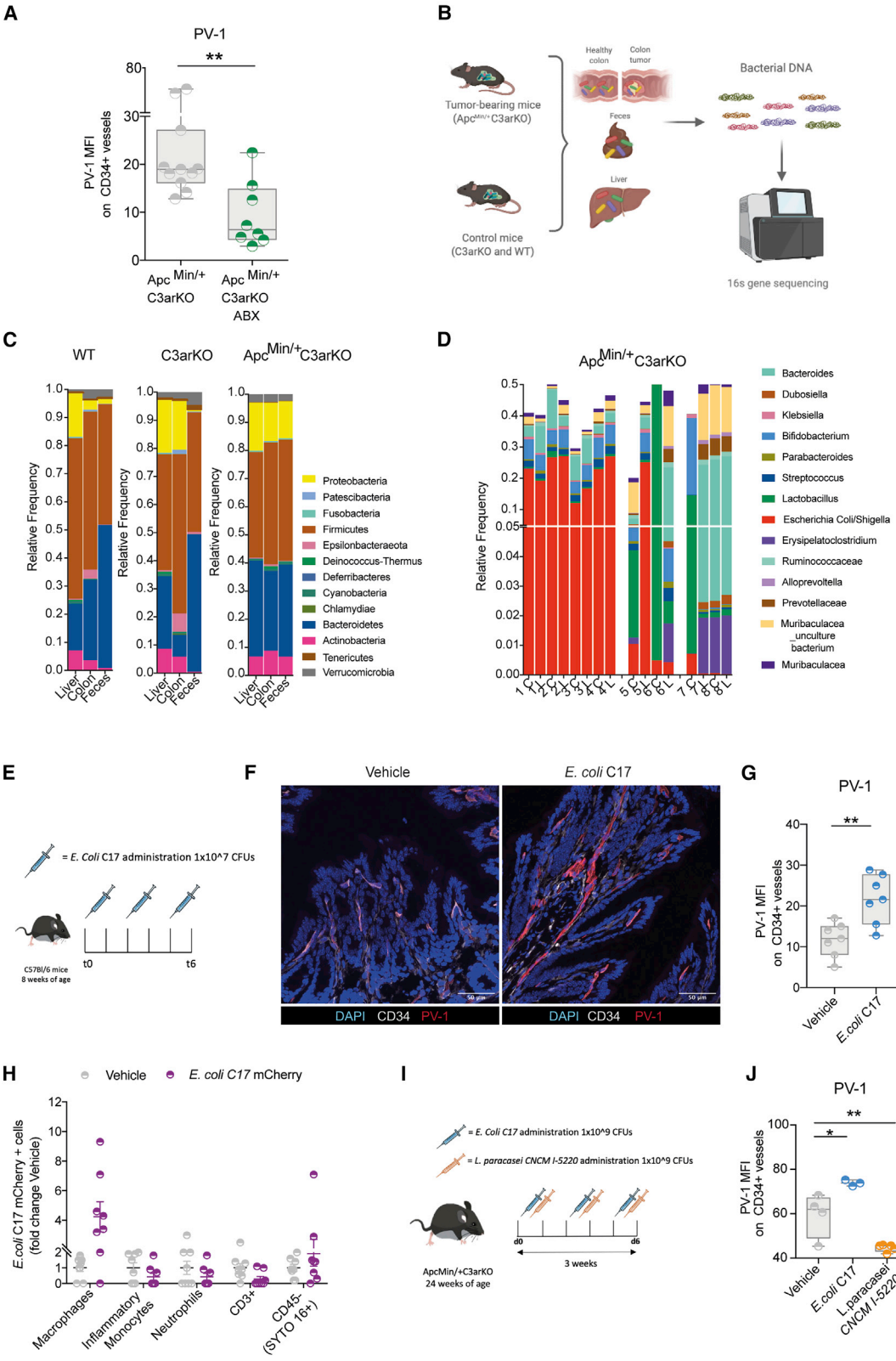



**Figure 5. Specific tumor-resident bacteria (such as *E. coli* C17) are directly causing GVB disruption**

(A) Mean fluorescence intensity (MFI) of PV-1 among CD34 intestinal blood vessels of Apc$^{Min/+}$C3arKO mice treated with (n = 8) or w/o ABX (n = 11).

(B) Scheme of 16s rRNA sequencing.

*(legend continued on next page)*

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004



patients. Thus, we extracted bacterial DNA from FFPE slices coming from primary tumor of PV-1$^{high}$ metastatic CRC patients (n = 10) and PV-1$^{low}$ not-metastatic CRC patients (n = 10). We observed that it was statistically more abundant in metastatic (PV-1$^{high}$) than in non-metastatic (PV-1$^{low}$) CRC patients, and it was also detectable in liver metastatic foci of CRC patients (Figures 6D and 6E). These data suggest that specific tumor-derived bacteria can trigger GVB disruption through the activation of a TTSS, allowing bacteria to colonize the liver and foster a PMN, triggering metastases formation.

## Spontaneous liver lesions in old tumor-bearing mice with altered GVB

To confirm the temporal relationship between GVB disruption, liver PMN onset and metastases formation, we searched for spontaneous liver lesions in aged Apc$^{Min/+}$C3arKO mice. A 35-week-old Apc$^{Min/+}$C3arKO mouse, with altered GVB, had macroscopic liver lesions (Figures 6F and 6G). The hematoxylin and eosin-stained sections of these lesions showed the presence of aberrant structures surrounded by fibrotic tissue (Figure 6G) positive for cytokeratin 20 (Figure 6H), a marker of intestinal epithelial cells currently used in the clinic to detect human CRC liver metastases (Tot, 2002). In line with the above results, we could also detect bacteria in this pathologic liver (Figure 6I). Unfortunately, because of primary tumor lethality (only 1.4% of mice survive longer than age 35 weeks, when liver metastases start developing) it has been impossible to perform a statistical analysis of this phenomenon or to study the role of bacteria in the metastatic process in this mouse model.

## Antibiotic treatment hampers metastases formation preventing GVB disruption

To overcome the limitations of Apc$^{Min/+}$C3arKO mice, we used another well-established model of liver metastases (Soares et al., 2014), which consists of the intrasplenic injection of murine MC38 pLUC colon cancer cells. To evaluate the contribution of intratumoral bacteria in this process, we intrasplenically injected MC38 pLUC colon cancer cells into the chemically induced AOM/DSS model of CRC. First, we demonstrated that the AOM/DSS model recapitulated the features of the Apc$^{Min/+}$C3arKO mice. Indeed, we observed an increased detection of PV-1 coupled with a higher presence of *E. coli* C17 in colons of tumor-bearing mice (Figures 7A–7C). Second, we evaluated the difference of MC38 pLUC colon cancer cells to reach the liver following intrasplenic injection in mice with or without intestinal tumors (Figure 7D). Bioluminescent *ex vivo* analyses showed more metastatic foci in livers of tumor-bearing mice compared with control mice (Figure 7E). Notably, this difference was not simply related to an increase in intestinal inflammation as DSS-treated mice displayed a number of liver lesions comparable with that of untreated mice and much lower than that of tumor-bearing mice (Figure 7E).

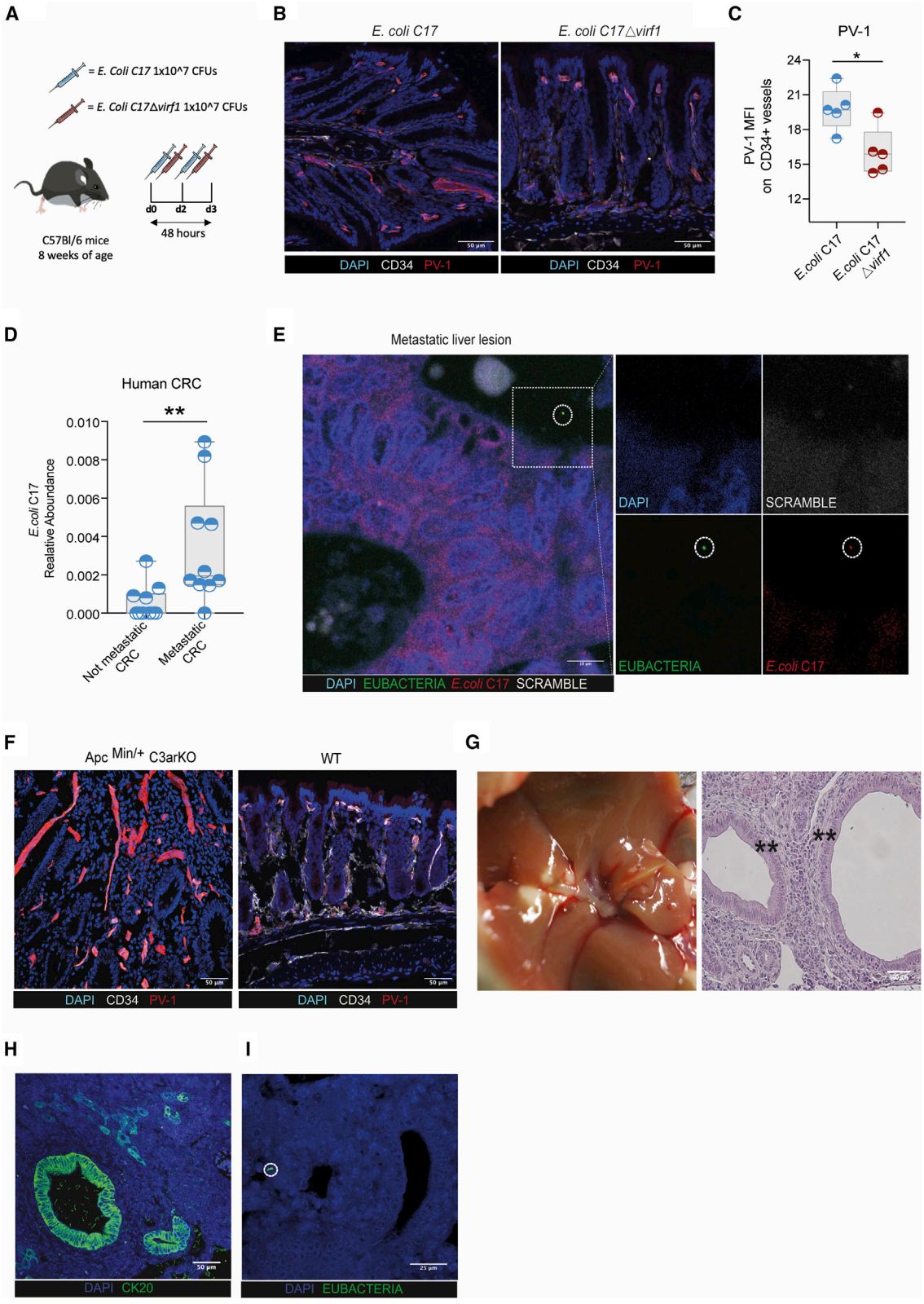
To directly assess the effect of microbiota on liver metastases onset, we treated AOM/DSS mice with ABX starting from 1 week before the intrasplenic injection (Figure 7D). Similar to Apc$^{Min/+}$C3arKO mice, ABX mix administration after tumor development had no impact on primary colon tumor numbers, but it decreased bacteria both in the liver and in the gut lumen (Figures 7F, S7A, and S7B). Together with reduced PV-1 detection in the primary tumor (Figure S7C), ABX-treated AOM/DSS mice showed significant reduction of metastatic foci compared with AOM/DSS untreated mice (Figure 7H). This was probably due to the defective formation of a hepatic PMN. Indeed, we observed that the liver of ABX-treated AOM/DSS mice showed reduced recruitment of innate immune cells and lower mRNA expression of *Tgf-β*, an essential protein for PMN formation (Peinado et al., 2017) (Figures 7I and S7D). Next, we evaluated the difference in the seeding capacity of MC38 pLUC colon cancer cells to the liver following intrasplenic injection in C57BL/6 mice pre-treated or not every other day for 6 days and for the following month with *E. coli* C17 (Figure S7E). After 30 days from intrasplenic injection, mice were sacrificed and livers were harvested for *ex vivo* quantification of metastases accumulation. Bioluminescent *ex vivo* analysis and macroscopic count showed more metastatic foci in livers of mice treated with *E. coli* C17 compared with control mice (Figures S7F and S7G).

We finally evaluated whether increased PV-1 detection per se was sufficient to drive PMN formation. Thus, we used a different model in which PV-1 is detectable in the absence of concomitant tumor formation (Figure S7H). In line with previous data from our laboratory (Mouries et al., 2019), 1 week of high-fat diet feeding drove an increase in PV-1 detection and translocation of bacteria to the liver as compared with standard diet-fed mice (Figures S7I–S7L). Nevertheless, we did not observe the development of the PMN in the liver (Figures S7M and S7N). These data suggest that PV-1 increased detection per se is not sufficient to drive

(C) Bar plots of the phylum-level microbial composition of paired colon, feces, and liver tissues from WT (n = 6), C3arKO (n = 6), and Apc$^{Min/+}$C3arKO (n = 8) mice at age 25 weeks. Relative frequencies are plotted as average of each sample type among the three different genotypes.
(D) Species level microbial composition of paired colon and liver tissues from Apc$^{Min/+}$C3arKO mice. For simplicity, we have represented the relative frequency of bacteria with statistically significant different balance between colon and liver of tumor-bearing mice and control mice (likelihood ratio test $p < 0.05$). The colors correspond to bacterial taxonomic class (n = 8 pairs, each pair represent one mouse).
(E) Scheme of *E. coli* C17 administration to C57BL/6 mice.
(F and G) (F) Confocal images and (G) relative MFI quantification of PV-1 (red) among CD34 intestinal blood vessels (gray) of C57BL/76 mice treated with *E. coli* C17 (n = 7) or vehicle (n = 7). Cell nuclei were stained with DAPI (blue). Scale bars, 50 μm.
(H) Flow cytometry analyses of *E. coli* mCherry localization in the liver of mice treated with *E. coli* C17mCherry (n = 8) or vehicle (n = 8). Data of one experiment are represented as mean ± SEM using scatter dot plots.
(I) Scheme of *E. coli* C17 and *L. paracasei CNCM I-5220* administration to Apc$^{Min/+}$C3arKO mice.
(J) MFI of PV-1 among CD34 intestinal blood vessels inside colon tumor of Apc$^{Min/+}$C3arKO mice treated with *E. coli* C17 (n = 3), *L. paracasei CNCM I-5220* (n = 5), or vehicle (n = 4). (A, G, and L) Data of one experiment are represented using box and whisker plots. Boxplots display values of minimum, first quartile, median, third quartile, and maximum. Statistical analyses were evaluated using (A–G) two-sided unpaired t test with Welch's correction or (J) ordinary one-way ANOVA with Bonferroni post-test two-sided; *$p < 0.05$, **$p < 0.01$.
See also Figures S5 and S6 and Table S3.

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004






(legend on next page)

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004



PMN maturation, but it has to be coupled with specific bacteria translocated into the liver, such as *E. coli* C17.

## DISCUSSION

Spreading of neoplastic cells to regional LNs correlates with poor CRC patient outcome and it is used as a prognostic marker of recurrence for treatment assignment (Laubert et al., 2012). However, it has been shown that distinct cancer cell clones from the primary tumor are able to colonize distant organs, presumably via the systemic blood circulation, suggesting that distant metastases could arise independently of LN metastases (Naxerova et al., 2017; Zhang et al., 2020). Venous blood from the intestine reaches the liver directly through the portal vein; therefore, it is the most logical CRC tumor cell-disseminating hematogenous route. Nonetheless, so far, no clinical prognostic marker was available to stratify CRC patients based on tumor blood vessel status. Here, we showed that the higher frequency of PV-1+ cells, highlighting increased blood vessel permeability in the primary tumor of CRC patients correlates with the development of metachronous distant metastases. Our analysis shows not only that PV-1 can be used as a prognostic marker for distal metastases, but that scoring CRC patients using both PV-1 status and LN involvement could better predict distant recurrence than using the two parameters independently.

We show that liver metastases are the consequence of a sequential series of events. Bacteria enter the tumor tissue and modify the GVB, then they migrate to the liver and foster the formation of a PMN which creates the soil for subsequent cancer cell seeding. We also started to unravel both the bacterial players and the molecular mechanism involved in this process. We found that a strain of *E. coli* (C17) could directly open the GVB, through a TTSS virulence factor (Virf)-dependent mechanism, and translocate into the liver where it could initiate the recruitment of immune cells contributing to PMN maturation and favoring metastases formation. The same strain (*E. coli* C17) could be detected in human CRC (both primary tumor and liver metastatic foci), suggesting that a similar mechanism may be used in human CRC and confirming a report showing that primary colon cancer and paired liver metastases are colonized by identical bacteria (Bullman et al., 2017).

This mechanism of metastasis formation may apply to other tumors and organs. Indeed, Ma et al. (2018) observed reduced tumors in the liver but increased neoplastic foci in the lung after oral antibiotic treatment in several models, including an intrasplenic injection of B16 tumor cells, which naturally metastasize to the lungs. It is possible that oral administration of antibiotics, decreases the recruitment of tumor cells to the liver and shifts metastases to the lung. The lung microbiota may also be involved in favoring a PMN in the lung. In line with this observation, it has been demonstrated that the lungs of vancomycin/neomycin-aerosolized mice showed a significant reduction of melanoma B16 lung metastases, together with a decrease in lung bacterial load (Le Noci et al., 2018). It would be interesting to evaluate whether by providing antibiotics via aerosol would inhibit intrasplenically injected B16 cell seeding also to the lung as well as the liver.

In conclusion, we here propose a mechanism of CRC metastases formation: the tumor and/or the tumor-associated microbiota induce GVB disruption, which allows bacteria to reach the liver and attract inflammatory cells. The latter contribute to the formation of a PMN and seeding of tumor cells leading to metastases formation. We observed that there are beneficial bacteria, such as *L. paracasei*, which instead restore a correct GVB. Additional investigations are required to further dissect the contribution of these bacteria in the control of the metastatic process. The GVB permeability marker PV-1 is thus a promising prognostic marker that, used alone or in association with LN status, can predict the development of distant metastases in CRC patients and can direct a more appropriate treatment strategy for LN negative CRC patients.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODELS AND SUBJECT DETAILS
  - Human samples

**Figure 6. *E. coli* C17 foster PV1 detection via a Virf1-dependent mechanism**

(A) Scheme of *E. coli* C17 and *E. coli* C17 *Δvirf1* administration to C57BL/6 mice.

(B and C) (B) Confocal images and (C) relative MFI quantification of PV-1 (red) increased detection among CD34 colonic vessels (gray) of C57BL/6 mice treated with *E. coli* C17 (n = 5) or *E. coli C17* Δ*virf1* (n = 5). Scale bars, 50 μm. Data of one experiment are represented using box and whisker plots. Boxplots display values of minimum, first quartile, median, third quartile, and maximum. Statistical analysis was evaluated using two-sided unpaired t test with Welch's correction. *$p < 0.05$, **$p < 0.01$.

(D) *E. coli* C17 relative abundance assessed by qPCR on bacterial DNA extracted from FFPE slices coming from primary tumor of not-metastatic (n = 10) and metastatic CRC patients (n = 10). Data are represented using box and whisker plots. Boxplots display values of minimum, first quartile, median, third quartile, and maximum. Statistical analysis was evaluated using two-sided Mann-Whitney test. **$p < 0.01$.

(E) FISH staining for *E. coli* C17 in one human metastatic liver. Tissue section was stained with eubacteria probes (green), *E. coli* C17 probes (red), and scramble probe (gray). Bacteria are encircled by white dashed circles. Scale bars, 10 μm.

(F) Confocal images show intestinal CD34+ blood vessels (gray) and PV-1 (red) staining from aged $Apc^{Min/+}$C3arKO colon tumor and healthy colon of age-matched WT mice.

(G) Pictures show hepatic lesions and the corresponding hematoxylin and eosin (H&E) staining. Double asterisks mark (**) aberrant structures. Scale bar, 100 μm.

(H) Confocal images of cytokeratin 20 (green).

(I) FISH staining of bacteria 16s rRNA (green, white circle) in the liver of $Apc^{Min/+}$C3arKO with aberrant liver structures. The sections were stained with DAPI (blue) for nuclei (B–I). Scale bars, 25 μm (I), 50 μm (F and H) and 100 μm (G).

See also Figure S6, Table S3.

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004



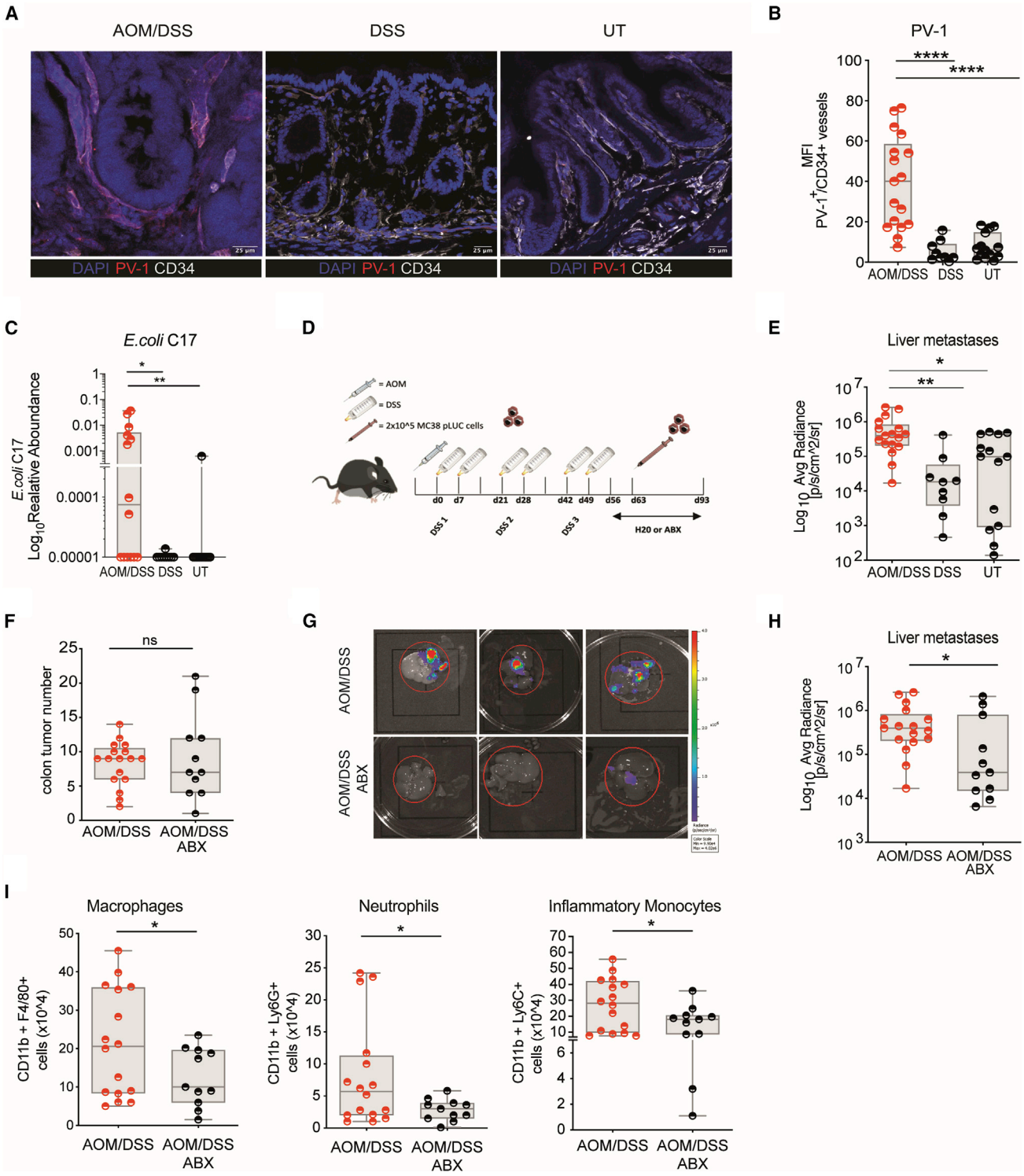



**Figure 7. Bacteria are the key factor able to boost distant metastases formation**

(A and B) (A) Confocal images and (B) relative MFI quantification of PV-1 (red) detection among CD34 (gray) colonic blood vessels of AOM/DSS (n = 14), DSS-treated (n = 9), and untreated (UT) (n = 14) mice at day 93 after intrasplenic injection. Cell nuclei were stained with DAPI (blue). Scale bars, 25 μm.

(C) *E. coli* C17 relative abundance in colon tissue of AOM/DSS (n = 14), DSS (n = 9), and UT (n = 13) mice, assessed by qPCR.

(D, E and H) (D) Scheme of the AOM/DSS experiment followed by intrasplenic injection with or without ABX treatment. Quantification of *ex vivo* signal related to MC38pLUC+ colon cancer cells accumulation in the liver of (E) AOM/DSS (n = 17), DSS-treated (n = 9), UT (n = 14), and (H) C57BL/6 AOM/DSS untreated versus ABX-treated intrasplenically injected mice. Multiple experiments (E–H) were performed simultaneously with a single control group (AOM/DSS).

*(legend continued on next page)*

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.ccell.2021.03.004.

## ACKNOWLEDGMENTS

We thank Dr. R. Spada, Dr. A. Doni, and Dr. F. Pasqualini for their help in imaging mass cytometry experiments, Dr. S. Freddi, Dr. M. Tillhon, and Dr. F. Montani for their assistance with PV-1 IHC, Dr. G. Fornasa and Dr. A. Melacarne for their help during experiments, Dr. C. Burrello for her precious scientific advice, the IEO and Humanitas Animal Facilities for the excellent animal husbandry. The research leading to these results has received funding from European Research Council Grant (M.R. HomeoGUT no. 615735), and from the Italian Foundation for Cancer Research (AIRC IG 22026; AIRC 5x1000 2018 ID.21147 – PI Alberto Mantovani, GL M.R.; AIRC 5x1000 2018ID.21091 program – PI Bardelli Alberto, GL M.R.; AIRC 5x1000 2019 ID.22757 – PI M.R.). Work in the P.P.D.F. lab was supported by the Italian Foundation for Cancer Research (IG 23060, MCO 10.000) and, in part, by the Italian Ministry of Health (Ricerca Corrente and 5x1000 funds). A.B. was supported by a FIRC-AIRC fellowship for Italy.

## AUTHOR CONTRIBUTIONS

A.B. ideated, conducted the experiments, and analyzed the data. S.C. contributed to experiments and data analysis. S.R. contributed to human CRC patient selection. G.B. performed PV-1 quantification. I.S., A.L., M.L., C.K., M.E., P.B., and A.A. provided technical support in *in vivo* experiments. S.G. performed the biostatistical analyses. D.B. performed bioinformatic analyses. F.A. and N.S. performed whole-genome shotgun sequencing of *E. coli* isolates. S.G. generated the $Apc^{Min/+}$C3arKO mice. M.G.J. and C.L. performed the IHC on human samples. E.R. generated *E. coli C17* Δ*Virf1*. A.S., A.M., and P.S.R. contributed to the selection of human CRC patients. S.P., P.P.D.F., G.V., and A.S. provided scientific discussion. C.P. and G.P. helped in designing experiments. M.R. conceived the study and interpreted the data. A.B. and M.R. wrote the manuscript.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 15, 2020
Revised: January 29, 2021
Accepted: March 8, 2021
Published: April 1, 2021

## REFERENCES

Asnicar, F., Thomas, A.M., Beghini, F., Mengoni, C., Manara, S., Manghi, P., Zhu, Q., Bolzan, M., Cumbo, F., May, U., et al. (2020). Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. Nat. Commun. https://doi.org/10.1038/s41467-020-16366-7.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. https://doi.org/10.1093/bioinformatics/btu170.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. https://doi.org/10.1038/s41587-019-0209-9.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. https://doi.org/10.3322/caac.21492.

Bullman, S., Pedamallu, C., Sicinska, E., Clancy, T.E., Zhang, X., Cai, D., Neuberg, D., Huang, K., Guevara, F., Nelson, T., et al. (2017). Analysis of *Fusobacterium* persistence and antibiotic response in colorectal cancer. Science *358*, 1443–1448, https://doi.org/10.1126/science.aal5240.

Costa-Silva, B., Ocean, A.J., Singh, S., Zhang, H., Thakur, B.K., Becker, A.A., Hoshino, A., Mark, M., Molina, H., Xiang, J., et al. (2015). Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver. Nat. Cell Biol. https://doi.org/10.1038/ncb3169.

Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. U S A. https://doi.org/10.1073/pnas.120163297.

Erler, J.T., Bennewith, K.L., Cox, T.R., Lang, G., Bird, D., Koong, A., Le, Q.T., and Giaccia, A. (2009). Hypoxia-induced lysyl oxidase is a critical mediator of bone marrow cell recruitment to form the premetastatic niche. Cancer Cell. https://doi.org/10.1016/j.ccr.2008.11.012.

Fridman, W.H., Pagès, F., Sauts-Fridman, C., and Galon, J. (2012). The immune contexture in human tumours: impact on clinical outcome. Nat. Rev. Cancer. https://doi.org/10.1038/nrc3245.

(F) Numbers of colon tumors at day 93 of AOM/DSS mice treated (n = 11) or not (n = 17) with ABX.
(G) Representative images of six different mice are shown. The color bar shows the radiant efficiency.
(I) Flow cytometric analysis of CD11b+ F4/80+ (macrophages-like), CD11b+ Ly6G+ (neutrophils), and CD11b+ Ly6c+ (inflammatory monocytes) cells in the liver of AOM/DSS mice treated (n = 11) or not (n = 17) with ABX. (B, C, E, F, H, and I) Data of three pooled experiments are represented using box and whisker plots. Boxplots display values of minimum, first quartile, median, third quartile, and maximum. Statistical analysis was evaluated using (B) ordinary one-way ANOVA with Bonferroni post-test two-sided, (C and E) Kruskal-Wallis Dunn's multiple comparison test and (F, H, and I) Mann-Whitney unpaired t test; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{****}p < 0.0001$. Each data point represents one mouse.
See also Figure S7.

Furet, J.-P., Firmesse, O., Gourmelon, M., Bridonneau, C., Tap, J., Mondot, S., DorÃ©, J., and Corthier, G. (2009). Comparative assessment of human and farm animal faecal microbiota using real-time quantitative PCR. FEMS Microbiol. Ecol. https://doi.org/10.1111/j.1574-6941.2009.00671.x.

Galon, J., Costes, A., Sanchez-Cabo, F., Kirilovsky, A., Mlecnik, B., Lagorce-Pagès, C., Tosolini, M., Camus, M., Berger, A., Wind, P., et al. (2006). Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science *5795*, 1960–1964, https://doi.org/10.1126/science.1129139.

Garrett, W.S. (2015). Cancer and the microbiota. Science *348*, 80–86.

Gebhardt, M.J., Jacobson, R.K., and Shuman, H.A. (2017). Seeing red; the development of pON. mCherry, a broad-host range constitutive expression plasmid for Gram-negative bacteria. PLoS One. https://doi.org/10.1371/journal.pone.0173116.

Giesen, C., Wang, H.A.O., Schapiro, D., Zivanovic, N., Jacobs, A., Hattendorf, B., Schüffler, P.J., Grolimund, D., Buhmann, J.M., Brandt, S., et al. (2014). Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. Nat. Methods. https://doi.org/10.1038/nmeth.2869.

Gonzalez-Avila, G., Sommer, B., Mendoza-Posada, D.A., Ramos, C., Garcia-Hernandez, A.A., and Falfan-Valencia, R. (2019). Matrix metalloproteinases participation in the metastatic process and their diagnostic and therapeutic applications in cancer. Crit. Rev. Oncol. https://doi.org/10.1016/j.critrevonc.2019.02.010.

Guglietta, S., Chiavelli, A., Zagato, E., Krieg, C., Gandini, S., Ravenda, P.S., Bazolli, B., Lu, B., Penna, G., and Rescigno, M. (2016). Coagulation induced by C3aR-dependent NETosis drives protumorigenic neutrophils during small intestinal tumorigenesis. Nat. Commun. *7*, 11037, https://doi.org/10.1038/ncomms11037.

Hiratsuka, S., Watanabe, A., Aburatani, H., and Maru, Y. (2006). Tumour-mediated upregulation of chemoattractants and recruitment of myeloid cells predetermines lung metastasis. Nat. Cell Biol. https://doi.org/10.1038/ncb1507.

Hiratsuka, S., Watanabe, A., Sakurai, Y., Akashi-Takamura, S., Ishibashi, S., Miyake, K., Shibuya, M., Akira, S., Aburatani, H., and Maru, Y. (2008). The S100A8-serum amyloid A3-TLR4 paracrine cascade establishes a pre-metastatic phase. Nat. Cell Biol. https://doi.org/10.1038/ncb1794.

Hoshino, A., Costa-Silva, B., Shen, T.L., Rodrigues, G., Hashimoto, A., Tesic Mark, M., Molina, H., Kohsaka, S., Di Giannatale, A., Ceder, S., et al. (2015). Tumour exosome integrins determine organotropic metastasis. Nature. https://doi.org/10.1038/nature15756.

Humbles, A.A., Lu, B., Nilsson, C.A., Lilly, C., Israel, E., Fujiwara, Y., Gerard, N.P., and Gerard, C. (2000). A role for the C3a anaphylatoxin receptor in the effector phase of asthma. Nature *406*, 998–1001, https://doi.org/10.1038/35023175.

Ichikawa, M., Williams, R., Wang, L., Vogl, T., and Srikrishna, G. (2011). S100A8/A9 activate key genes and pathways in colon tumor progression. Mol. Cancer Res. https://doi.org/10.1158/1541-7786.MCR-10-0394.

Koppolu, V., Osaka, I., Skredenske, J.M., Kettle, B., Hefty, P.S., Li, J., and Egan, S.M. (2013). Small-molecule inhibitor of the *Shigella flexneri* master virulence regulator VirF. Infect. Immun. https://doi.org/10.1128/IAI.00919-13.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods. https://doi.org/10.1038/nmeth.1923.

Le Noci, V., Guglielmetti, S., Arioli, S., Camisaschi, C., Bianchi, F., Sommariva, M., Storti, C., Triulzi, T., Castelli, C., Balsari, A., et al. (2018). Modulation of pulmonary microbiota by antibiotic or probiotic aerosol therapy: a strategy to promote immunosurveillance against lung metastases. Cell Rep. https://doi.org/10.1016/j.celrep.2018.08.090.

Lee, J.W., Stone, M.L., Porrett, P.M., Thomas, S.K., Komar, C.A., Li, J.H., Delman, D., Graham, K., Gladney, W.L., Hua, X., et al. (2019). Hepatocytes direct the formation of a pro-metastatic niche in the liver. Nature. https://doi.org/10.1038/s41586-019-1004-y.

Lemoinne, S., Cadoret, A., Rautou, P.E., El Mourabit, H., Ratziu, V., Corpechot, C., Rey, C., Bosselut, N., Bardu, V., Wendum, D., et al. (2015). Portal myofibroblasts promote vascular remodeling underlying cirrhosis formation through the release of microparticles. Hepatology. https://doi.org/10.1002/hep.27318.

Letunic, I., and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. https://doi.org/10.1093/nar/gkz239.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550, https://doi.org/10.1186/s13059-014-0550-8.

Luo, J., and Xiong, C. (2013). Youden index and associated cut-points for three ordinal diagnostic groups. Commun. Stat. Simulation Comput. https://doi.org/10.1080/03610918.2012.661906.

Di Martino, M.L., Falconi, M., Micheli, G., Colonna, B., and Prosseda, G. (2016). The multifaceted activity of the VirF regulatory protein in the *Shigella* lifestyle. Front. Mol. Biosciences. https://doi.org/10.3389/fmolb.2016.00061.

Ma, C., Han, M., Heinrich, B., Fu, Q., Zhang, Q., Sandhu, M., Agdashian, D., Terabe, M., Berzofsky, J., Fako, V., et al. (2018). Gut microbiome-mediated bile acid metabolism regulates liver cancer via NKT cells. Science. https://doi.org/10.1126/science.aan5931.

McMurdie, P.J., and Holmes, S. (2014). PLOS computational Biology: waste not, want not: why rarefying microbiome data is inadmissible. PLoS Comput. Biol.

Milani, C., Hevia, A., Foroni, E., Duranti, S., Turroni, F., Lugli, G.A., Sanchez, B., Martin, R., Gueimonde, M., van Sindersen, D., et al. (2013). Assessing the fecal microbiota: an optimized ion torrent 16S rRNA gene-based analysis protocol. PLoS One. https://doi.org/10.1371/journal.pone.0068739.

Mouries, J., Brescia, P., Silvestri, A., Spadoni, I., Sorribas, M., Wiest, R., Mileti, E., Galbiati, M., Invernizzi, P., Adorini, L., et al. (2019). Microbiota-driven gut vascular barrier disruption is a prerequisite for non-alcoholic steatohepatitis development. J. Hepatol. https://doi.org/10.1016/j.jhep.2019.08.005.

Naxerova, K., Reiter, J.G., Brachtel, E., Lennerz, J.K., van de Wetering, M., Rowan, A., Cai, T., Clevers, H., Swanton, C., Nowak, M.A., et al. (2017). Origins of lymphatic and distant metastases in human colorectal cancer. Science *357*, 55–60, https://doi.org/10.1126/science.aai8515.

Neher, R.A., and Bedford, T. (2018). Real-time analysis and visualization of pathogen sequence data. J. Clin. Microbiol. https://doi.org/10.1128/JCM.00480-18.

Nejman, D., Livyatan, I., Fuks, G., Gavert, N., Zwang, Y., Leore, T., Rotter-maskowitz, A., et al. (2020). The human tumor microbiome is composed of tumor type-specific intra-cellular bacteria *980*, 973–980.

Laubert, T., Habermann, J.K., Hemmelmann, C., Kleemann, M., Oevermann, E., Bouchard, R., Hildebrand, P., Jungbluth, T., Burk, C., Esnaashari, H., et al. (2012). Metachronous metastasis- and survival-analysis show prognostic importance of lymphadenectomy for colon carcinomas. BMC Gastroenterol. *12*, 24, https://doi.org/10.1186/1471-230X-12-24.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J *17*, https://doi.org/10.14806/ej.17.1.200.

Ohl, M.E., and Miller, S.I. (2001). *Salmonella*: a model for bacterial pathogenesis. Annu. Rev. Med. https://doi.org/10.1146/annurev.med.52.1.259.

Pasolli, E., Asnicar, F., Manara, S., Zolfo, M., Karcher, N., Armanini, F., Beghini, F., Manghi, P., Tett, A., Ghensi, P., et al. (2019). Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. Cell. https://doi.org/10.1016/j.cell.2019.01.001.

Peinado, H., Zhang, H., Matei, I.R., Costa-Silva, B., Hoshino, A., Rodrigues, G., Psaila, B., Kaplan, R.N., Bromberg, J.F., Kang, Y., et al. (2017). Pre-metastatic niches: organ-specific homes for metastases. Nat. Rev. Cancer *17*, 302–317, https://doi.org/10.1038/nrc.2017.6.

Pita-Fernández, S., Alhayek-Aí, M., González-Martín, C., López-Calviño, B., Seoane-Pillado, T., and Pértega-Díaz, S. (2015). Intensive follow-up strategies improve outcomes in nonmetastatic colorectal cancer patients after curative surgery: a systematic review and meta-analysis. Ann. Oncol. https://doi.org/10.1093/annonc/mdu543.

Qian, B.-Z., Li, J., Zhang, H., Kitamura, T., Zhang, J., Campion, L.R., Kaiser, E.A., Snyder, L.A., and Pollard, J.W. (2011). CCL2 recruits inflammatory

monocytes to facilitate breast-tumour metastasis. Nature *475*, 222–225, https://doi.org/10.1038/nature10138.

Rakoff-Nahoum, S., Paglino, J., Eslami-Varzaneh, F., Edberg, S., and Medzhitov, R. (2004). Recognition of commensal microflora by Toll-like receptors is required for intestinal homeostasis. Cell. https://doi.org/10.1016/j.cell.2004.07.002.

Rantakari, P., Jäppinen, N., Lokka, E., Mokkala, E., Gerke, H., Peuhu, E., Ivaska, J., Elima, K., Auvinen, K., and Salmi, M. (2016). Fetal liver endothelium regulates the seeding of tissue-resident macrophages. Nature. https://doi.org/10.1038/nature19814.

Roggero, P., Liotto, N., Pozzi, C., Braga, D., Troisi, J., Menis, C., Giannì, M.L., Berni Canani, R., Paparo, L., Nocerino, R., et al. (2020). Analysis of immune, microbiota and metabolome maturation in infants in a clinical trial of *Lactobacillus paracasei* CBA L74-fermented formula. Nat. Commun. https://doi.org/10.1038/s41467-020-16582-1.

Rossi, E., Motta, S., Mauri, P., and Landini, P. (2014). Sulfate assimilation pathway intermediate phosphoadenosine 5′-phosphosulfate acts as a signal molecule affecting production of curli fibres in *Escherichia coli*. Microbiology (United Kingdom). https://doi.org/10.1099/mic.0.079699-0.

R Core Team (2017). R: A language and environment for statistical computing (R Foundation for Statistical Computing). https://www.R-project.org/.

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. Bioinformatics. https://doi.org/10.1093/bioinformatics/btu153.

Seubert, B., Grünwald, B., Kobuch, J., Cui, H., Schelter, F., Schaten, S., Siveke, J.T., Lim, N.H., Nagase, H., Simonavicius, N., et al. (2015). Tissue inhibitor of metalloproteinases (TIMP)-1 creates a premetastatic niche in the liver through SDF-1/CXCR4-dependent neutrophil recruitment in mice. Hepatology. https://doi.org/10.1002/hep.27378.

Soares, K.C., Foley, K., Olino, K., Leubner, A., Mayo, S.C., Jain, A., Jaffee, E., Schulick, R.D., Yoshimura, K., Edil, B., and Zheng, L. (2014). A preclinical murine model of hepatic metastases. J. Vis. Exp. https://doi.org/10.3791/51677.

Spadoni, I., Pietrelli, A., Pesole, G., and Rescigno, M. (2016). Gene expression profile of endothelial cells during perturbation of the gut vascular barrier. Gut Microbes *7*, 540–548.

Spadoni, I., Zagato, E., Bertocchi, A., Paolinelli, R., Hot, E., Di Sabatino, A., Caprioli, F., Bottiglieri, L., Oldani, A., Viale, G., et al. (2015). A gut-vascular barrier controls the systemic dissemination of bacteria. Science (New York, N.Y.) *350*, 830–834, https://doi.org/10.1126/science.aad0135.

Stan, R.V., Tse, D., Deharvengt, S.J., Smits, N.C., Xu, Y., Luciano, M.R., McGarry, C.L., Buitendijk, M., Nemani, K.V., Elgueta, R., et al. (2012). The diaphragms of fenestrated endothelia: gatekeepers of vascular permeability and blood composition. Dev. Cell. https://doi.org/10.1016/j.devcel.2012.11.003.

Tot, T. (2002). Cytokeratins 20 and 7 as biomarkers: usefulness in discriminating primary from metastatic adenocarcinoma. Eur. J. Cancer *38*, 758–763.

Vijay-Kumar, M., Aitken, J.D., Carvalho, F.A., Cullender, T.C., Mwangi, S., Srinivasan, S., Sitaraman, S.V., Knight, R., Ley, R.E., and Gewirtz, A.T. (2010). Metabolic syndrome and altered gut microbiota in mice lacking Toll-like receptor 5. Science (New York, N.Y.). https://doi.org/10.1126/science.1179721.

Wattam, A.R., Davis, J.J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., Conrad, N., Dietrich, E.M., Disz, T., Gabbard, J.L., et al. (2017). Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. Nucleic Acids Res. https://doi.org/10.1093/nar/gkw1017.

Wickham, H. (2009). Ggplot2: Elegant Graphics for Data Analysis, 2nd Edition (Springer). https://doi.org/10.1007/978-0-387-98141-3.

Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). ClusterProfiler: an R package for comparing biological themes among gene clusters. OMICS A J. Integr. Biol. https://doi.org/10.1089/omi.2011.0118.

Zagato, E., Mileti, E., Massimiliano, L., Fasano, F., Budelli, A., Penna, G., and Rescigno, M. (2014). *Lactobacillus paracasei* CBA l74 metabolic products and fermented milk for infant formula have anti-inflammatory activity on dendritic cells in vitro and protective effects against colitis and an enteric pathogen in Vivo. PLoS One. https://doi.org/10.1371/journal.pone.0087615.

Zhang, C., Zhang, L., Xu, T., Xue, R., Yu, L., Zhu, Y., Wu, Y., Zhang, Q., Li, D., Shen, S., et al. (2020). Mapping the spreading routes of lymphatic metastases in human colorectal cancer. Nat. Commun. https://doi.org/10.1038/s41467-020-15886-6.

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004



## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Antibodies | | |
| anti-human CD31 | DAKO | Cat #M0823; RRID:AB_2114471 |
| anti-human CD3 | DAKO | Cat # M7254; RRID:AB_2631163 |
| anti-human PV-1 | Sigma Aldrich | Cat #HPA002279; RRID:AB_1079636 |
| anti-mouse PV-1 | BD Biosciences | Cat #553849; RRID:AB_395086 |
| anti-mouse CD34 Alexa Fluor 488 | eBioscience | Cat # 53-0341-82; RRID:AB_2866440 |
| anti-mouse CD34 Alexa Fluor 647 | BD Biosciences | Cat # 560230; RRID:AB_1645200 |
| anti-human SOX9 | Abcam | Cat # ab185230; RRID:AB_2715497 |
| anti-mouse CK20 | Santa Cruz Biotechnology | Cat # sc17113; RRID:AB_2234445 |
| anti-mouse CD16/CD32 | eBioscience | Cat # 16-0161-85; RRID:AB_468899 |
| anti-mouse CD45.2 V450 | BD Biosciences | Cat # 560697; RRID:AB_1727495 |
| anti-mouse CD45.2 PerCP-Cy5.5 | BD Biosciences | Cat # 552950; RRID:AB_394528 |
| anti-mouse CD3 PerCP-Cy5.5 | Biolegend | Cat # 100328; RRID:AB_893318 |
| anti-mouse CD3 BV421 | Biolegend | Cat # 100228; RRID:AB_2562553 |
| anti-mouse CD11b Alexa Fluor 488 | BD Biosciences | Cat # 557672; RRID:AB_396784 |
| anti-mouse F4/80 PE-Cy7 | eBioscience | Cat # 25-4801-82; RRID:AB_469653 |
| anti-mouse Ly6C APC | BD Biosciences | Cat # 560595; RRID:AB_1727554 |
| anti-mouse Ly6G PE | BD Biosciences | Cat # 551461; RRID:AB_394208 |
| anti-mouse Ly6G BV510 | Biolegend | Cat # 127633; RRID:AB_2562937 |
| SYTO 16 green fluorescent nucleic acid stain | Invitrogen | Cat # S7578; RRID: AB_2890138 |
| Fixable Viability Dye eFluor® 510 | BD Biosciences | Cat # 564406; RRID:AB_2869572 |
| Fixable Viability Dye eFluor® 780 | BD Biosciences | Cat # 565388; RRID:AB_2869673 |
| anti-human α-SMA Tag #141Pr | Fluidigm | Cat #3141017D; RRID: AB_2890139 |
| anti-human CD31 Tag #151Eu | Fluidigm | Cat #3151025D; RRID: AB_2890140 |
| anti-human CD44 Tag #153Eu | Fluidigm | Cat #3153029D; RRID: AB_2890141 |
| anti-human Collagen type I Tag #169Tm | Fluidigm | Cat #3169023D; RRID:AB_2810857 |
| anti-human Cytokeratin 7 Tag #164Dy | Fluidigm | Cat #3164028D; RRID: AB_2890142 |
| anti-human Ki-67 Tag #168Er | Fluidigm | Cat #3168022D; RRID:AB_2811061 |
| anti-human SOX9 Tag #147Sm | Fluidigm | Cat #3147022D; RRID: AB_2890143 |
| Cell-IDTM-intercalator-Ir | Fluidigm | Cat # 201192A; RRID: AB_2890144 |
| Cy3-AffiniPure Donkey Anti-Rat IgG (H+L) | Jackson ImmunoResearch | Cat # 712-165-153; RRID:AB_2340667 |
| Anti-mouse Ly6G Monoclonal Antibody, Unconjugated | BD Biosciences | Cat # 551459; RRID:AB_394206 |
| Anti-mouse F4/80 Monoclonal Antibody, Unconjugated | Thermo Fisher Scientific | Cat # MF48000; RRID:AB_10376289 |
| Bacterial and Virus Strains | | |
| *E. coli* C17 | This paper | N/A |
| *E. coli* mCherry C17 | This paper | N/A |
| *E. coli* C17 virF::cat mutant (*E. coli C17 ΔvirF1*) | This paper | N/A |
| *Lactobacillus paracasei* CNCM I-5220 | N/A | |
| Biological Samples | | |

(*Continued on next page*)

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004





*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Training Cohort: Formalin fixed paraffin embedded samples of human CRC tumor and healthy colon (approved by the Ethics Committee of European Institute of Oncology License No. IEO/0799 and Humanitas Clinical and Research Center License No. Humanitas12/19) | European Institute of Oncology / Humanitas Clinical and Research Center | N/A |
| Validation Cohort: Formalin fixed paraffin embedded samples of human CRC tumor (approved by the Ethics Committee of INCLIVIA Intituto de Investigatiòn Sanitaria License No. PI15/02180 and Humanitas Clinical and Research Center License No. Humanitas12/19) | INCLIVIA Intituto de Investigatiòn Sanitaria / Humanitas Clinical and Research Center | N/A |
| **Chemicals, Peptides, and Recombinant Proteins** | | |
| Azoxymethane (AOM) | Sigma-Aldrich | Cat # A5486 |
| Dextran Sulphate Sodium (DSS) | TdB | Cat # DB001 |
| vancomycin hydrochloride | GoldBio | Cat # V-200-25 |
| neomycin sulfate | GoldBio | Cat # N-620-100 |
| ampicillin sodium | GoldBio | Cat # A-301-100 |
| metronidazole | MPBiomedicals | Cat # 155710 |
| lysozyme | Sigma-Aldrich | Cat # L6876 |
| 4',6-Diamidino-2-Phenylindole, Dihydrochloride (DAPI) | Invitrogen | Cat # D1306 |
| VECTASHIELD Mounting Medium | Vector Laboratories | Cat # H-1000 |
| FITC-Dextran 4 kDa MW | Sigma-Aldrich | Cat # 46944-500MG-F |
| XenoLight D-Luciferin - K+Salt | PerkinElmer | Cat # 122799 |
| Collagenase D | Roche | Cat # 11088866001 |
| Poly(vinylpolypyrrolidone) (PVPP) | Sigma-Aldrich | Cat # 77627-25G |
| MRS broth with Tween 80 | Biolife italiana | Cat # 4017292 |
| Tryptic Soy Broth | BD Difco | Cat #211825 |
| **Critical Commercial Assays** | | |
| BOND Polymer Refine Detection Kit | Leica Biosystems | Cat # DS9800 |
| BOND Epitope Retrieval Solution 1 (pH6) | Leica Biosystems | Cat # AR9961 |
| BOND Primary Antibody Diluent | Leica Biosystems | Cat # AR9352 |
| antigen retrieval reagent, pH 9 | Agilent | Cat # S236784-2 |
| Maxpar Water | Fluidigm | Cat # 201069 |
| Quick-RNA MiniPrep | Zymo Research | Cat # ZYR1055 |
| ImProm-II Reverse Transcriptase kit | Promega | Cat # A3803 |
| Fast Sybr Green PCR kit | Applied Biosystems | Cat # 4385614 |
| G'NOME DNA isolation kit | MPBiomedicals | Cat # 112010600 |
| QIAamp DNA Microbiome kit | QIAGEN | Cat # 51704 |
| QIAquick PCR Purification Kit | QIAGEN | Cat # 28104 |
| Diva Decloaker 10x | Biocare Medical | Cat # DV2004 |
| Rodent Block M | Biocare Medical | Cat # RBM961 |
| Rat on Mouse HRP-Polymer kit | Biocare Medical | Cat # RT517L |
| Betazoid DAB Chromogen Kit | Biocare Medical | Cat # BDB2004 |
| SMART-Seq v4 Ultra Low Input RNA Plus Kit | Clontech-Takara | Cat # 634894 |
| Nextera XT DNA Library Preparation Kit | Illumina | Cat # FC-131-1096 |

(*Continued on next page*)

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004



*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| **Deposited Data** | | |
| RNaseq dataset | This paper | Sequence Read Archive (SRA) BioProject: PRJNA647545 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA647545) |
| 16S rRNA dataset | This paper | Sequence Read Archive (SRA) BioProject: PRJNA647545 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA647545) |
| *E. coli* C17 complete genome | This paper | Gene Bank database BioProject: PRJNA647545 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA647545) |
| Macro for Fiji | This paper | https://zenodo.org/record/4560402#.YDeP1XnSJaQ. |
| Experimental Models: Cell Lines | | |
| MC38 pLUC cells | This paper | N/A |
| **Experimental Models: Organisms/Strains** | | |
| Mouse: $Apc^{Min/+}$C3arKO | Guglietta et al., 2016 | N/A |
| Mouse: C3aRKO | Humbles et al., 2000 | Cat # JAX:005712; RRID:IMSR_JAX:005712 |
| Mouse: C57BL/6J | Charles River | Cat # JAX:000664; RRID:IMSR_JAX:00066 |
| **Oligonucleotides** | | |
| FISH probes, see Table S4 | This paper | N/A |
| *Tnf-α* QuantiTect Primer Assay | QIAGEN | Cat # QT00104006 |
| *Tgf-β* QuantiTect Primer Assay | QIAGEN | Cat # QT00145250 |
| Primers, see Table S5 | This paper | N/A |
| **Recombinant DNA** | | |
| Plasmid: pLuc | Addgene | Addgene plasmid #17477; RRID:Addgene_17477 |
| Plasmid: pONmCherry | Addgene | Addgene plasmid #84821; RRID:Addgene_84821 |
| **Software and Algorithms** | | |
| Fiji | ImageJ | http://fiji.sc; RRID: SCR_002285 |
| MCD TM Viewer | Fluidigm | https://www.fluidigm.com/binaries/content/documents/fluidigm/consumables/software/mcd-viewer/mcd-viewer/fluidigm%3Afile |
| CellProfiler | Broad Institute | https://cellprofiler.org; RRID: SCR_007358 |
| HistoCAT | https://www.nature.com/articles/nmeth.4391 | https://bodenmillergroup.github.io/histoCAT/ |
| Living Image v4.3.1 | Perkin Elmer | http://www.perkinelmer.com/catalog/category/id/living%20image%20software; RRID: SCR_014247 |
| FlowJo v10.5.3 | TreeStar | https://www.flowjo.com/solutions/flowjo; RRID: SCR_008520 |
| PhyloPhlAn v3.0 | Asnicar et al., 2020 | http://segatalab.cibio.unitn.it/tools/phylophlan/; RRID: SCR_013082 |
| Prokka | Seemann (2014) | https://narrative.kbase.us/#catalog/modules/ProkkaAnnotation; RRID: SCR_014732 |
| Roary | Neher and Bedford (2018) | https://sanger-pathogens.github.io/Roary/; RRID: SCR_018172 |
| iTOL v5 | Letunic and Bork, 2019 | https://itol.embl.de; RRID: SCR_018174 |

(*Continued on next page*)

*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| clusterProfiler | Yu et al. (2012) | https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html; RRID: SCR_016884 |
| R v3.6.3 | R Core Team, 2017 | https://www.r-project.org/ ; RRID: SCR_001905 |
| FastQC v0.11.8 | Babraham Institute | https://www.bioinformatics.babraham.ac.uk/projects/fastqc/; RRID: SCR_014583 |
| Trimmomatic v0.39 | Bolger et al. (2014) | http://www.usadellab.org/cms/?page=trimmomatic; RRID: SCR_011848 |
| cutadapt v1.18 | Martin, 2011 | https://cutadapt.readthedocs.io/en/stable/index.html; RRID: SCR_011841 |
| Bowtie2 v2.3.4.3 | Langmead and Salzberg (2012) | http://bowtie-bio.sourceforge.net/bowtie2/index.shtml; RRID: SCR_016368 |
| Qiime2 v2019.7 | Bolyen et al., 2019 | https://qiime2.org/; RRID: SCR_018074 |
| DESeq2 v1.26.0 | Love et al., 2014 | https://bioconductor.org/packages/release/bioc/html/DESeq2.html; RRID: SCR_015687 |
| ggplot2 v3.3.0 | Wickham, 2009 | https://ggplot2.tidyverse.org/; RRID: SCR_014601 |
| Prism | GraphPad | http://www.graphpad.com/; RRID: SCR_002798 |
| **Other** | | |
| 60% lard high-fat diet | Envigo | Cat # TD.06414 |
| Columbia agar plates (5% sheep blood) | Oxoid | Cat # PB5039A |
| 0,1 mm Zirconia/Silica Beads | BioSpec Products | Cat # 11079101z |
| Count Bright Absolute Counting Beads | Invitrogen | Cat # C36950 |

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Maria Rescigno (maria.rescigno@hunimed.eu).

### Materials availability

This study did not generate unique reagents.

### Data and code availability

Original data have been deposited to Zenodo Repository: https://zenodo.org/record/4560402#.YDeP1XnSJaQ. The accession number for the RNaseq, 16S rRNA datasets and *E. coli* C17 whole genome sequence reported in this paper is BioProject: PRJNA647545 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA647545). Macro for FIJI generated in this study have been deposited to Zenodo Repository:https://zenodo.org/record/4560402#.YDeP1XnSJaQ.

## EXPERIMENTAL MODELS AND SUBJECT DETAILS

### Human samples

#### *CRC training cohort*

Formalin fixed paraffin-embedded (FFPE) samples of colorectal cancer and healthy colon were retrieved from European Institute of Oncology (Milan, Italy) and Humanitas Clinical and Research Center (Rozzano, Milan, Italy) Biobanks. 157 patients were enrolled into a clinical protocol entitled ‘‘Ruolo della barriera vascolare intestinale nell’ambito dei tumori del colon’’ (IEO/0799). The remaining 22 patients were enrolled into a clinical protocol entitled ‘‘Identification of a new vascular marker correlating with colon cancer metastatization’’ (Humanitas12/19). A total of 179 patients were included in the training cohort:

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004



- 79 CRC patients that developed distant metachronous metastases (not local recurrence) within 10 years;
- 100 CRC patients did not develop distant metachronous metastases within 10 years.

As ancillary study we have analyzed 10 colon samples belonging to patients with healthy colon. These healthy colons were removed for technical reason during surgery of ovary carcinoma.

Patients included in the study, selected from the institutional patient registry, granted all these characteristics:

- They had histologically proven colon adenocarcinoma or proximal rectal cancer suitable for curative surgery and adjuvant chemotherapy if indicated;
- They underwent surgery with curative intent between 2000-2014 (in order to have at least 5 years of follow-up);
- They signed a written informed consent for general research purposes.

Patients were excluded from the study if:

- They had previous or concomitant malignancies, except for adequately treated basal or squamous cell skin cancer, *in situ* cervical cancer, or other cancer for which the patient has been disease-free for five years prior to colon cancer diagnosis;
- They had any serious underlying medical conditions or serious co-morbidities, which could impair prognosis during 5-years follow-up besides resected colorectal cancer;
- They had known inflammatory bowel diseases (i.e. ulcerative colitis, Crohn's disease);
- They had distal rectal cancer or they underwent radiotherapy/chemotherapy before surgery;

Demographics and clinical characteristics from this training cohort are in Table S1.

#### *CRC validation cohort*

Formalin fixed paraffin-embedded (FFPE) samples of 51 colorectal cancer patients were retrieved from INCLIVIA Intituto de Investigatiòn Sanitaria (Valencia, Spain) and Humanitas Clinical and Research Center (Rozzano, Milan, Italy) Biobanks. 29 patients were enrolled into a clinical protocol entitled *"Enfermedad minima residual en canceres colorectales de alto riesgo resecados. Valor de las biopsias liquidas en el seguimiento y analisis de la heterogeneidad tumoral"* (PI15/02180). The remaining 22 patients were enrolled into a clinical protocol entitled "*Identification of a new vascular marker correlating with colon cancer metastatization*" (Humanitas12/19). A total of 51 patients were included in the validation cohort:

- 24 CRC patients that developed distant metachronous metastases (not local recurrence) within 4 years;
- 27 CRC patients did not develop distant metachronous metastases within 4 years.

Demographics and clinical characteristics from the training cohort are maintained in the validation cohort and they are listed in Table S2.

Hematoxylin and eosin (H&E) stained sections from each sample were subjected to pathology review to confirm that the tumor specimen was histologically consistent with the allowable colon adenocarcinoma.

### Mouse models

All mice used were on a C57BL/6 background. C57BL/6 8 weeks old mice were purchased from Charles River. $Apc^{Min/+}$C3arKO mice were obtained by crossing C57BL/6J- $Apc^{Min/+}$ with C3arKO mice in our facility (Guglietta et al., 2016). C3arKO mice were generated in the laboratory of Craig Gerard (Childrens' Hospital Boston, Harvard MS). Colony was maintained crossing male $Apc^{Min/+}$C3arKO with female $Apc^{+/+}$C3arKO (C3arKO) littermates. Experiments were performed using both male and female, unless otherwise specified. Mice were housed under specific-pathogen-free (SPF) conditions at Campus IFOM-IEO (Milan, Italy) and Humanitas Clinical and Research Center (Rozzano, Milan, Italy). All animal experiments were performed under protocols (n.139/15, 662/18, 167/2020-PR and 927-2020-PR) approved by the Italian Ministry of Health, and consistent with national (D.L. N. 26, G.U. March 4, 2014) and international law and policies (EEC Council Directive 2010/63/EU).

#### *AOM/DSS model of colon tumorigenesis*

Newly purchased eight weeks old C57BL/6J male mice were randomized into 5 mice/cage and housed for one week to normalize gut microbiome. For the AOM/DSS experiment, mice were intraperitoneally injected with azoxymethane (AOM, Sigma-Aldrich # A5486) 10 mg/Kg of body weight. After 7 days, AOM injected mice received Dextran Sulphate Sodium (DSS) 1% w/v in drinking water (group AOM/DSS). As a control group, mice were not injected with AOM, but received Dextran Sulphate Sodium (DSS) 1% w/v in drinking water (group DSS). Two cohorts received DSS 1% w/v in drinking water for 7 days (AOM/DSS and DSS-treated). After 7 days of DSS treatment, both groups were allowed to recover for 14 days. This schedule was repeated for 3 cycles. Body weight and survival were assessed periodically (two times per week).

#### *Antibiotics administration*

After one week of recovery from the last cycle of DSS, mice were randomly assigned into water ($H_20$ control group) or antibiotic (ABX) treatment group. Mice of ABX group received a three-antibiotic cocktail in the drinking water containing vancomycin hydrochloride (V-200-25, final concentration 0.5g/L, GoldBio), neomycin sulfate (# N-620-100, final concentration 1g/L, GoldBio) and ampicillin sodium (A-301-100, final concentration 1g/L, GoldBio) as previously reported(Rakoff-Nahoum et al., 2004). Since it is known that mice generally keep off drinking the antibiotic cocktail for the bad taste of metronidazole, we gavaged mice with metronidazole (2 mg/mouse in

200 ul, MPBiomedicals) every other day. Fresh antibiotic cocktail was replaced every four days. After one week of ABX treatment, mice were intrasplenically injected with MC38 pLUC cells. Multiple experiments were performed simultaneously with a single control group (AOM/DSS). The same ABX mix was used to treat $Apc^{Min/+}$C3arKO, C3arKO and WT mice for one week.

***High fat diet experiment***

For the High Fat Diet (HFD) experiment, eight weeks old C57BL/6J male mice were maintained on chow diet as control diet or 60% lard high-fat diet (HFD, TD.06414) for one week. The HFD TD.06414 was purchased from Envigo. At the experimental end points, mice were sacrificed for organ harvest.

## METHODS DETAILS

### Immunohistochemistry

Immunohistochemical analysis of tumor/healthy colon was performed on FFPE tissue sections. Immunohistochemical staining was performed in a Bond Max Automated Immunohistochemistry Vision Biosystem (Leica Microsystems GmbH, Wetzlar, Germany) using the Bond Polymer Refine Detection Kit (DS9800). 3-μm-thick sections were prepared from FFPE human colorectal cancer tissue blocks, deparaffinized, pre-treated with the Epitope Retrieval Solution 1 (pH6) at 100°C (20 min for CD31, 40 min for PV-1 and 20 min for CD3) and then incubated for 30 min with primary antibody anti-human CD31 (DAKO, #M0823, final concentration = 6.8 ug/ml), anti-human PV-1 (Sigma Aldrich, # HPA002279 1:200) or anti-human CD3 (DAKO, #M7254 1:50) antibodies diluted in Bond Primary Antibody Diluent for PV-1 and CD31 (AR9352) or in Staining Buffer for CD3 (PBS and 0.1% BSA). Serial sections of the same FFPE sample were stained for PV-1 and CD31 respectively. BIOP VSI reader on Fiji (ImageJ) software was used for CD3 quantification. For anti-mouse Ly6G and F4/80 staining, sections of FFPE liver tissues were dewaxed in xylene and rehydrated in ethanol. Sections were then pretreated in a microwave oven (two cycles for 3 minutes each at 800 W) in H2O plus Diva Decloaker 10x (Biocare Medical, #DV2004). Endogenous peroxidase was blocked for 20 min in H2O containing 2% H2O2. Unspecific sites were blocked with Rodent Block M (Biocare Medical, #RBM961) 30 minutes and tissues were incubated for two hours with affinity-purified Ig against Ly6G 1:200 (Clone 1A8; BD Bioscience, #551459) or F4/80 1:100 (Clone BM8; Invitrogen #MF48000) in PBS supplemented with 0.1% BSA. Rat on Mouse HRP-Polymer kit (Biocare Medical #RT517L) was used as secondary antibody. After washing, slides were developed with DAB (3,30-diaminobenzidine) (Betazoid DAB Chromogen Kit; Biocare Medical #BDB2004) and counterstained with Hematoxylin. Tissues were dehydrated with ethanol, mounted with Eukitt and analyzed with an Olympus BX61 virtual slide scanning system.

### Analysis of PV-1 protein detection on human samples

All tissue sections stained for PV-1 and CD31 were scored blindly by the pathologist (G.B.). He calculated the percentage of PV-1+ among the CD31+ vessels. If two tissue sections for one patient were analyzed (i.e. two samples from distinct areas of the same tumor), the two sections were scored independently and paired at the end. Tumors with discordant scores on the two sections were upgraded to the highest score. A detailed description of the scoring system, together with representative scoring results, are provided in Figure S1A. Finally, the association between PV-1 detection and survival outcomes was tested by a third investigator (S.Ga), who did not participate in the scoring process.

### Fluorescent *in situ* hybridization (FISH)

Fluorescent *in situ* hybridization was performed on 5-μm sections from FFPE blocks. Tissue sections were washed in 0.1 M Tris-HCl pH 7.4 for 15 minutes, after which gram-positive bacterial cell walls were hydrolyzed using lysozyme (10 mg/mL) for 30 minutes at 37°C. The used probes (EUB 1, 2, 3 and *E. Coli C17*, listed in Table S4) were designed to specifically target different regions of the bacterial 16S ribosomal RNA (16S rRNA) or specifically *E. coli C17*. The scramble probe (Table S4) is a "non-sense probe" and it has been used as FISH negative control. All the probes were manufactured by Sigma-Aldrich and labeled with Alexa Fluor 488, Cy3 or Alexa Fluor 647. Slides were incubated with 5 ng/μl of each probe diluted in prewarmed hybridization buffer (0.9 M NaCl, 20 mM Tris-HCl pH 7.4, 0.01% SDS) overnight at 50°C in a humid chamber. Slides were then washed twice with a pre-warmed washing buffer (0.9 M NaCl, 20 mM Tris-HCl pH 7.4) at 50°C. Tissues were counterstained with DAPI (1:5000) and mounted with VECTA SHIELD Mounting Medium. Confocal images were acquired using SP8 confocal microscope through HCX PL APO 40/63X (NA 1.25) oil immersion objective. For SOX9 double staining, after FISH protocol, we proceed with the protocol for immunofluorescence (described below) after FISH protocol.

### Immunofluorescence

Mousee colon tissues were fixed overnight (o/n) in PLP buffer (1% paraformaldehyde, L-Lysine 0.2M pH 7.4 and 25 mg NaIO4). Then, they were then washed, dehydrated with 20% sucrose for at least 4 hours and included in OCT compound (Sakura). 10-μm cryosections were rehydrated, blocked with 0.1M Tris- HCl pH 7.4, 2% fetal bovine serum (FBS), 0.3% Triton X-100 and stained with the following antibodies: anti-mouse PV-1 (clone MECA-32, 1:100 BD pharmigen), anti-mouse CD34-Alexa Fluor 488 (clone RAM34, 1:100 eBioscience) or anti-mouse CD34-Alexa Fluor 647 (clone RAM34, 1:50 eBioscience), anti-human SOX9 (ab185230, 1:250 Abcam) and anti-mouse CK20 (clone G-20, 1:50 Santa Cruz Biotechnology SC17113). Primary antibodies were incubated o/n at 4°C. Slices were then incubated with the appropriate fluorophore-conjugated secondary antibody. Before imaging, nuclei were counterstained with (DAPI). Confocal microscopy was performed on: 1) Leica TCS SP5/SP8 laser confocal scanner mounted on a Leica DMI

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004



6000B inverted microscope equipped with motorized stage. Violet (405nm laser diode), blue (488nm argon laser), yellow (561nm laser diode) and red (633nm laser diode) laser lines that have been used for excitation. All images were acquired with an HCX PL APO 40X/63X (NA 1.25) oil immersion objective. Leica LAS AF was used for acquisitions. 2) Leica TCS SP8 laser scanning confocal microscope, equipped with 405, 488, 552 and 638nm diode laser. Images were acquired with a HC PL FLUOTAR 40x/1.30 oil and HC PL APO CS2 63/1.40 oil objectives. Leica LASx software was used for acquisition. FiJi software package was used for images analysis and fluorescence quantification.

We have designed different Macros for FiJi in order to quantify the mean fluorescent intensity (MFI) of PV-1 among the CD34 positive vessels (Macro n°1), quantify the number of bacteria for each field of view. Each bacterium was calculated as positive for eubacteria staining and negative for scramble signal (Macro n°2) and quantify the number of bacteria that were located near SOX9+ cancer cells (Macro n°3). Each macro is available in Zenodo Repository at https://zenodo.org/record/4560402#.YDeP1XnSJaQ.

## Imaging Mass Cytometry (IMC)

### *Tissue preparation and staining*

Paraffin-embedded tissue section of human metastatic liver, collected from IEO Biobank and belonging to a patient enrolled in the study (see above section), was dewaxed in xylene followed by alcohol series rehydration steps. Heat-induced epitope retrieval was conducted in antigen retrieval reagent pH 9 1X (Agilent # S236784-2) for 30 minutes at 96°C. After immediate cooling and washing in Maxpar Water, the slide was blocked in 3% BSA in Maxpar water (blocking buffer) for 45 minutes at room temperature. Sample was incubated in a wet chamber overnight at 4°C with a mixture of 12 metal-conjugated antibodies diluted in blocking buffer. The antibodies used for this study were: α-SMA (clone 1A4, Tag #141Pr, Product #3141017D, 1:200), CD31 (clone EPR3094, Tag #151Eu, Product #3151025D, 1:100), CD44 (clone IM7, Tag #153Eu, Product #3153029D, 1:100), Collagen type I (polyclonal, Tag #169Tm, Product #3169023D, 1:400), Cytokeratin 7 (clone RCK105, Tag #164Dy, Product #3164028D, 1:100), Ki-67 (clone B56, Tag #168Er, Product #3168022D, 1:50) and SOX9 (clone EPR14335, Tag #147Sm, Product #3147022D, 1:50). Samples were then washed in 0.2% Triton X in Maxpar PBS for 8 minutes and exposed to Cell-IDTM-intercalator-Ir 1:400 (Catalog number 201192A Fluidigm) for 30 minutes at RT for nuclei staining. Slides were washed in Maxpar water twice for three minutes. Sections were airdried at least for 20 minutes at room temperature before IMC analysis.

## Imaging mass cytometry acquisition

Images were acquired with the Hyperion Imaging System (Fluidigm). Briefly, slide was inserted into the Hyperion laser ablation chamber (Giesen et al., 2014) where the sample was scanned by a pulsed UV laser: each pulse samples the stained tissue from a $1\mu m^2$ tissue spot for a comprehensive analysis of the selected region of interest (ROIs). From each spot, a plume is formed, ionized and detected using CyTOF technology (Fluidigm). For each spot, identified with a pre-determined coordinate, metal-tagged antibodies are detected simultaneously by CyTOF and their digitalized signals are used to generate an image with MCD™ Viewer software (Fluidigm). Images were then processed with Fiji, CellProfiler and HistoCAT softwares.

## FITC-Dextran permeability assay

Mice were starved over-night and then orally administered through gavage with ~400 mg/Kg FITC-Dextran (4 kDa; Sigma-Aldrich, # 46944-500MG-F). Blood was collected from the heart after four hours and the concentration of FITC-Dextran in plasma samples was measured as fluorescence intensity (Clariostar Plus Microplate Reader; BMG Labtech).

## Intra-splenic injection

A murine colon cancer cell line (MC38) stably transfected with the firefly luciferase (Luc) gene (Addgene plasmid #17477) was generated. To establish a model of liver metastasis mice were anesthetized with ketamine/xylazine. Then the spleens were exposed to allow direct intrasplenic injections of $2 \times 10^5$ Luc-labelled MC38 (MC38pLuc) cells in 50 μL 1mM EDTA in PBS on day 63 after the AOM/DSS protocol or 24 hours after the last *E. coli C17* administration. Ten minutes after the injection of tumor cells, spleens were removed using a cauterizer. Mice were sacrificed 30 days after injection of the tumor cells, and liver were analyzed immediately. For *ex vivo* imaging, mice were administered by intra-peritoneal injection with d-luciferin (150 mg/kg; Perkin Elmer). Eight minutes later, mice were sacrificed, the livers were collected and acquired using counted using the *ex vivo* IVIS Lumina III system (Perkin Elmer). Collected images were analyzed using Living Image 4.3.1 software (Perkin Elmer): ROIs were drawn around each liver and radiant efficiency calculated. Approval for these studies was obtained from the Committee on Animal Research of Ministry of Health.

## Flow cytometry on liver isolated cells

Mice were perfused with sterile PBS and livers were harvested in complete medium (RPMI, 5% fetal bovine serum North American, 1% Penicillin/Streptomycin and 1% Glutamine). Livers were enzymatically digested for 30 minutes at 37°C with 1 mg/mL Collagenase D (1 mg/ml, Roche). Collected cells were submitted to a discontinuous Percoll gradient (80/40) to separate mononuclear cells from hepatocytes and stromal cells through centrifugation for 20 minutes at 2000 g. Cell were surface-labeled with the indicated antibodies for 30 minutes 4°C. Dead cells were excluded by using Fixable Viability Dye eFluor® 510 or 780 (564406 and 565388 respectively, BD Biosciences, 1:1000). The immune cells were incubated with anti-CD16/CD32 antibody (clone 93, eBioscience, 1:100) and stained with the following antibodies: anti-mouse CD45.2 Pacific Blue or PerCP-Cy5.5 (clone 104, BD Biosciences, 1:200), anti-mouse CD3 PerCP-Cy5.5 (clone 145-2c11, BD Biociences, 1:200) or B421 (clone 17A2, BioLegend, 1:200), anti-mouse CD11b

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004





FITC (clone M1/70, BD Biosciences, 1:200), anti-mouse F4/80 PECy7 (clone BM8, eBioscience, 1:200), anti-mouse Ly6C APC (clone AL-21, BD Biosciences, 1:200), anti-mouse Ly6G PE (clone 1A8, BD Biosciences, 1:200) or BV510 (clone 1A8, BioLegend, 1:200). For SYTO16 detection, cells were then permeabilized using an intracellular staining kit (eBioscience, #00-5523-00) and stained with SYTO16 green fluorescent nucleic acid (Invitrogen, # S7578).

Flow cytometry was performed on a FACSCanto II and Fortessa (BD Biosciences) platform and results were analyzed using FlowJo software version 10.5.3. Neutrophils were identified as CD11b+ Ly6G+ cells in the CD45+ CD3- gate, inflammatory monocytes were identified as (not neutrophils) CD11b+ Ly6Chigh cells, macrophages were identified as (not neutrophils) CD11b+ F4/80+ cells and not immune cells (*E. coli mCherry C17* experiment) were identified as CD45- and SYTO16+ in the gate of live cells. Every gate was done on live cells. Absolutes numbers were calculated using Count Bright Absolute Counting Beads (Invitrogen #C36950).

### RT-qPCR assay

Liver tissues were homogenized using Tissue Lyser (Qiagen). Total RNA was purified using Quick-RNA MiniPrep (Zymo Research). cDNA synthesis was performed using ImProm-II Reverse Transcriptase kit (Promega) following manufacturer's instruction and random primers (0,5 ug/ul, Invitrogen). To check the presence of *E. coli* C17 in colon tissues and feces, we extracted bacterial DNA using a protocol described below, then we searched for *E. coli* C17 presence using designed strain specific primers targeting *E. coli* C17. For the detection of *Tnf-*$\alpha$ and *Tgf-*$\beta$ expression, QuantiTect Primer Assays (Qiagen), QT00104006 and QT00145250 respectively, were used. Real-time PCR reactions were carried out using the Fast Sybr Green PCR kit (QuantiStudio 7 Flex RrealTime PCR, Applied Biosystems). The relative expression levels were calculated by the $\Delta\Delta$CT method after normalization to the average of *Rpl32* or 16s rRNA level, for mouse gene expression and amplification of bacterial genes respectively. Primers sequences are listed in Table S5.

Stool bacterial load was quantified by analyzing the level of 16s rRNA gene expression in mice treated or not with antibiotics. We calculated the Ct average of mice not treated with antibiotics; we divided this value for each Ct value of every single mouse. Then, we transformed each measurement in a logarithmic value.

### RNA QC, library preparation and sequencing

RNA quality control was performed with the Agilent 4200 Tape Station system using the High Sensitivity RNA ScreenTape analysis kit (Agilent, Santa Clara, CA, USA), only RNAs having a RIN>7.5 were used for library preparation. Libraries for mRNA sequencing were prepared starting from 2 ng of total RNA for each sample by using the SMART-Seq v4 Ultra Low Input RNA Plus Kit (Clontech-Takara). The full-length cDNA synthetized with this kit was processed with the Nextera XT DNA Library Preparation Kits (Illumina, San Diego, CA, USA). Final libraries were checked using the Agilent TapeStation 4200 using the High Sensitivity DNA ScreenTape analysis kit. All samples were sequenced on an Illumina NextSeq 550 at an average of 20 M 76-bp single-end reads.

### RNASeq data analysis

Quality control of raw data was performed with fastqc v0.11.8. Quality filtering was performed with trimmomatic v0.39 setting the following parameters: CROP:72 AVGQUAL:30. High-quality reads were aligned to the mouse reference genome (GRCm38 primary assembly genome) using STAR (v2.6.1a). Reads were assigned to genes with featureCounts (v1.6.4) using the Mus_musculus.GRCm38.101.gtf gene transfer file (GTF) as a reference annotation file for genomic feature boundaries. Exploratory data analysis, and analysis of differential gene expression were performed using DESeq2 (v1.28.1) R package. Differentially expressed genes (DEGs) were selected with padj < 0.05 - Benjamini–Hochberg multiple test correction (FDR).

### Bacterial CFUs from liver tissue

Livers from $Apc^{Min/+}$C3arKO, C3arKO, WT ($Apc^{Min/+}$ littermates) and C57BL/6J mice were aseptically harvested and incubated 20 minutes at 37°C with gentamycin (50 μg/ml). Organs were then digested with collagenase D (1 mg/ml, Roche) for 30 minutes at 37°C. Cells were lysed with 1% sodium-deoxycholate and plated on Columbia agar plates with 5% Sheep Blood (Oxoid) to evaluate intracellular bacteria. Agar plates were incubated at 37°C for 3-5 days under aerobic and anaerobic conditions.

### Bacterial DNA extraction from fecal samples

DNA from fecal pellets was extracted with G'NOME DNA isolation kit (MP) following a published protocol (Furet et al., 2009). Briefly, fecal pellets, stored at -80°C, were homogenized in 550 μl Cell Suspension Solution (G NOME DNA Kit). After addition of 50 μl RNase Mix (G NOME DNA Kit) and 100 μl Cell Lysis/Denaturing Solution (G NOME DNA Kit) samples were incubated at 55°C for 30 minutes. After adding 25 μl Protease Mix (G NOME DNA Kit) samples were incubated for further 2 hours at 55°C. Samples then underwent mechanical disruption of bacterial cells with 0,1 mm zirconia/silica beads (BioSpec #11070101z) in FastPrep®-24 homogenizer (MP Biomedicals). Lysates were retrieved. Beads were washed three times with 400 μl of TENP buffer (50 mM Tris pH 8, 20 mM EDTA pH 8, 100 mM NaCl, 1% PVPP). Supernatants were pooled with the original lysate and precipitated with isopropanol. DNA pellet was resuspended in 400 μl water and incubated with 100 μl of Salt Out Mixture (G NOME DNA Kit) to remove impurities. Samples were then precipitated in 100% ethanol and DNA pellet washed with 70% ethanol. DNA pellets were dried and resuspended in water.

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004



### Bacterial DNA extraction from colon and liver samples

Murine colons and livers were harvested in 2 mL tubes and stored at -80°C until further usage. After tissue homogenization we used DNA Microbiome Kit (QIAamp®, #51704, QIAGEN) to deplete the eukaryotic DNA from the sample. Bacterial DNA from paraffin blocks of CRC patients was extracted using the same kit with some adaptations. Paraffin slices (3-5, 10μm thick) underwent three cycles of deparaffinization (xylene) and rehydration (ethanol 100%). After centrifugation with a speed vacuum, protein digestion was done at 56 C with constant shaking (600RPM) for three hours. The lysates underwent bead beating with 200μl 0.1mm zirconia/silica beads (Biospec - #11079101z) for 10 minutes at full speed. Starting from this step, we followed the G NOME DNA isolation kit (MP). Lysates were retrieved. Glass beads were washed three times with 400 μl of TENP buffer (50 mM Tris pH 8, 20 mM EDTA pH 8, 100 mM NaCl, 1% PVPP). Supernatants were pooled with the original lysate and precipitated with isopropanol. DNA pellet was re-suspended in 400 μl water and incubated with 100 μl of Salt Out Mixture (G NOME DNA Kit) to remove impurities. Samples were then precipitated in 100% ethanol and DNA pellet washed with 70% ethanol. DNA pellets were dried and resuspended in water.

### 16S rRNA analyses

Fecal material and organs were sampled from $Apc^{Min/+}$C3arKO, C3arKO and WT mice at 25 weeks of age and bacterial DNA purified as described above. The subsequent steps involving sample preparation for sequencing conducted in the laboratory of Prof. Marco Ventura (GenProbio s.r.l., Parco Area delle Scienze n.11/A 43124 University of Parma, Italy). The bioinformatics analyses were conducted by one of the authors (D.B). V3 region of the 16S rRNA gene sequence was amplified from extracted DNA using primer pair Probio_Uni/Probio_Rev (Milani et al., 2013). 16S rRNA gene amplification and amplicon checks were carried out, and sequencing of amplicon libraries was performed using a MiSeq (Illumina) at the DNA sequencing facility of GenProbio srl. Following sequencing, the quality of the fastq files was checked with FASTQC and reads were trimmed to 150 nt with Trimmomatic v0.39 (Bolger et al., 2014). Sequences of amplification primers and reads with more than 3 unknown (N) nucleotides were removed using cutadapt v1.18. Alignment to the mouse reference genome (GRCm38) was performed with Bowtie2 (Langmead and Salzberg, 2012) in order to remove host-derived sequences. High quality and cleaned sequences were analyzed using the Qiime2 platform (v2019.7) (Bolyen et al., 2019). The qiime dada2 denoise-paired command was used to denoise reads, remove chimeras and output the Amplicon Sequence Variants (ASVs). A phylogenetic tree was built using the qiime phylogeny align-to-tree-mafft-fasttree command and diversity measures (alpha- and β-diversity indices) were calculated using the qiime diversity core-metrics-phylogenetic function with a sampling depth of 10000 sequences. Community alpha diversity was evaluated by Chao1 index and represented by box-and-whisker plot. Community dissimilarities (β-diversity) were evaluated by Bray-Curtis distance and represented by a PCoA plot. The qiime diversity alpha-group-significance function was used to test differences of alpha diversity indices across experimental groups with Kruskal Wallis pairwise test. The qiime diversity beta-group-significance function was used to assess differences in the microbiome composition across the different experimental groups with Permutational Multivariate Analysis of Variance (PERMANOVA). Q2-feature-classifier, trained on the SILVA132 99% OTUs, specifically on the V3 region, was used to perform taxonomic classification. Raw taxonomic counts were processed with DESeq2 package as recommended by McMurdie and Holmes for 16S microbiome data (McMurdie and Holmes, 2014). Bacteria showing different abundance balance between colon and liver of tumor-bearing mice and control mice were selected testing the interaction of tissue type and genotype ($Apc^{Min/+}$C3arKO or WT) with Likelihood Ratio test.

### *Escherichia coli* isolation and amplification

To isolate tumor associated bacteria, tumors and livers collected from $APC^{Min/+}$C3arKo mice were homogenized in sterile phosphate buffered saline solution (PBS), centrifuged, and both supernatants and pellets were plated on Columbia Sheep Blood Agar plates and incubate in aerobic conditions at 37°C. We proceed with the isolation and screening of different colonies using colony PCR techniques. Briefly, cells from a single colony were picked up from agar plates by using a sterile pipette tip and suspended into lysis buffer (Tris/EDTA, 0.2% SDS, and 10 mM EDTA). The lysate was then centrifuged, and the supernatant was collected and used as PCR template. PCR amplification was performed on the 16SrRNA amplifying V1-V4 with 28F (5'-GAGTTTGATCNTGGCTCAG-3'), 907R (5'-CCGTCAATTCMTTTRAGTTT-3') and V5-V7 with 1392R (5′-ACGGGCGGTGTGTRC-3′), 926F (5′-AAACTYAAAKGAATTGACGG-3′). PCRs were performed with Phusion High-Fidelity DNA Polymerase (2 U/μL) (New England Biolabs) in a total volume of 50 μl. Each cycle consisted of 60 s at 98°C, 45 s at 63°C, and 60 s at 72°C, with a final extension of 5 min. The PCR products were subjected to 1,5% agarose gel electrophoresis for analysis and purified with QIAquick PCR Purification Kit (Qiagen) according to the manufacturer's protocol. PCR products were purified, sequenced and blasted for bacteria identification. For the taxonomic characterization of the isolate genomes, we used PhyloPhlAn (Asnicar et al., 2020) (version 3.0) to assign the species-level genome bins (SGBs (Pasolli et al., 2019)) to the isolate genomes with the following parameters: "-d SGB.Jan19 –add_ggb –add_fgb". To further characterize the isolate genomes and according to the assigned putative taxonomic label from the phylogenetic analysis, we used Prokka (Seemann, 2014) and Roary (Neher and Bedford, 2018) to annotate and perform a pangenome analysis, respectively. In particular, the gene annotations performed with Prokka was run using the default parameters. The pangenome analysis with Roary was executed using the following parameters: "-e -z -g 1000000". The resulting tree was visualized using iTOLv5. The genetic features of the *Escherichia coli* isolate C17 were compared to an internal database of 32 *Escherichia coli* strains. The presence or absence of a gene in each strain was represented with a heatmap. Genes found only in the C17 isolate were analyzed with cluster Profiler (Yu et al., 2012) in order to identify enriched KEGG molecular pathways. In Table S3 we have reported a description of genes present in *E. coli* C17. This table was generated using PATRIC database (Wattam et al., 2017) as reference.

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004





### Construction of *E. coli* mCherry C17

*E. coli* C17 was transformed by electroporation using pONmCherry plasmid, a gift from Howard Shuman (Gebhardt et al., 2017)(Addgene plasmid # 84821). 5 ml of exponential phase cultures *E. coli* C17 (0D600= 0.6) were harvested by centrifugation at 6000 rpm for 15 min. After washing the pellet twice with 5 ml ice-cold washing solution and once with 1ml of ice-cold 10% glycerol, cells were re-suspended 1:100 ml of 10% ice-cold glycerol, ready for electroporation. For each transformation, 40-μl cells were mixed with around 100 ng of plasmid DNA and incubated on ice for 10 min. The mixture was then transferred to a 2-mm gap cuvette (Bio-Rad, USA) and electroporation was performed at 2,500 V, 200 Ω and 25 μF, using a Gene Pulser Xcell TM electroporator (Bio-Rad, Munchen, Germany). Then the electroporated cells were transferred in 500μl of fresh SOB medium supplemented with $MgCl_2$ (1M, 10 μl/ml) and Glucose (1M, 20 μl/ml) allowing them to recover by incubation at 37°C for 1 h. Cells (100 ul) were then plated onto selective agar plates (TSB-chloramphenicol agar plates) to recover transformants. Expression of the mCherry gene did not alter the growth rate of the bacterium.

### *E. coli* C17 and *E. coli* MCherry C17 growth and *in vivo* administration conditions

For culturing of *E. coli* C17, Tryptic Soy Broth (TBS, BD Difco Cat 211825) medium was used as a broth. In indicated cases (*E. Coli* mCherry C17), TSB medium was supplemented with 30 μg/ml of chloramphenicol (SIGMA, Cat #CO378-5G). For *in vivo* experiment of GVB impairment, liver bacteria translocation and PMN maturation, stationary-phase cultures were obtained by growing bacteria overnight for approximately 15 h in TSB broth w or w/o chloramphenicol at 37°C under well-aerated conditions (170 rpm on an orbital shaker). Exponential phase cultures were in turn prepared by diluting stationary phase cultures until an OD 600 of 0.05 in fresh broth and allowing further incubation at 37°C until an OD600 of 0.6 was reached. Then, 10^7 CFUs of *E. coli* C17 or *E. coli mCherry* C17 resuspended in 200 μl of vehicle (Sodium carbonate) were fed to mice by gavage every other day for one week. Bacteria were prepared fresh for every administration. Mice were sacrificed 24 hours after last *E. coli C17* or *E. coli mCherry* C17 administration. For *E. coli* C17 administration to $Apc^{Min/+}$C3arKO mice, 10^9 CFUs resuspended in 200 μl of vehicle (Sodium carbonate) were fed to mice by gavage every other day for three weeks. Bacteria were prepared fresh for every administration. Mice were sacrificed 24 hours after last *E. coli C17* or *L. paracasei* administration.

### *Lactobacillus paracasei CNCM I-5220* growth and *in vivo* administration conditions

Lactobacillus paracasei *CNCM I-5220* (Heinz Italia SpA, Latina, Italy), International Depository Accession Number LMG P-24778 was grown in MRS broth (Biokar diagostics) in semi-anaerobic conditions. For *in vivo* experiment of GVB impairment, liver bacteria translocation and PMN maturation on $Apc^{Min/+}$C3arKO mice, stationary-phase cultures were obtained by growing bacteria overnight for approximately 15 h in MRS broth at 37°C in semi-anaerobic conditions. Exponential phase cultures were in turn prepared by diluting stationary phase cultures until an OD 600 of 0.2 in fresh broth and allowing further incubation at 37°C until an OD600 of 0.6 was reached. Then, 10^9 CFUs of *L. paracasei CNCM I-5220* resuspended in 200 μl of vehicle (Sodium carbonate) were administered to mice by gavage every other day for three weeks. Bacteria were freshly prepared for each administration. Mice were sacrificed 24 hours after last *E. coli C17* or *L. paracasei CNCM I-5220* administration.

### Knock-out of Virf1 from *E. coli* C17 and *in vivo* experiment

*Escherichia coli* C-17 virF::cat mutant derivative was constructed through homologous recombination using the λ-red technique (Datsenko and Wanner, 2000). The protein products of the red genes (Gam, Exo, and Beta) enable the highly efficient recombination of PCR products (containing a chloramphenicol resistance cassette) flanked by short (50 bp) nucleotide sequences, homologous to the target sequence. Thus, PCR product for mutant construction were obtained amplifying the resistance cassette flanked by the FRT sequences from the genome of the *E. coli* MG1655ΔcysH strain (Rossi et al., 2014) using primer pairs *virF1*_FRT_for and *virF1*_FRT_rv (Table S5). Correct insertion of the resistance cassette in the *virF* gene was verified through PCR using the *virF1*_for and *virF1*_rev primers pair (Figure S6F and Table S5). For *in vivo* experiment of GVB impairment and liver bacteria translocation 10^7 CFUs of *E. coli* C17 or *E. coli C17* Δ*virf1* resuspended in 200 μl of vehicle (Sodium carbonate) were fed to mice by gavage every day for two days. Mice were sacrificed 24 hours after last *E. coli* C17 or *E. coli C17* Δ*virf1* administration.

## QUANTIFICATION AND STATISTICAL ANALYSIS

For the human part, we used receiver-operating-characteristic (ROC) curves to determine the best cut-off value of PV-1. The area under the ROC curve (AUC) was calculated, and the threshold that was optimal (in terms of sensitivity and specificity for the identification of recurrent patients) in the training cohort was determined on the basis of the Youden Index that maximizes the sum of sensitivity and specificity. The Youden index not only summaries the discriminatory accuracy of a diagnostic test but also provides a ready-to-use optimal cut-off point for the purpose of future diagnosis. This index is defined (Luo and Xiong, 2013) as J(t) = Se(t) + Sp(t) −1, essentially a combinatory index of sensitivity and specificity at a cut-point t, that in our case is 65%. Chi-square tests were used to calculate the associations between PV-1 and clinical prognostic factors. Overall survival (OS) was calculated from diagnosis to death (event) or last follow-up (censored). Progression free survival (PFS) was calculated from diagnosis to appearance of relapse, or death or last follow-up (censored). Disease free survival (DFS) was calculated from diagnosis to appearance of relapse or other second cancer or death or last follow-up (censored). For the validation cohort, we could not calculate DFS and OS because none of the enrolled patients in the validation cohort developed other types of tumors and they are all alive, except for two of

Please cite this article in press as: Bertocchi et al., Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver, Cancer Cell (2021), https://doi.org/10.1016/j.ccell.2021.03.004



them. Once stratified based on PV-1 detection, patient's subsets were compared for survival outcomes, using Kaplan-Meier survival curves and multivariable analyses based on the Cox proportional hazards methods. Differences in Kaplan-Meier curves were tested for statistical significance using the long-rank test. Hazard ratios with 95% confidence intervals were estimated with multivariable Cox regression models adjusted for significant prognostic factors and confounding factors. All analyses were performed with R and SAS software, version 9.2 (SAS Institute).

For murine part, the sample size for animal studies were guided by previous murine studies in our laboratory. Results are represented as mean ± standard error mean (SEM) or using box plots showing the interquartile range; the horizontal lines show the median values and the whiskers indicate the minimum-to-maximum range. Statistical significance, using GraphPad Prism software, was evaluated with: Mann-Whitney two-sided unpaired t-test (for comparison between two groups without Gaussian distribution), two-sided Unpaired t-test (for comparison between two groups with Gaussian distribution), Kruskal-Wallis test (for comparison between more than two groups without Gaussian distribution) and One-Way ANOVA test (for comparison between more than two groups with Gaussian distribution) followed by Bonferroni's or Dunn's post-test using GraphPad Prism software. Gaussian distribution was evaluated using Column Statistic (D'Agostino & Pearson normality test) in GraphPad Prism software. Following combined experiments, outliers were detected with the Grubbs' test and excluded from the analysis. Data display normal variance. A probability value of * $P<0.05$ was considered to be significant. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessments.