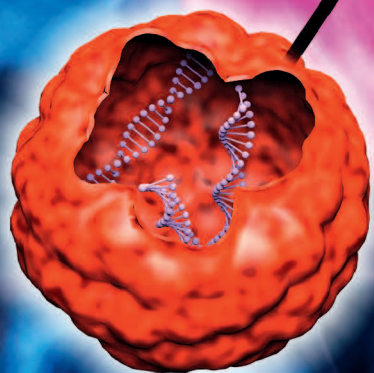
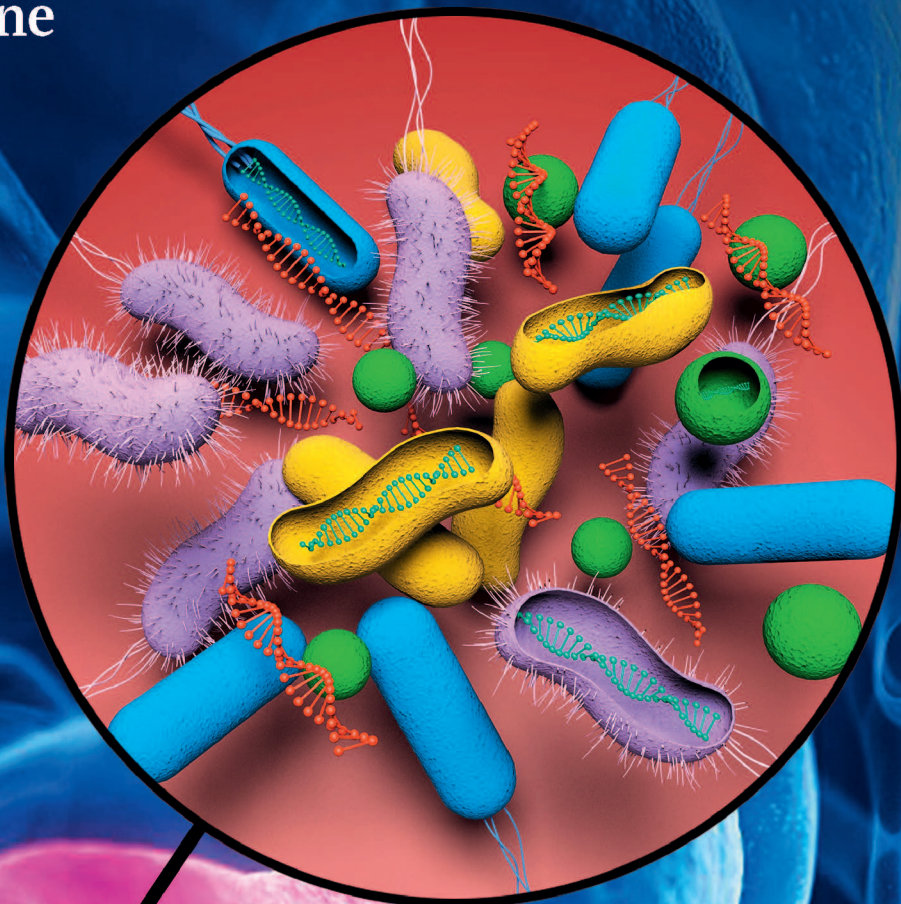


# Microbiome and Molecular Characterization of Neoplasms in Pancreas and Intestine



Shan Li

李山

# **Microbiome and Molecular Characterization of Neoplasms in Pancreas and Intestine**

**Shan Li**

The studies presented in this thesis were performed in the Department of Gastroenterology and Hepatology, Erasmus MC–University Medical Center, Rotterdam, the Netherlands.

The research was funded by:  
Erasmus MC Grant  
China Scholarship Council (CSC)

Financial support for the printing of this thesis was kindly provided by:  
Erasmus University Rotterdam

Copyright © 2020 by Shan Li.

For articles that are not published open access, the copyright has been transferred to the respective publisher. No part of this thesis may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, without written permission from the author or when appropriate from the publisher.

Cover design: the author of this thesis & LetPub ([www.letpub.com](http://www.letpub.com))

Cover layout: Ridderprint ([www.ridderprint.nl](http://www.ridderprint.nl))

Printed by: Ridderprint ([www.ridderprint.nl](http://www.ridderprint.nl))

ISBN: 978-94-6375-987-8

# **Microbiome and Molecular Characterization of Neoplasms in Pancreas and Intestine**

Microbioom en moleculaire karakterisering  
van neoplasma's in de pancreas en darm

## **Thesis**

**to obtain the degree of Doctor from the  
Erasmus University Rotterdam  
by command of the  
rector magnificus**

Prof.dr. R.C.M.E. Engels

**and in accordance with the decision of the Doctorate Board.**

**The public defence shall be held on**

*Thursday 27<sup>th</sup> August 2020 at 15:30*

by

**Shan Li**

born in Kaizhou, Chongqing, China

**Erasmus University Rotterdam**





## **Doctoral Committee:**

### **Promotor:**

Prof.dr. M.P. Peppelenbosch

### **Other Members:**

Prof.dr. M.J. Bruno

Prof.dr. R. Fodde

Prof.dr. I.D. Nagtegaal

### **Copromotor:**

Dr. M.J.M. Smits

# Contents

Chapter 1	General introduction and outline of this thesis	1
Chapter 2	Pancreatic cyst fluid harbors a unique microbiome Shan Li, Gwenny M. Fuhler, Nahush BN, Tony Jose, Marco J. Bruno, Maikel P. Peppelenbosch, Sergey R. Konstantinov <b>Microbiome.</b> 2017, 5:147.	15
Chapter 3	Bacterial biofilms in colorectal cancer initiation and progression Shan Li, Sergey R. Konstantinov, Ron Smits, Maikel P. Peppelenbosch <b>Trends in Molecular Medicine.</b> 2017, 23(1):18-30.	53
Chapter 4	Bacterial biofilms as a potential contributor to mucinous colorectal cancer formation Shan Li, Maikel P. Peppelenbosch, Ron Smits <b>BBA-Reviews on Cancer.</b> 2019, 1872(1):74-79.	79
Chapter 5	Unravelling posttranslational modifications of ATOH1 driving colorectal cancer towards mucinous differentiation Shan Li, Werner Helvensteijn, Raymond Poot, Jeroen Demmers, Wenhui Wang, Zhijiang Miao, Menggang Liu, Bingting Yu, Marla Lavrijsen, Buyun Ma, Shanshan Li, Pengyu Liu, Maikel P. Peppelenbosch, Ron Smits <b>Manuscript in Preparation.</b>	97
Chapter 6	Commonly observed <i>RNF43</i> mutations retain functionality in attenuating Wnt/ $\beta$ -catenin signaling and unlikely confer Wnt-dependency onto colorectal cancers Shan Li, Marla Lavrijsen, Aron Bakker, Marcin Magierowski, Katarzyna Magierowska, Pengyu Liu, Wenhui Wang, Maikel P. Peppelenbosch, Ron Smits <b>Oncogene.</b> 2020, 39:3458-3472.	145
Chapter 7	Summary	195
Chapter 8	Discussion	203
Chapter 9	Dutch summary (Nederlandse samenvatting)	213
Appendices		221
	List of Publications	223
	PhD Portfolio	225
	Curriculum Vitae	227
	Acknowledgements	229



# **Chapter 1**

## **General Introduction and Outline of This Thesis**





The first chapter gives a brief background and context of pancreatic cystic neoplasms, colorectal cancer and its mucinous subtype, and introduces the oncogenic Wnt/ $\beta$ -catenin signaling pathway, cancer-associated microbiome, and cancer-promoting bacterial biofilm. The aims and outline of this thesis are provided at the end of this chapter.

## **Pancreatic cystic neoplasms and their premalignant potential**

Pancreatic cystic neoplasms, also known as pancreatic cysts, are fluid-filled lesions lined by neoplastic or non-neoplastic epithelium and located in the pancreas. An unexpectedly high prevalence (12.9%) of the pancreatic cysts can be observed in the general population (1). In most cases, these cystic pancreatic lesions are asymptomatic and, therefore, do not require any surgical intervention or specific treatment (1,2). However, it is noteworthy that a minor, but not to be overlooked subset of the cystic lesions is considered a precancerous condition that has the potential to develop into pancreatic cancer (1,2). The most common premalignant cyst is the mucinous subtype, including intraductal papillary mucinous neoplasm (IPMN) and mucinous cystic neoplasm (MCN). In contrast, pseudocyst and the serous subgroup such as serous cystadenoma (SCA) generally behave as benign lesions (2). Despite the premalignant potential of pancreatic mucinous neoplasms, patients with such cystic neoplasms usually need only observation (2). To date, decision-making prior to surgery for the precancerous cystic lesions mainly depends on endoscopic ultrasound-guided fine-needle aspiration (EUS-FNA) for cytology and biochemical analysis of cyst fluid, to determine e.g. levels of carcinoembryonic antigen (CEA) and carbohydrate antigen 19-9 (CA19-9) (2). However, these strategies still lack sufficient predictive strength for accurately determining the therapeutic plans (2). An unmet need therefore exists to identify alternative markers that are present in the EUS-FNA aspirated fluids. Intriguingly, intestinal microbiota has emerged as an important contributor to neoplastic lesions in the digestive system and other organs over the last decades (3–5). It is however not yet identified whether gastrointestinal microbiota is present in the pancreatic cyst fluid. Performing investigations into linking the gut microbiome

to the pancreatic cystic lesions will open a new avenue for better understanding of pancreatic cystogenesis and carcinogenesis, and may thus aid in the decision-making for lesion management.

### **Colorectal cancer and its mucinous subtype**

Colorectal cancer (CRC) is the third most prevalent cancer (10.2% of the total cases), and is the second leading cause (9.2%) of cancer-related death in the world (6). Underlying molecular mechanisms for colorectal carcinogenesis have been extensively investigated. A large aggregation of genetic alterations and epigenetic changes contribute to colon tumorigenesis and development (7–9). Importantly, accumulated evidence supports that aberrant activation of the Wnt/ $\beta$ -catenin signaling pathway is a hallmark of the great majority of CRCs, owing to mutations in components of this pathway (9–16).

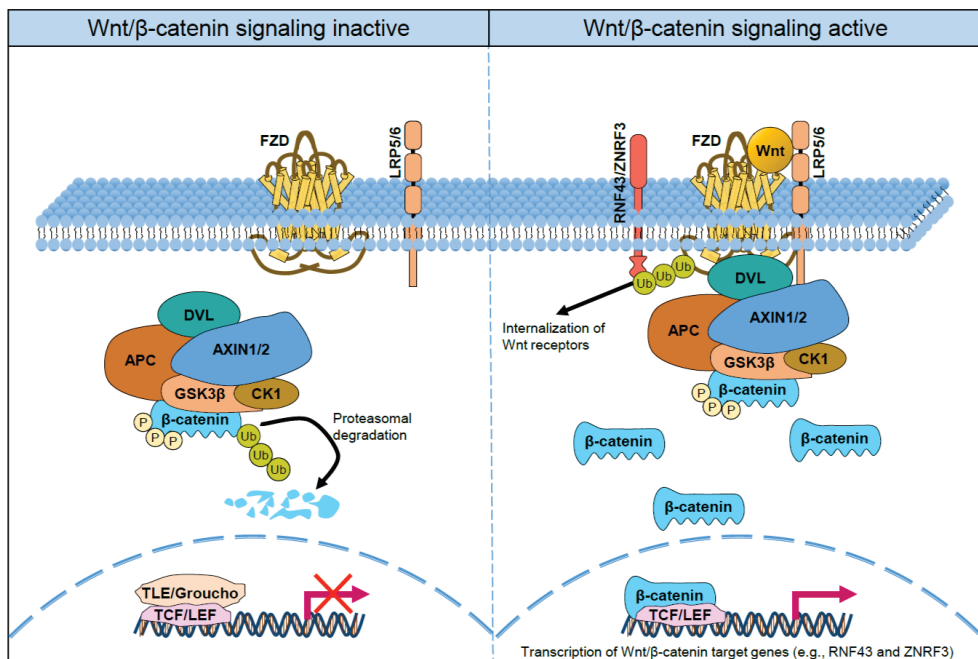
Mucinous CRC (MCC) is an important histological subtype of CRC, which is arbitrarily defined as a tumor composed of more than 50% extracellular mucin pools, accounting for 10–15% of all the CRCs (17,18). This CRC subtype with such prominent mucinous component is positively associated with more frequent multiple metastases, worse survival, and resistance to chemoradiotherapy as compared with non-mucinous CRC (17,19). In fact, the mucinous component within a tumor mass *per se* does not seem to be a detrimental phenotype, as it is a commonly observed feature of well-differentiated lesions (18). Of note, intestinal mucus has tumor suppressive effects under homeostatic conditions. It essentially functions as a viscous lubricant and epithelia-coating shield against pathogen colonization, metabolic insults, and noxious luminal contents, thereby effectively protecting colonic mucosa from constantly and abundantly present environmental challenges (20,21). However, the mucus existing in MCC serves as a barrier that limits chemotherapeutic agents and immune cells penetrating into the solid tumor mass. Underlying mechanisms through which the prominent mucin-producing differentiation is regulated in this CRC subgroup remain unclear. Because of lack of knowledge of this subgroup, it is more difficult to identify the best treatment options to improve outcome.

Importantly, it is now well known that the transcription factor ATOH1 is a master regulator in driving intestinal progenitors towards secretory cell lineages, including mucin-producing goblet, enteroendocrine, and Paneth cells, for maintaining gut homeostasis (22). As a good example, transgenic mouse models that intestinal-specific lack functional activity of ATOH1, exhibited not merely a significant reduction in mucus secretion and abnormality of the mucin-producing goblet cells, but also showed a high predisposition for developing tumors in the intestine (23–25). Moreover, several lines of evidence suggest that ATOH1 (in)directly regulates expression of *MUC2*, encoding a well-known secreted mucin that makes up most of the mucus in MCC (26,27). Thus, understanding of ATOH1-regulated mucin-producing differentiation during tumor formation in the colon may better explain MCC pathogenesis.

### **The key role of the Wnt/ $\beta$ -catenin signaling pathway in pancreatic and colorectal neoplasms**

The crucial regulatory machinery of the Wnt/ $\beta$ -catenin pathway is the destruction complex that targets  $\beta$ -catenin for degradation (14). This complex is mainly comprised of adenomatous polyposis coli (APC), glycogen synthase kinase-3 $\beta$  (GSK-3 $\beta$ ), casein kinase-1 (CK1), and the scaffold proteins AXIN1/AXIN2 (14) (**Figure 1**). Aberrant activation of this pathway is more commonly observed in colorectal than in pancreatic cancers. Inactivating mutations in the tumor suppressor gene *APC* are thought to be an early event that drives tumor initiation in the vast majority (~80%) of CRCs (9–13). Activating mutations of *CTNNB1* encoding the  $\beta$ -catenin protein itself or inactivating mutations in genes such as *AXIN1* and *AXIN2* also take part in such  $\beta$ -catenin-dependent carcinogenesis (10–13,15). Mutation of the *RNF43* gene is also frequently observed in CRC. It encodes an upstream regulator of Wnt/ $\beta$ -catenin signaling that negatively regulates Wnt receptors, which, interestingly is also frequently mutated in pancreatic and other neoplasms (10,16,28–36). Importantly, *RNF43* mutations are believed to confer Wnt-dependency to colorectal and pancreatic cancers (37–39). These *RNF43*-mutant tumors rely on Wnt ligand-mediated activation of the  $\beta$ -catenin

pathway for their proliferation. Inhibitors targeting Wnt secretion have therefore been deemed a promising candidate in the treatment of *RNF43* mutated tumors (37–39). However, there has not been a comprehensive functional analysis of the *RNF43* inactivating mutations to evaluate the extent to which these mutations affect oncogenic  $\beta$ -catenin signaling and the responsiveness of such mutant tumors to these Wnt secretion inhibitors.



**Figure 1.** Overview of Wnt/ $\beta$ -catenin signaling pathway. **Left, Wnt/ $\beta$ -catenin signaling inactive state.** In the absence of Wnt signals,  $\beta$ -catenin is phosphorylated by the destruction complex, which is mainly comprised of APC, GSK3 $\beta$ , CK1, and AXIN1/AXIN2. Phosphorylated  $\beta$ -catenin is targeted for proteasomal degradation after ubiquitination. In this manner TCF/LEF-mediated transcriptional activation is repressed by a complex formed by TCF/LEF itself and transcriptional repressors such as TLE/Groucho in the nucleus. **Right, Wnt/ $\beta$ -catenin signaling active state.** Upon Wnt ligands binding to the Wnt receptor complex that contains Frizzled (FZD) and LRP5/6, Dishevelled (DVL) is activated, which transduces sequential signaling cascades within the cytoplasm, eventually disrupting  $\beta$ -catenin degradation. This causes cytoplasmic accumulation and nuclear translocation of  $\beta$ -catenin, thereby activating TCF/LEF-mediated transcription of target genes such as *RNF43* and *ZNRF3*. Both *RNF43* and its homologue *ZNRF3* interact with and ubiquitinate FZD and LRP proteins, leading to internalization of the cell surface Wnt receptors. As such, the two E3 ubiquitin ligases constitute a negative feedback regulation of  $\beta$ -catenin transactivation. This figure, in part, was created with modified objects derived from the ScienceSlides ([www.scienceslides.com](http://www.scienceslides.com)).

## Gut microbiome and cancer

It is estimated that up to 15% of human malignancies are attributable to microbes that inhabit the gastrointestinal tract (40–42). *Helicobacter pylori* infection is one of the best examples of bacterial involvement in inflammation and its association with cancer (43,44). However, different from the *H. pylori*-induced carcinogenesis in the stomach, inflammation and cancers in other organs such as the large intestine are not associated with pathogenic effects of a single specific pathogen, but are linked to microbe-host-microbe interactions among a multispecies microbial community (41,45–51). It has been well accepted that an interplay of bacterial dysbiosis (alteration in abundance and diversity of microbiota), persistent inflammation, and epithelial barrier defects can orchestrate cancer development (52) (**Figure 2**).


Gut microbiota controls the balance between homeostasis and disease status. Commensal microbiota not only leads to enhanced gut immunity against invasive bacterial infections, but also in some cases behaves as a pathogen that induces an inflammatory response, which is a common hallmark of cancers (53). Under homeostatic conditions, commensal microbes have crucial roles in providing protection for the host intestine or other organs from pathogen colonization. The ‘healthy’ microorganisms are capable of mitigating or even avoiding tissue injury (52,54). These symbiotic bacteria promote production of short-chain fatty acids (SCFAs) such as butyrate, and thereby enhance the intestinal immunity and confer host defenses to epithelium against pathogen colonization (52,55,56). Interestingly, butyrate also inhibits tumor cell proliferation and induces apoptosis, highlighting its potent anti-cancer effects (57). Moreover, beneficial microbiota could alleviate cancer-eliciting inflammation by regulating the host inflammatory response. For instance, colonization of *Lactobacillus murinus* in mice resulted in elevated levels of regulatory T (Treg) cells, reduced inflammatory infiltration, and decreased expression of proinflammatory cytokines, thereby reducing the incidence of dextran sodium sulfate (DSS)-induced colitis (58). However, it is interesting to note that ‘normal’ microbiota may occasionally act as pathobionts, because some commensal bacteria also have the potential to provoke a pro-oncogenic type 17 T helper cell (Th17)-



mediated host inflammatory response. For instance, epithelial colonization by adhesive, but not non-adhesive, segmented filamentous bacterium (SFB) that is a commensal bacterium, specifically induced the development of Th17 cells (59,60).

It is noteworthy that bacteria that only provoke inflammation but not tissue impairment are insufficient for tumor development. In contrast, bacteria that induce both tissue injury and inflammation may have the potential to drive carcinogenesis (7). Tissue injury caused by some invasive pathogens appears to be an early event for tumorigenesis. The enterotoxigenic *Bacteroides fragilis* (ETBF) that promotes colon tumorigenesis is a good example for supporting such notions (61,62). This bacterium generates a *B. fragilis* toxin (BFT, also known as fragilysin) that induces tissue damage and prolonged colitis, and concurrence of IL-17 generation and activation of Stat-3 proliferation signaling, contributing to colon tumorigenesis (61,63). More importantly, the loss of epithelial barrier induced by the ‘driver pathogen’ ETBF leads to bacterial dysbiosis and thereby favors subsequent colonization of other ‘passenger’ opportunistic pathogens onto the epithelium, thus enhancing the host pro-inflammatory response and bacterial pro-oncogenic activity (64).

---

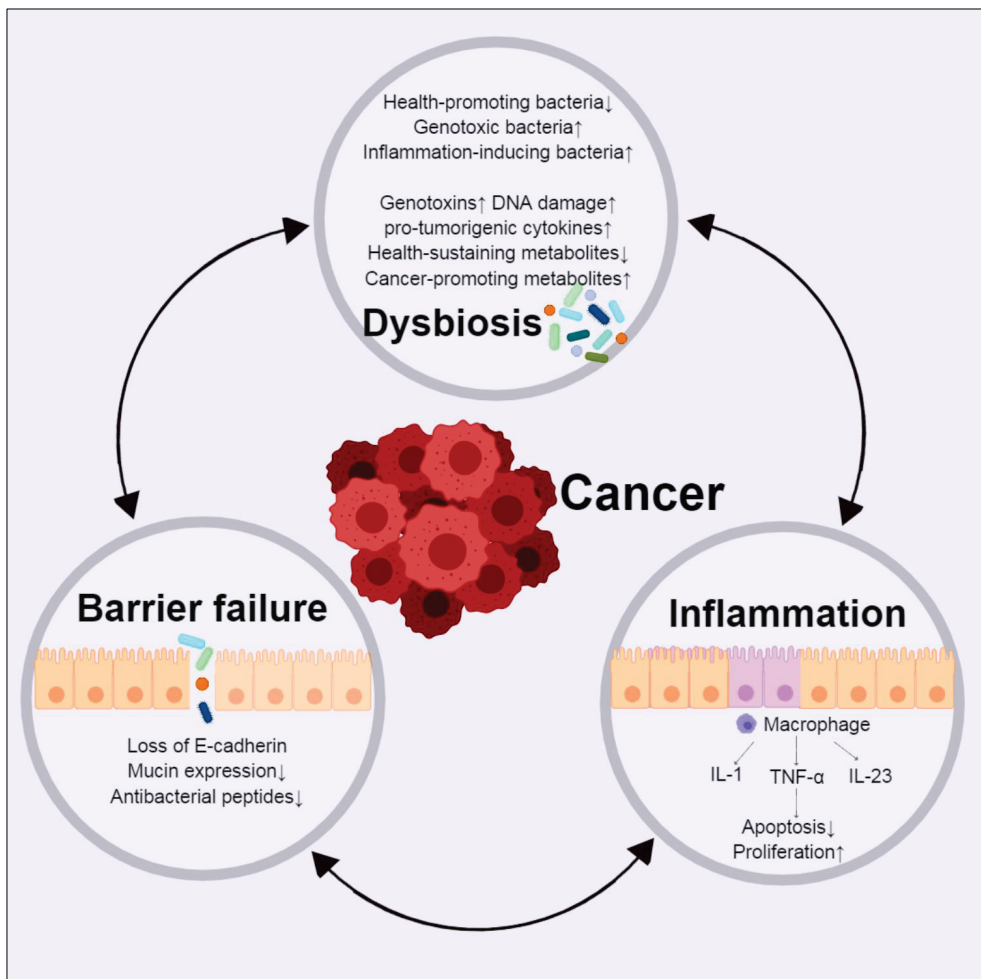


**Figure 2.** The mechanisms by which host–microbiota interplays influence bacteria-induced carcinogenesis. Microbial oncogenesis is highly associated with bacterial dysbiosis, epithelial barrier dysfunction, and inflammatory status, which are mutually reinforcing. Under homeostatic condition, host tissue regeneration and immunity, and commensal microbiota cooperatively maintain host-microbiota symbiosis. Defects in epithelial barrier, such as compromised tissue integrity (e.g., E-cadherin loss), decreased mucin expression, and reduced production of antibacterial peptides, tip the balance towards bacterial dysbiosis. In this way, the composition and diversity of microbiota are changed, resulting in an increase in abundance of pathogens. These pathogenic bacteria produce genotoxins and cancer-causing metabolites, thereby promoting carcinogenesis. Moreover, microbial metabolites and bacterial components such as lipopolysaccharide (LPS) induce activation of macrophages, promoting the release of IL-1, TNF- $\alpha$ , and IL-23 into the tumor microenvironment. Secretion of these pro-inflammatory and pro-carcinogenic cytokines is capable of leading to enhanced dysbiosis and a more severe disruption of the barrier. This figure, adapted from Schwabe et al. (52), was created with Biorender (<https://biorender.com/>).

(see figure on next page)

## Cancer-promoting bacterial biofilm

Bacterial biofilm is a higher-order spatial organization of multispecies microbial aggregates, encompassed by a polymeric matrix including polysaccharides, proteins, and other components. Biofilm is not only a microbial block mass but also a structurally and dynamically complex biological system, which usually exhibits exacerbated pathogenic effects to the host (65–68). Previously, the formation of bacterial biofilm was considered to be one of the most important ways for pathogens to acquire their antibiotic-resistant phenotype in infectious diseases and/or bacteria-



**Figure 2** (see legend on previous page)

induced inflammation. Such spatial organization of colonic microbial communities has recently been associated with CRC tumorigenesis (51,69–71). The multispecies biofilm formed by some invasive bacteria such as *Fusobacteria* may augment the pro-inflammatory and pro-oncogenic effects of the gut microbiome in CRC initiation and progression (51,69,70). Bacterial biofilms enhance the microbe-host-microbe interplay that contributes to CRC development through promoting an inflammatory response, disrupting the intestinal epithelial barrier, and generating bacterial polyamine metabolites (51,70).

Interestingly, there exists a substantial overlapping phenotype between biofilm-associated colonic tumors and mucinous CRC. Accumulating evidence suggests that MCC predominantly arises in the right-sided colon, which appears also to be the case for tumor-associated bacterial biofilms (18,72,73). Moreover, MCC well correlates with microsatellite instability (MSI) and CpG island methylator phenotype (CIMP) status, which are frequently associated with increased intestinal inflammation (18,72–74). The generation of cytokines, such as TNF- $\alpha$ , IL-22 and others, enhanced by the inflammatory response, can lead to elevated mucus production in CRC cells (75–82). Given the preferential occurrence of bacterial biofilm in the right-sided colon and its involvement in exacerbated inflammatory response, there is a probable link between bacterial biofilm and MCC.

### **Aim of this thesis**

As noted above, the general aims of this thesis are: 1) to investigate the microbial profiles of pancreatic cyst fluid (PCF), and to explore the potential link of the PCF microbiome to pancreatic cystogenesis and tumorigenesis. 2) to provide deep insights into the relationship between gut microbiota, bacterial biofilm, carcinogenesis, and molecular features of mucinous differentiation during tumor formation in colon. 3) to molecularly discover the modulations of ATOH1-regulated mucinous differentiation. 4) to study the functions of commonly observed *RNF43* inactivating mutations contributing to activation of oncogenic Wnt/ $\beta$ -catenin signaling.

---

**Outline of this thesis**

- Chapter 1.** I have given a brief genetic background and microbial context of neoplasms in the pancreas and colorectum.
- Chapter 2.** We show that pancreatic cyst fluid is not sterile. This specific niche harbors a unique microbiome that is different from other organs. Putative pro-oncogenic pathogens are present among the PCF microbiome, which may contribute to neoplastic progression in the pancreas.
- Chapter 3.** We review the impacts of the cancer-eliciting microbiome on the host organs. We also provide insights into links between the bacterial biofilm and colorectal cancer, and discuss the underlying mechanisms of bacterial biofilm contributing to colorectal carcinogenesis.
- Chapter 4.** We review already-known molecular characterizations of mucinous colorectal cancer, and discuss the potential relationship between bacterial biofilm and the mucinous phenotype of colorectal cancer.
- Chapter 5.** We investigate the modulations of posttranslational modifications of ATOH1 contributing to mucinous differentiation in colorectal cancer and thereby demonstrate a novel mechanism through which ATOH1 is stabilized.
- Chapter 6.** We perform a comprehensive mutational analysis of the *RNF43* gene in various cancers, and we carry out biochemical and functional analyses of *RNF43* inactivating mutations in regulating  $\beta$ -catenin signaling in colorectal cancer cells. We show that C-terminal inactivating mutations of *RNF43* retain function that inactivates Wnt/ $\beta$ -catenin signaling and thus fail to induce  $\beta$ -catenin-dependent carcinogenesis.
- Chapter 7 & 8.** I summarize and discuss the findings from this thesis, in which the integration of microbiome and molecular characterization of neoplasms in the pancreas and large intestine will provide a new paradigm for cancer development.

## References

1. M. L. Kromrey *et al.*, Prospective study on the incidence, prevalence and 5-year pancreatic-related mortality of pancreatic cysts in a population-based study. *Gut* **67**, 138-145 (2018).
2. A. Stark, T. R. Donahue, H. A. Reber, O. J. Hines, Pancreatic cyst disease: a review. *JAMA* **315**, 1882-1893 (2016).
3. S. Yachida *et al.*, Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nat. Med.* **25**, 968-976 (2019).
4. W. S. Garrett, Cancer and the microbiota. *Science* **348**, 80-86 (2015).
5. E. Elinav, W. S. Garrett, G. Trinchieri, J. Wargo, The cancer microbiome. *Nat. Rev. Cancer* **19**, 371-376 (2019).
6. F. Bray *et al.*, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**, 394-424 (2018).
7. J. Terzic, S. Grivennikov, E. Karin, M. Karin, Inflammation and colon cancer. *Gastroenterology* **138**, 2101-2114 (2010).
8. E. R. Fearon, Molecular genetics of colorectal cancer. *Annu. Rev. Pathol.-Mech. Dis.* **6**, 479-507 (2011).
9. J. M. Carethers, B. H. Jung, Genetics and genetic biomarkers in sporadic colorectal cancer. *Gastroenterology* **149**, 1177-1190 (2015).
10. N. The Cancer Genome Atlas, Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330 (2012).
11. C. Albuquerque, E. R. Bakker, W. van Veelen, R. Smits, Colorectal cancers choosing sides. *BBA-Rev. Cancer* **1816**, 219-231 (2011).
12. S. J. Leedham *et al.*, A basal gradient of Wnt and stem-cell number influences regional tumour distribution in human and mouse intestinal tracts. *Gut* **62**, 83-93 (2013).
13. M. Christie *et al.*, Different APC genotypes in proximal and distal sporadic colorectal cancers suggest distinct WNT/  $\beta$ -catenin signalling thresholds for tumourigenesis. *Oncogene* **32**, 4675-4682 (2013).
14. V. S. Li *et al.*, Wnt signaling through inhibition of  $\beta$ -catenin degradation in an intact Axin1 complex. *Cell* **149**, 1245-1256 (2012).
15. C. Albuquerque *et al.*, Colorectal cancers show distinct mutation spectra in members of the canonical WNT signaling pathway according to their anatomical location and type of genetic instability. *Genes Chromosomes Cancer* **49**, 746-759 (2010).
16. M. Giannakis *et al.*, *RNF43* is frequently mutated in colorectal and endometrial cancers. *Nat. Genet.* **46**, 1264-1266 (2014).
17. J. Verhulst, L. Ferdinande, P. Demetter, W. Ceelen, Mucinous subtype as prognostic factor in colorectal cancer: a systematic review and meta-analysis. *J. Clin. Pathol.* **65**, 381-388 (2012).
18. N. Hugen, G. Brown, R. Glynne-Jones, J. H. de Wilt, I. D. Nagtegaal, Advances in the care of patients with mucinous colorectal cancer. *Nat. Rev. Clin. Oncol.* **13**, 361-369 (2016).
19. L. J. Mekenkamp *et al.*, Mucinous adenocarcinomas: poor prognosis in metastatic colorectal cancer. *Eur. J. Cancer* **48**, 501-509 (2012).
20. Y. S. Kim, S. B. Ho, Intestinal goblet cells and mucins in health and disease: recent insights and progress. *Curr. Gastroenterol. Rep.* **12**, 319-330 (2010).
21. M. E. Johansson *et al.*, The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 15064-15069 (2008).
22. Q. Yang, N. A. Bermingham, M. J. Finegold, H. Y. Zoghbi, Requirement of Math1 for secretory cell lineage commitment in the mouse intestine. *Science* **294**, 2155-2158 (2001).
23. K. Yang *et al.*, Interaction of *Muc2* and *Apc* on Wnt signaling and in intestinal tumorigenesis: potential role of chronic inflammation. *Cancer Res.* **68**, 7313-7322 (2008).
24. A. Kazanjian, T. Noah, D. Brown, J. Burkart, N. F. Shroyer, Atonal homolog 1 is required for growth and differentiation effects of Notch/ $\gamma$ -secretase inhibitors on normal and cancerous intestinal epithelial cells. *Gastroenterology*, 918-928 (2010).
25. W. Bossuyt *et al.*, Atonal homolog 1 is a tumor suppressor gene. *PLoS Biol.* **7**, e39 (2009).
26. G. Li *et al.*, SCF/c-KIT signaling promotes mucus secretion of colonic goblet cells and development of mucinous colorectal adenocarcinoma. *Am. J. Cancer Res.* **8**, 1064-1073 (2018).
27. P. Shen *et al.*, SCF/c-KIT signaling increased Mucin2 production by maintaining Atoh1 expression in mucinous colorectal adenocarcinoma. *Int. J. Mol. Sci.* **19**, 1541 (2018).
28. G. L. Ryland *et al.*, Mutational landscape of mucinous ovarian carcinoma and its neoplastic precursors. *Genome Med.* **7**, 87 (2015).



29. G. L. Ryland *et al.*, *RNF43* is a tumour suppressor gene mutated in mucinous tumours of the ovary. *J. Pathol.* **229**, 469-476 (2013).
30. K. Wang *et al.*, Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer. *Nat. Genet.* **46**, 573-582 (2014).
31. Y. Jiao *et al.*, Whole-exome sequencing of pancreatic neoplasms with acinar differentiation. *J. Pathol.* **232**, 428-435 (2014).
32. N. Waddell *et al.*, Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature* **518**, 495-501 (2015).
33. Y. Zou *et al.*, *RNF43* mutations are recurrent in Chinese patients with mucinous ovarian carcinoma but absent in other subtypes of ovarian cancer. *Gene* **531**, 112-116 (2013).
34. T. Eto *et al.*, Impact of loss-of-function mutations at the *RNF43* locus on colorectal cancer development and progression. *J. Pathol.* **245**, 445-455 (2018).
35. H. H. N. Yan *et al.*, *RNF43* germline and somatic mutation in serrated neoplasia pathway and its association with BRAF mutation. *Gut* **66**, 1645-1656 (2017).
36. B. H. Min *et al.*, Dysregulated Wnt signalling and recurrent mutations of the tumour suppressor *RNF43* in early gastric carcinogenesis. *J. Pathol.* **240**, 304-314 (2016).
37. Z. Steinhart *et al.*, Genome-wide CRISPR screens reveal a Wnt-FZD5 signaling circuit as a druggable vulnerability of *RNF43*-mutant pancreatic tumors. *Nat. Med.* **23**, 60-68 (2017).
38. X. Jiang *et al.*, Inactivating mutations of *RNF43* confer Wnt dependency in pancreatic ductal adenocarcinoma. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 12649-12654 (2013).
39. B. K. Koo, J. H. van Es, M. van den Born, H. Clevers, Porcupine inhibitor suppresses paracrine Wnt-driven growth of *Rnf43;Znrf3*-mutant neoplasia. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7548-7550 (2015).
40. D. M. Parkin, The global health burden of infection-associated cancers in the year 2002. *Int. J. Cancer* **118**, 3030-3044 (2006).
41. A. D. Kostic *et al.*, Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome Res.* **22**, 292-298 (2012).
42. V. Narayanan, M. P. Peppelenbosch, S. R. Konstantinov, Human fecal microbiome-based biomarkers for colorectal cancer. *Cancer Prev. Res. (Phila.)* **7**, 1108-1111 (2014).
43. J. G. Kusters, A. H. van Vliet, E. J. Kuipers, Pathogenesis of *Helicobacter pylori* infection. *Clin. Microbiol. Rev.* **19**, 449-490 (2006).
44. B. J. Marshall, J. R. Warren, Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet* **1**, 1311-1315 (1984).
45. A. Machida-Montani *et al.*, Atrophic gastritis, *Helicobacter pylori*, and colorectal cancer risk: a case-control study. *Helicobacter* **12**, 328-332 (2007).
46. M. Castellarin *et al.*, *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome Res.* **22**, 299-306 (2012).
47. J. C. Arthur *et al.*, Microbial genomic analysis reveals the essential role of inflammation in bacteria-induced colorectal cancer. *Nat. Commun.* **5**, 4724 (2014).
48. A. D. Kostic, E. Chun, M. Meyerson, W. S. Garrett, Microbes and inflammation in colorectal cancer. *Cancer Immunol. Res.* **1**, 150-157 (2013).
49. A. D. Kostic *et al.*, *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe* **14**, 207-215 (2013).
50. M. R. Rubinstein *et al.*, *Fusobacterium nucleatum* promotes colorectal carcinogenesis by modulating E-cadherin/ $\beta$ -catenin signaling via its FadA adhesin. *Cell Host Microbe* **14**, 195-206 (2013).
51. C. M. Dejea *et al.*, Microbiota organization is a distinct feature of proximal colorectal cancers. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18321-18326 (2014).
52. R. F. Schwabe, C. Jobin, The microbiome and cancer. *Nat. Rev. Cancer* **13**, 800-812 (2013).
53. D. Hanahan, R. A. Weinberg, Hallmarks of cancer: the next generation. *Cell* **144**, 646-674 (2011).
54. N. Gagliani, B. Hu, S. Huber, E. Elinav, R. A. Flavell, The fire within: microbes inflame tumors. *Cell* **157**, 776-783 (2014).
55. K. M. Maslowski *et al.*, Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. *Nature* **461**, 1282-1286 (2009).
56. M. H. Kim, S. G. Kang, J. H. Park, M. Yanagisawa, C. H. Kim, Short-chain fatty acids activate GPR41 and GPR43 on intestinal epithelial cells to promote inflammatory responses in mice. *Gastroenterology* **145**, 396-406 (2013).
57. D. E. Serban, Gastrointestinal cancers: influence of gut microbiota, probiotics and prebiotics. *Cancer Lett.* **345**, 258-270 (2014).

58. C. Tang *et al.*, Inhibition of Dectin-1 signaling ameliorates colitis by inducing *Lactobacillus*-mediated regulatory T cell expansion in the intestine. *Cell Host Microbe* **18**, 183-197 (2015).
59. K. Atarashi *et al.*, Th17 cell induction by adhesion of microbes to intestinal epithelial cells. *Cell* **163**, 367-380 (2015).
60. I. I. Ivanov *et al.*, Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell* **139**, 485-498 (2009).
61. S. Wu *et al.*, A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nat. Med.* **15**, 1016-1022 (2009).
62. A. Boleij *et al.*, The *Bacteroides fragilis* toxin gene is prevalent in the colon mucosa of colorectal cancer patients. *Clin. Infect. Dis.* **60**, 208-215 (2015).
63. E. C. Wick *et al.*, Stat3 activation in murine colitis induced by enterotoxigenic *Bacteroides fragilis*. *Inflamm. Bowel Dis.* **20**, 821-834 (2014).
64. H. Tjalsma, A. Boleij, J. R. Marchesi, B. E. Dutilh, A bacterial driver-passenger model for colorectal cancer: beyond the usual suspects. *Nat. Rev. Microbiol.* **10**, 575-582 (2012).
65. P. K. Singh *et al.*, Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms. *Nature* **407**, 762-764 (2000).
66. A. Swidsinski, J. Weber, V. Loening-Baucke, L. P. Hale, H. Lochs, Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. *J. Clin. Microbiol.* **43**, 3380-3389 (2005).
67. N. Steinberg, I. Kolodkin-Gal, The matrix reloaded: how sensing the extracellular matrix synchronizes bacterial communities. *J. Bacteriol.* **197**, 2092-2103 (2015).
68. L. Hall-Stoodley, J. W. Costerton, P. Stoodley, Bacterial biofilms: from the natural environment to infectious diseases. *Nat. Rev. Microbiol.* **2**, 95-108 (2004).
69. C. M. Dejea *et al.*, Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria. *Science* **359**, 592-597 (2018).
70. C. H. Johnson *et al.*, Metabolism links bacterial biofilms and colon carcinogenesis. *Cell Metab.* **21**, 891-897 (2015).
71. J. L. Drewes *et al.*, High-resolution bacterial 16S rRNA gene profile meta-analysis and biofilm status reveal common colorectal cancer consortia. *NPJ Biofilms Microbiomes* **3**, 34 (2017).
72. N. Hugen *et al.*, The molecular background of mucinous carcinoma beyond MUC2. *J. Pathol. Clin. Res.* **1**, 3-17 (2015).
73. J. R. Hyngstrom *et al.*, Clinicopathology and outcomes for mucinous and signet ring colorectal adenocarcinoma: analysis from the National Cancer Data Base. *Ann. Surg. Oncol.* **19**, 2814-2821 (2012).
74. H. Tanaka *et al.*, *BRAF* mutation, CpG island methylator phenotype and microsatellite instability occur more frequently and concordantly in mucinous than non-mucinous colorectal cancer. *Int. J. Cancer* **118**, 2765-2771 (2006).
75. P. Dharmani, J. Strauss, C. Ambrose, E. Allen-Vercos, K. Chadee, *Fusobacterium nucleatum* infection of colonic cells stimulates MUC2 mucin and tumor necrosis factor alpha. *Infect. Immun.* **79**, 2597-2607 (2011).
76. Y. Xue *et al.*, Host inflammatory response inhibits *Escherichia coli* O157:H7 adhesion to gut epithelium through augmentation of mucin expression. *Infect. Immun.* **82**, 1921-1930 (2014).
77. D. H. Ahn *et al.*, TNF-alpha activates MUC2 transcription via NF-kappaB but inhibits via JNK activation. *Cell. Physiol. Biochem.* **15**, 29-40 (2005).
78. F. Ishibashi *et al.*, Contribution of ATOH1+ cells to the homeostasis, repair, and tumorigenesis of the colonic epithelium. *Stem Cell Rep.* **10**, 27-42 (2018).
79. K. Sugimoto *et al.*, IL-22 ameliorates intestinal inflammation in a mouse model of ulcerative colitis. *J. Clin. Invest.* **118**, 534-544 (2008).
80. J. E. Turner, B. Stockinger, H. Helmby, IL-22 mediates goblet cell hyperplasia and worm expulsion in intestinal helminth infection. *PLoS Pathog.* **9**, e1003698 (2013).
81. S. Brand *et al.*, IL-22 is increased in active Crohn's disease and promotes proinflammatory gene expression and intestinal epithelial cell migration. *Am. J. Physiol. Gastrointest. Liver Physiol.* **290**, G827-838 (2006).
82. K. Fukushima *et al.*, Atonal homolog 1 protein stabilized by tumor necrosis factor alpha induces high malignant potential in colon cancer cell line. *Cancer Sci.* **106**, 1000-1007 (2015).

# Chapter 2

## Pancreatic Cyst Fluid Harbors a Unique Microbiome

Shan Li<sup>1</sup>, Gwenny M. Fuhler<sup>1</sup>, Nahush BN<sup>3</sup>, Tony Jose<sup>3</sup>, Marco J. Bruno<sup>1</sup>,  
Maikel P. Peppelenbosch<sup>1,2</sup>, Sergey R. Konstantinov<sup>1,4</sup>

<sup>1</sup>Department of Gastroenterology and Hepatology, Erasmus MC-University Medical Center, Rotterdam, The Netherlands.

<sup>2</sup>Erasmus Medical Center Cancer Institute, Rotterdam, The Netherlands.

<sup>3</sup>Cleverage Biocorp Private Limited, Bangalore, India.

<sup>4</sup>Janssen Vaccines and Prevention B.V., Leiden, The Netherlands.

**Microbiome.** 2017, 5:147.



## Abstract

**Background:** It is clear that specific intestinal bacteria are involved in the development of different premalignant conditions along the gastrointestinal tract. An analysis of the microbial constituents in the context of pancreatic cystic lesions has, however, as yet not been performed. This consideration prompted us to explore whether endoscopically obtained pancreatic cyst fluids (PCF) contain bacterial DNA and to determine the genera of bacteria present in such material.

**Methods:** Total DNA was isolated from 69 PCF samples. Bacterial 16S rRNA gene-specific PCR was performed followed by Sanger sequencing and de novo deep sequencing for the V3-V4 variable region of 16S rRNA gene.

**Results:** We observed that 92.8% of the samples were positive in conventional PCR, and that 100% of selected PCF samples (n = 33) were positive for bacterial microbiota as determined by next generation sequencing (NGS). Comprehensive NGS data analysis of PCF showed the presence of 408 genera of bacteria, of which 17 bacterial genera were uniquely abundant to PCF, when compared to the Human Microbiome Project (HMP) database and 15 bacterial microbiota were uniquely abundant in HMP only. *Bacteroides* spp., *Escherichia/Shigella* spp., and *Acidaminococcus* spp. which were predominant in PCF, while also a substantial *Staphylococcus* spp. and *Fusobacterium* spp. component was detected.

**Conclusion:** These results reveal and characterize an apparently specific bacterial ecosystem in pancreatic cyst fluid samples and may reflect the local microbiota in the pancreas. Some taxa with potential deleterious functions are present in the bacterial abundance profiles, suggesting that the unique microbiome in this specific niche may contribute to neoplastic processes in the pancreas. Further studies are needed to explore the intricate relationship between pathophysiological status in the host pancreas and its microbiota.

**Keywords:** 16S rRNA gene; Pancreatic cyst fluid; Cystic tumors; *Fusobacterium* spp.; Bacterial translocation; NGS; Human Microbiome Project.



## Background

Pancreatic cysts are fluid-filled neoplasms that can be detected with a frequency of up to 2% in the general asymptomatic adult population and have a low, but not negligible risk for evolving into pancreatic ductal adenocarcinoma, while accounting for up to 5% of the total incidence of pancreatic cancerous lesions (1,2). The vast majority of cysts are coincidentally found during cross-sectional imaging done for other reasons than cyst-related symptoms. Optimal clinical management of pancreatic cysts remains controversial, but there is consensus in the field that increased insight into the molecular pathogenesis of pancreatic cysts may guide development of rational strategies in this respect. Unfortunately, the etiology of pancreatic cysts remains largely obscure.

Progress with respect to understanding the nature and natural history of pancreatic cysts is compounded by the presence of different types of this lesion. Grutzmann et al. and Farrell et al. have attempted to classify the different types of pancreatic cyst lesions and distinguish among others as intra-ductal papillary mucinous neoplasms (IPMNs), mucinous cystic neoplasms (MCN), serous cystadenomas (SCA), and pseudocysts (3–5). In general, it is assumed that IPMN and MCN pose a higher risk of developing into carcinoma, with IPMN being more prevalent compared with MCN (1,6). IPMNs are further classified as main branch, side branch, or mixed types, based on the involvement of the duct in the pancreas (5). Presently, there are no validated biomarkers to identify cystic lesions that require surgical resection and this constitutes a major challenge in this field. Although pancreatic lesions develop into malignancy in only up to 3% of cases, 10% of patients with such lesions undergo resection (5) suggesting the need for superior clinical tests and patients' stratification prior surgery. Currently, the decision for resection of pancreatic cyst lesions and/or continued monitoring is made according to the Sendai guidelines after evaluation of different clinical tests (7–9). The available clinical tests include different biochemical analyses, cytology, pathological assessment of fine needle biopsy or aspiration material, endoscopic ultra-sonography (EUS), and radiological diagnosis such as endoscopic retrograde cholangiopancreatography (ERCP), magnetic resonance cholangiopancreatography (MRCP), and whole-body

computerized tomography (CT). The inter-observer agreement, however, between different modalities remains moderate (10). Therefore, a set of preoperative biochemical analyses have been increasingly used in clinical decision-making. This includes the study of cyst fluids and serum for the characteristic presence of carcinoembryonic antigen (CEA), cancer antigen 19.9 (CA19.9), cancer antigen 72.4 (CA72.4), cancer antigen 15.3 (CA15.3), pancreatic amylase, and mucin antigens, along with other cyst characteristics (11,12). Other tests are based on specific analysis of different genetic modalities like K-RAS mutation and integrity, but efforts attempting to provide clinical validation for such tests have largely proven unsuccessful (1,6,13,14). Increased insight into the factors that facilitate the development of cystic lesions would evidently benefit the identification of tests capable of providing guidance for clinical management of asymptomatic patients exhibiting pancreatic cysts.

Intriguingly, the human gut microbiome has emerged recently as an important environmental factor linked to the development of different intestinal and extra-intestinal malignancies (15–17). In the stomach, *Helicobacter pylori* remains the archetypical example of a prokaryotic organism that can initiate a cascade of molecular events finally leading to full-blown cancer, whereas in the colon, various organisms and especially *Fusobacteria* have been linked to the appearance of dysplasia [reviewed in (17)]. Whether the appearance of pancreatic cysts is linked to the presence of bacteria *per se* and if so whether specific types of bacteria are associated to the presence of cystic pancreatic neoplasms has remained unexplored. If, however, the presence of pancreatic cyst can be linked to the microbiome, this would entail a significant step forward with respect to our understanding of pancreatic cystogenesis.

The above-mentioned considerations prompted us to explore the potential microbial component of pancreatic cyst fluid. The results show that presence of bacterial DNA is common to such material and that especially *Fusobacterium* spp. and *Bacteroides* spp. are prominently present in such material. As some of these bacterial species have been linked to dysplastic processes elsewhere in the tracts, a causal link between the presence of such bacteria and pancreatic neoplasm may also exist and the results may indicate that bacterial colonization of pancreatic cyst fluid is a regular phenomenon.

## Methods

### Patient samples and pancreatic cyst fluid collection

A cohort of 69 patients with suspected cystic lesions was established between the period of 2008 and 2013 (**Table 1** and **Additional file 1**). The pancreatic cyst fluids (PCF) were collected after a signed informed consent from these patients, who were undergoing endoscopic ultrasound fine needle aspiration (EUS-FNA) at the Department of Gastroenterology, Erasmus MC, The Netherlands. Following collection, pancreatic cyst fluids were transferred to the laboratory and stored at  $-150^{\circ}\text{C}$  until analysis.

### DNA isolation

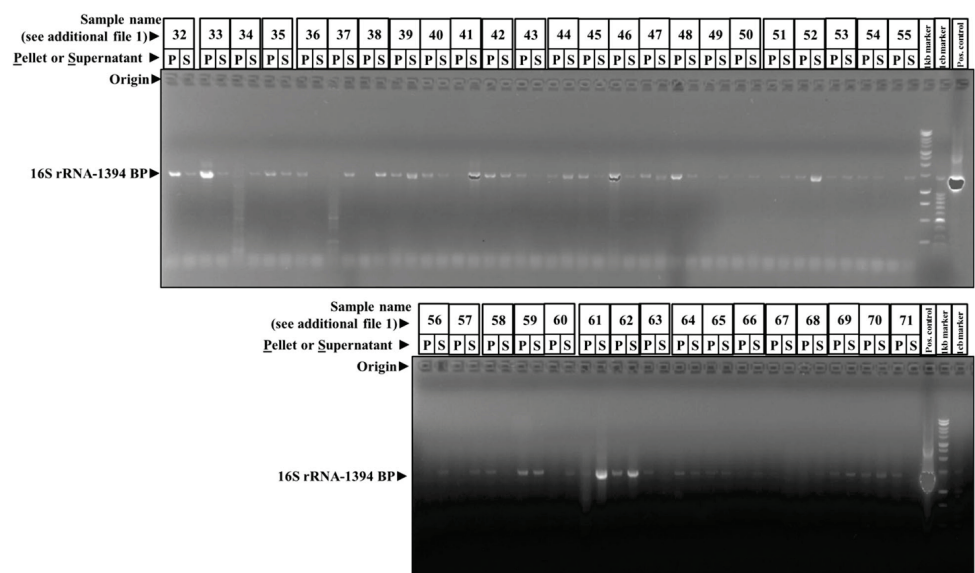
Approximately 300  $\mu\text{l}$  from 69 pancreatic cyst fluid samples were used for total DNA isolation. After bead beating (Fast Prep®-24 Instrument), the supernatant and pellet were separated by centrifugation at 13000 rpm for 1 min and both pellet and supernatant were used for total DNA isolation, using the Wizard DNA isolation kit as specified by the manufacturer's protocol (catalog no. A1620, Promega BNL B.V, The Netherlands). Isolated DNA was equilibrated in the DNA rehydration solution from the kit and quantified on nanodrop-2000 spectrophotometer (Isogen Life Science BV, De Meern, The Netherlands). Pancreatic cyst fluid DNA was diluted to 1 ng/ $\mu\text{l}$  for the PCR analyses and subsequently stored at  $-20^{\circ}\text{C}$ .

### PCR analyses

Total DNA isolated from pancreatic cyst fluids was used for bacterial 16S rRNA gene detection using conventional PCR. GoTaq® Flexi DNA polymerase kit (Promega, the Netherlands) and universal 16S rRNA primers were used (Bacteria cPCR-27F: 5'-AGAGTTTGATCCTGGCTCAG-3', Bacteria cPCR-1401R: 5'-CGGTGTGTACAAGACCC-3', 1394 bp product size, PCR condition,  $95^{\circ}\text{C}$  for 6 min; 30 cycles of  $95^{\circ}\text{C}$  denaturation for 30s,  $50^{\circ}\text{C}$  primer annealing for 30 s and  $72^{\circ}\text{C}$  elongation for 90 s; and final elongation of 7 min at  $72^{\circ}\text{C}$ ). The 16S rRNA gene amplicons were run on 1% agarose gels (Sigma Aldrich, the Netherlands), and the positive samples were selected by the presence of 16S rRNA gene 1394 bp band (**Fig. 1**).

**Table 1** Characteristics and clinicopathological features of the patients with pancreatic cyst

Patient characteristics	IPMN ( <i>n</i> = 27)	MCN ( <i>n</i> = 13)	Others ( <i>n</i> = 11)	Pseudocyst ( <i>n</i> = 9)	Serous cystadenoma ( <i>n</i> = 9)
Types of cysts (%) ( <i>n</i> = 69)	Main branch IPMN	2.9%	No definite clinical diagnosis (others)	6.0%	
	Mixed type IPMN	5.8%	NET (others)	3.0%	
	Multifocal side branch IPMN	5.8%	Acinar cell carcinoma (others)	1.0%	
	Side branch IPMN	17.0 %	Simple cyst (others)	3.0%	
	IPMN	7.2%	Cystic GLST; no communication PD (others)	3.0%	
Total percentage of samples in each pancreatic cyst	39.10%	18.8%	15.90%	13.0%	13.0%
Resected sample percentage	11.6%	13%	4.3%	2.9%	0.0%
Dysplasia	No Dysplasia	66.7%	53.8%	100%	100%
	Adenoma	11.1%	38.5%	0.0%	0.0%
	Moderate dysplasia	11.1%	7.7%	0.0%	0.0%
	Carcinoma in situ	11.1%	0.0%	0.0%	0.0%
Gender	Male	15.9%	0.0%	7.2%	2.9%
	Female	21.7%	17.4%	7.2%	10.1%
	Not available	1.4%	1.4%	0.0%	0.0%
Average age	68 years 22 days	11 months	53 years 7 months 14 days	62 years 5 months 14 days	54 years 8 months 19 days
					59 years 9 months 18 days



**Fig. 1** Abundance of bacterial DNA in pancreatic cyst fluid. A series of pancreatic cyst fluids (patient characteristics can be found by linking sample name to the patient information provided through additional Table 1) were exposed to bead beating and centrifugation followed by DNA extraction from both the pellet as well as the supernatant. Subsequently, the DNA was probed for the presence of sequence coding for bacterial 16S RNA through PCR. Reaction products were resolved by gel electrophoresis employing an agarose gel. The results show that bacterial DNA is commonly found in pancreatic cyst fluids.

## 16S rRNA gene sequencing analyses

### Sanger sequencing

Sanger sequencing was done in order to identify the bacterial 16S rRNA genes present in the PCF. The PCR products generated using universal primers of 16S rRNA were sequenced using F primer 5'-CTTAGGAATGAGACAGAGATG-3' through LGC Genomics GMBH, Germany. The chromatograms were analyzed, sequences were curated and identified using web databases of integrated microbial genomes and 16S rRNA-specific nucleotide BLAST. The sequences have been deposited at NCBI under gene accession identifiers MF061964–MF061990.

### Deep sequencing

From this cohort of 69 PCF samples, DNA of 33 samples [(47.8%), IPMN (n = 9), MCN (n = 7), Pseudocysts (n = 8), and SCA (n = 9)] was selected for de novo 16S rRNA gene

amplicon sequencing by Macrogen (South Korea). 16S rRNA gene amplicon sequencing was performed on the V3-V4 variable region using Illumina Miseq adapter PCR followed by clustering and sequencing. The raw images generated are used by MCS (MiSeq Control Software v2.4.1.3) for system control and base calling through an integrated primary analysis software called RTA (real-time analysis. v1.18.54.0). The BCL binary (base calls) is converted into FASTQ utilizing illumine package MSR (MiSeq Reporter).

### **Deep sequencing data analysis**

Two types of deep sequencing analyses were performed, one with paired-end reads alone discarding the single pair reads leading to 408 genera identification, and one with paired-end reads and single pair reads leading to identification of 785 genera (18). The data generated via paired-end reads alone was further used for downstream analysis for comparison of the PCF bacteria population vs human microbiome project samples of 13 different body sites and comparisons between the types of clinically defined cysts types and resection (**Additional file 1**).

### **Pre-processing of amplicon reads and sequence data QC**

Illumina Miseq sequence data was quality checked using FastQC and Rqc software's (19). Base call quality distribution, percentage of bases with quality scores above Q20 and Q30, GC percentage, and sequencing adapter contamination were used for assessing the data quality. The details of the number of reads and base quality obtained for each sample are provided in **Additional file 2**.

### **Taxonomic classification and OTU clustering**

For downstream analysis of the metagenome data, "mothur" software bundle was used (20). The quality-filtered sequence reads were imported into mothur, and the read pairs were aligned with each other to form contigs based on sequence consensus. This results in longer contigs that span the targeted V3-V4 hypervariable region. These contigs were screened and only those between 470 and 500 bp were taken for further

analysis. Contigs with ambiguous base calls were rejected; high-quality contigs were checked for identical sequences and duplicates merged.

The filtered contigs were aligned to a known database for 16S rRNA [Silva V.119] (21). These classified contigs were filtered for any undesired lineage from the taxonomy file. Any ambiguous contigs aligning to untargeted regions [other than V3-V4] were discarded. Using UCHIME algorithm, chimeric contigs were flagged and removed; a known reference for chimeric sequence was cross referenced (22). These final set of contigs were then phylotype binned into operational taxonomic units (OTU) based on the Silva V.119 database and the abundance of each OTU in the population was estimated.

### **Comparison of the taxonomic classification of the pancreatic microbiomes with those from other body sites**

In order to identify microbial signature of pancreatic cysts, a comparative analysis was carried out against organ-specific bacterial profiles obtained from the NIH Human Microbiome Project (HMP) database (23). The pancreatic microbial profile of the current study was compared with microbial profiles of 13 different body sites downloaded from HMP (SRA numbers provided in **Additional file 3**). The notion that bacteria in the gut may affect diseases outside the gut and vice versa is gradually accepted, and we therefore considered both anatomical sites in close proximity to the pancreas as well as more distant sites.

The comparative analysis was carried out between anterior nares, antecubital fossa, buccal mucosa, gingiva, hard palate, mid vagina, palatine tonsils, posterior fornix, retroauricular crease, saliva, stool, throat, and tongue. Although phylotype-based analysis is limited as compared to distance-based OTU classification, it allows for investigating the relationship of the PCF microbiota to previously characterized microbes in the HMP database.

Two analyses were performed for identifying the unique microbiota when PCF commensals were compared to the bacterial communities associated with 13 different body sites. First, a pairwise binomial test was carried out against each organ in order

to identify significantly ( $FDR < 0.05$ ,  $FC > 3$ ) and absolute abundance difference of 10 abundant bacterial species in pancreas. Second, in order to identify the relatedness of the PCF microbiomes with those from other body sites, we compared bacterial communities of the PCF samples to those of taxonomic profiles from 13 other body sites from the HMP database by using principal component analysis (PCA). Third, statistical comparisons of these microbiomes were performed using the STAMP analysis package. The statistical tests used was ANOVA and Welch's  $t$  test, with Benjamini-Hochberg multiple test corrections. Those taxa having a higher abundance in the pancreatic cysts (with corrected  $P$  value  $< 0.01$ ) were identified as those that are specifically abundant in the PCF.

### **Diversity analysis between the cyst types**

Using STAMP statistical analysis package, those taxa having a higher abundance between the test groups IPMN vs pseudocyst vs MCN vs SCA were identified as specific for the groups to be classified based on the bacterial population. The cysts were grouped based on the pathological classifications, and they are tested using White's non parametric  $t$  test, two tailed with Benjamini-Hochberg multiple test corrections of  $P$  value ( $P < 0.05$ ) (24). The measured levels of CEA and CA19.9 ranges were used for classification of cysts as cysts, benign cysts, and malignant cysts according to study conducted by Talar-Wojnarowska et al. (25). Then the classified data was used for the specific bacterial taxa identification between the resected types using the same parameters.

### **Statistics and calculations**

All the statistical analysis were done using excel, Graphpad Prism 5.0., and STAMP statistical analysis software. When appropriate, results were Bonferroni corrected in Graphpad analysis and Benjamini-Hochberg multiple test corrections in STAMP.  $P$  values  $< 0.05$  (between cysts and resection types) and  $< 0.01$  [between body sites (NGS)] were considered significant.



## Results

### Study sample characteristics

For analysis of the potential microbial component of pancreatic cyst fluid (PCF), material collected from 69 patients was used. The characteristics of these patients are listed in **Table 1** and **Additional file 1**. Of the PCF analyzed, 27 were obtained from patients harboring an IPMN (39.1%), including two patients with main branch IPMN (2.9%), four patients with mixed type IPMN (5.8%), 12 patients with a unifocal side-branch IPMN (17.4%), and four patients presenting with a multifocal side-branch IPMN (5.8%). In five patients, the type of IPMN was unclassified (7.2%). In addition, our cohort contained 13 patients with an MCN (18.8%), nine patients with pseudocysts (13.0%), and nine patients with serous cystadenomas (13.0%). Finally, the cohort contained 11 patients with apparently multiple forms of cystic lesions, gastrointestinal stromal tumor (GIST), neuroendocrine tumors (NET), or having no definitive clinical diagnosis and for this study, these patient were classified as “others” (15.9%). We concluded that this cohort would allow the study of potential microbiological constituents of PCF and to relate results to the clinical phenotype of the patient from which the fluid was obtained.

### Bacterial DNA is commonly present in EUS-FNA-collected pancreatic cyst fluids

As it is as yet unknown whether pancreatic cystic fluid hosts a microbiological component, analyzed the EUS-FNA-collected pancreatic cyst fluid obtained from our cohort for the presence of significant amounts of bacterial DNA. Importantly, we found that the majority of these fluid samples were rich for bacterial DNA, with 16S rRNA PCR demonstrating the presence of bacterial DNA in 64 (92.8%) out of the 69 samples (**Table 2** and **Fig. 1**). The presence of bacterial DNA in cyst fluid did not statistically relate to the type of lesion from which it was obtained ( $P$  value > 0.99;  $\chi^2$  test): in mucinous cystic neoplasms, 100% of samples contained significant amounts of bacterial DNA; in IPMNs, this number was 92.6%; in pseudocysts, 88.9% of PCF were positive for bacterial DNA; in serous cystadenomas, 88.9%; whereas in the group of others which included GIST, NET, and clinically undefined samples, 90.9% displayed significant amounts of bacterial DNA (**Additional file 1** and **Table 2**).

**Table 2** Bacterial ecosystems characteristics identified from the pancreatic cyst fluids using PCR, Sanger sequencing, and next-generation sequencing (NGS)

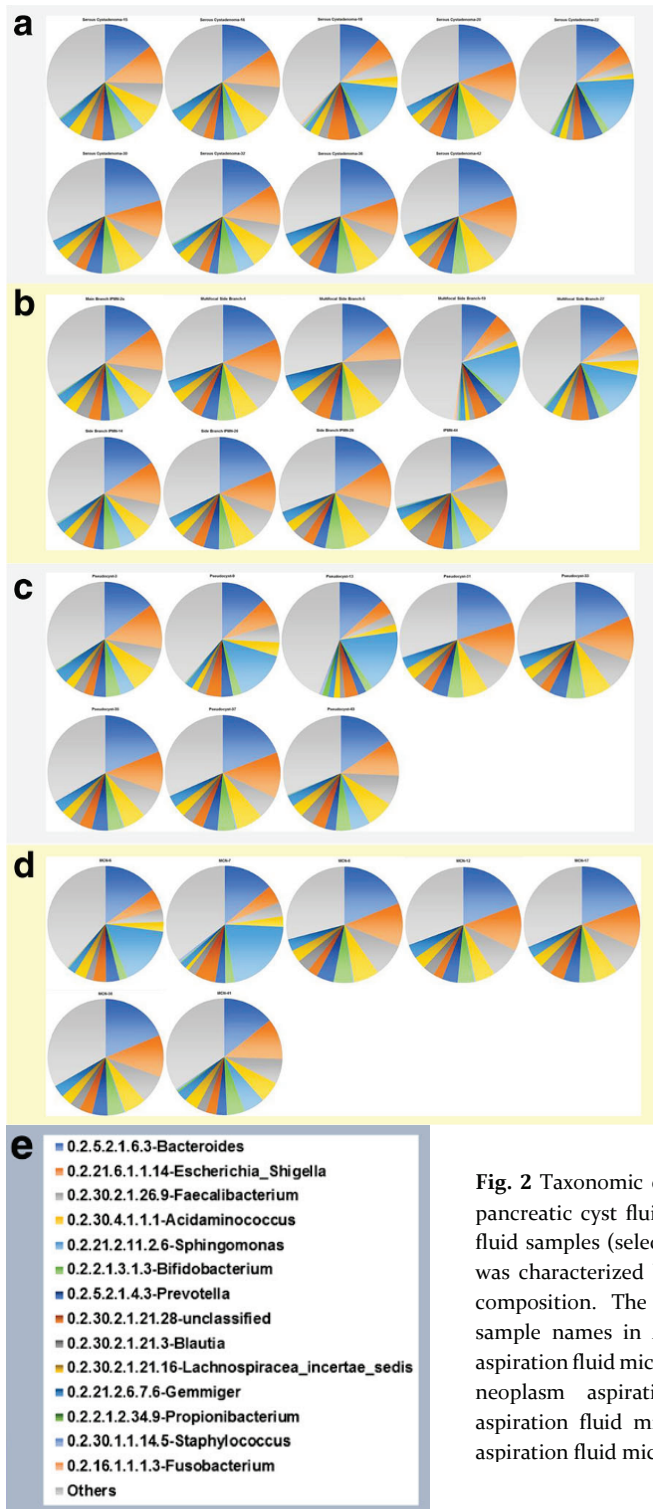
Bacterial ecosystem characteristics					
Patient characteristics	IPMN ( <i>n</i> = 27)	MCN ( <i>n</i> = 13)	Others ( <i>n</i> = 11)	Pseudocysts ( <i>n</i> = 9)	Serous cystadenoma ( <i>n</i> = 9)
16S rRNA PCR (universal 16S rRNA gene primers)	Bacteria present  7.4%	100%	90.9%	88.9%	88.9%
Sanger sequencing	Bacteria detected  <i>Bacillus</i> spp., <i>Fusobacterium</i> spp., <i>Orpinomyces</i> spp., <i>Anaerococcus</i> spp., <i>Caldimonas</i> spp., <i>Acinetobacter</i> spp., <i>Bacillus</i> spp.	<i>Fusobacterium</i> spp., <i>Bacillus</i> spp., <i>Orpinomyces</i> spp., <i>Microcystis</i> spp., <i>Staphylococcus</i> spp.	<i>Fusobacterium</i> spp.	<i>Caldimonas</i> spp., <i>Propionibacterium</i> spp., <i>Fusobacterium</i> spp., <i>Curvibacter</i> spp., <i>Escherichia</i> spp., <i>Bacillus</i> spp.	<i>Arthrobacter</i> spp., <i>Bacillus</i> spp., <i>Bacteroides</i> spp., <i>Ruminococcus</i> spp.
16S rRNA (NGS) ( <i>n</i> = 33)	Bacteria present  100% ( <i>n</i> = 9)	100% ( <i>n</i> = 7)	NA	100% ( <i>n</i> = 8)	100% ( <i>n</i> = 9)
Bacteria detected by 16S rRNA gene V3-V4 variable region NSG ( <i>n</i> = 33)	<i>Bacteroides</i> —15.45% <i>Escherichia/Shigella</i> —9.88% <i>Faecalibacterium</i> —8.57% <i>Acidaminococcus</i> —5.75% <i>Sphingomonas</i> —4.87% <i>Others</i> —55.49%	<i>Bacteroides</i> —17.06% <i>Escherichia/Shigella</i> —10.17% <i>Faecalibacterium</i> —6.95% <i>Acidaminococcus</i> —5.22% <i>Sphingomonas</i> —6.48% <i>Others</i> —54.12%	NA	<i>Bacteroides</i> —16.59% <i>Escherichia/Shigella</i> —10.55% <i>Faecalibacterium</i> —6.81% <i>Acidaminococcus</i> —6.23% <i>Sphingomonas</i> —5.40% <i>Others</i> —54.42%	<i>Bacteroides</i> —16.73% <i>Escherichia/Shigella</i> —9.97% <i>Faecalibacterium</i> —6.64% <i>Acidaminococcus</i> —6.24% <i>Sphingomonas</i> —4.81% <i>Others</i> —55.62%

NA: not applicable.

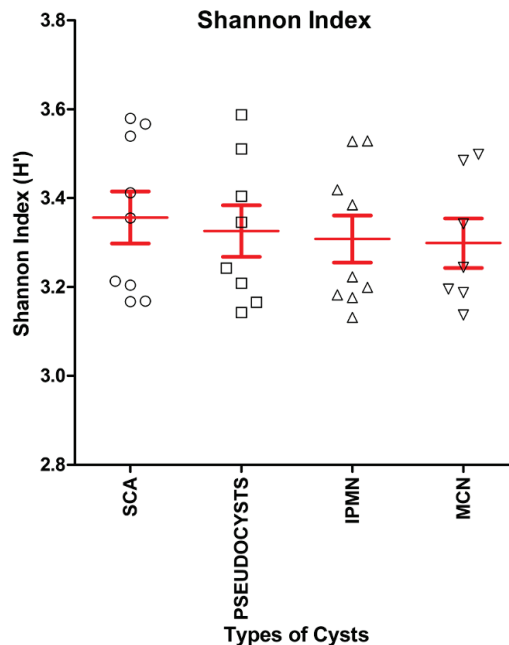
### **Microbial composition of pancreatic cyst fluid with a differential clinical aspect**

The presence of a microbial component in PCF raises obvious questions as to the identity of the organisms apparently present in such fluid. Sanger sequencing of PCR products generated using universal 16S rRNA demonstrated that *Fusobacterium* spp. is present in 13 out of all 69 PCF samples (18.84%). Another predominating bacterium in PCF was *Bacillus* spp. which was present in 16 out of 69 (23.19%) samples. The presence of other bacteria was also noted which included *Ruminococcus* spp., *Staphylococcus* spp., *Caldimonas* spp., *Arthrobacter* spp., *Acinetobacter* spp., *Bacteroides* spp., *Orpinomyces* spp., and *Anaerococcus* spp. (**Table 2**).

To confirm the presence of these bacteria and gain more insight into the bacterial composition of the PCF, DNA obtained from the 23 fluids containing the highest apparent concentration of 16S rRNA copies were sent out for 16S rRNA de novo sequencing of V3-V4 region, using universal 16S rRNA primers which should allow identification of the bacteria present in such fluid at least on genus level. The results (shown in **Fig. 2** and **Additional file 2**) are consistent with the presence of diverse bacterial ecosystems in such fluids, with as most predominant genera present *Bacteroides*, *Escherichia/Shigella*, and *Acidaminococcus*, but in total, 408 different genera were detected in the 33 samples analyzed, of which 93 genera were found in at least 50% of PCF samples analyzed. Different types of cysts were not statistically different with respect to microbial composition: for none of the 408 genera, a Bonferroni-corrected statistically significant difference in abundance was detected when serous cystadenoma, pseudocysts, IPMN, or mucous cystic neoplasm-derived fluids were compared. Furthermore, when low *P* values (Bonferroni-uncorrected *P* value < 0.05 but > 0.0001) were considered, it appeared that such findings were limited to very low abundant organisms (< 0.05% of all bacteria) that were relative low in one of the groups compared, more indicative of technical detection problems rather than a reflection of true biological differences between the groups (**Additional file 2**). In apparent agreement with the notion that the ecological niche provided by cyst fluid is relatively similar between different cyst manifestations is also the observation that the Shannon index for ecological diversity (*H'*) (**Fig. 3**) is not different between the different groups (*P* value = 0.99; one-way ANOVA).



**Fig. 2** Taxonomic composition of microbiological genera of pancreatic cyst fluid. DNA isolated from 33 pancreatic cyst fluid samples (selected based on strong PCR signals for 16S) was characterized by 16S rRNA NGS for its microbiological composition. The numbers indicated correspond to the sample names in Additional file 1. **a** Serous cystadenoma aspiration fluid microbiome. **b** Intraductal papillary mucinous neoplasm aspiration fluid microbiome. **c** Pseudocyst aspiration fluid microbiome. **d** Mucinous cystic neoplasm aspiration fluid microbiome. **e** Legend of the pie chart.



**Fig. 3** Shannon index for ecological microbial diversity in pancreatic cyst fluids. Employing the NGS results from PCF-derived DNA, the Shannon index for ecological diversity was calculated. There are no significant differences in the ecological diversity of the microbiome in the different types of PCF.

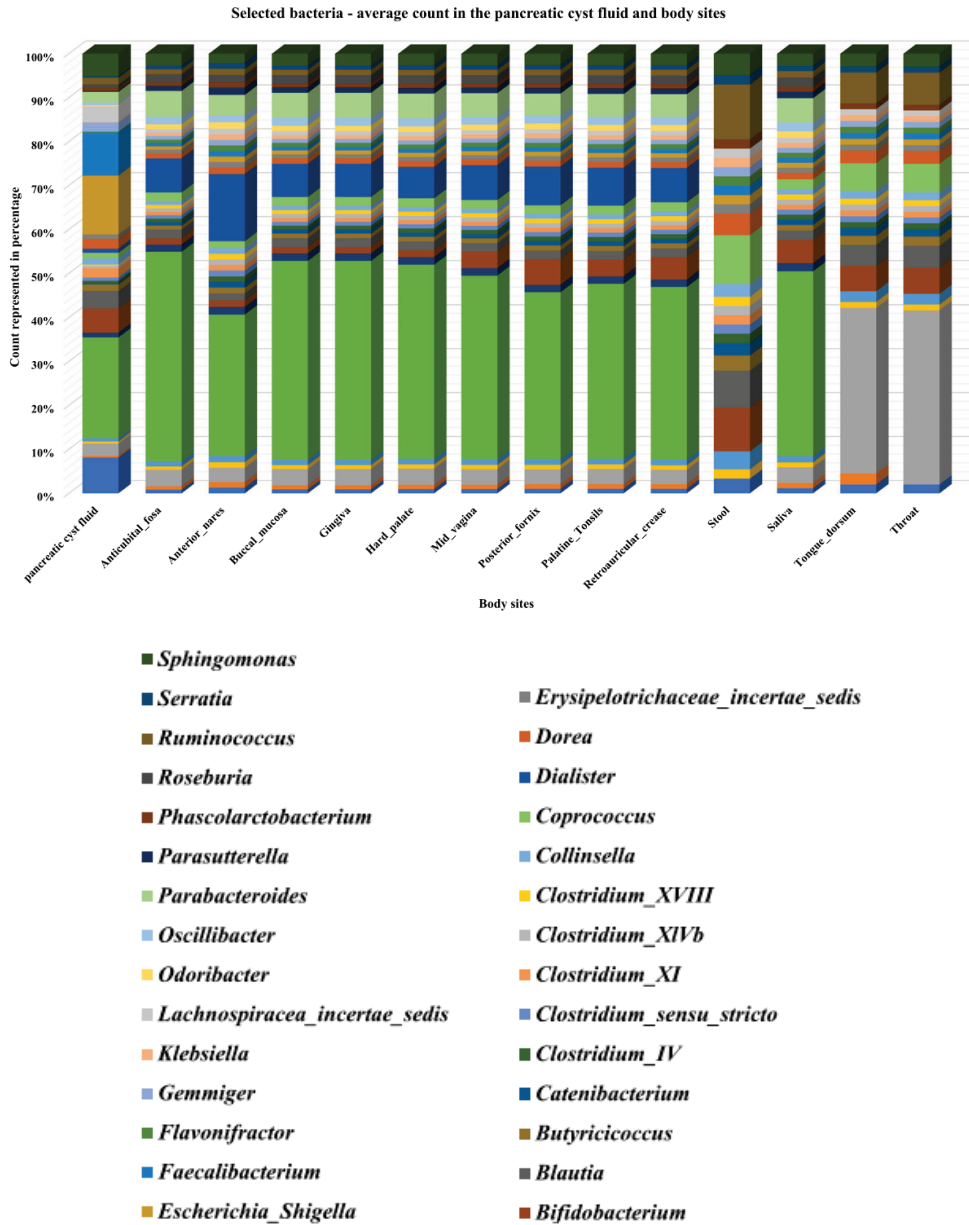
### Diversity within the cyst type and resections

There were no significant difference in the diversity of bacterial microbiota seen in the 33 sample group of deep sequenced data, between the cyst types, or between the resected vs non resected groups or between the CA19.9 or CEA range-based classification groups. The results were in accordance with the bacterial presence in the different types of cysts based on the universal primer PCR sequencing.

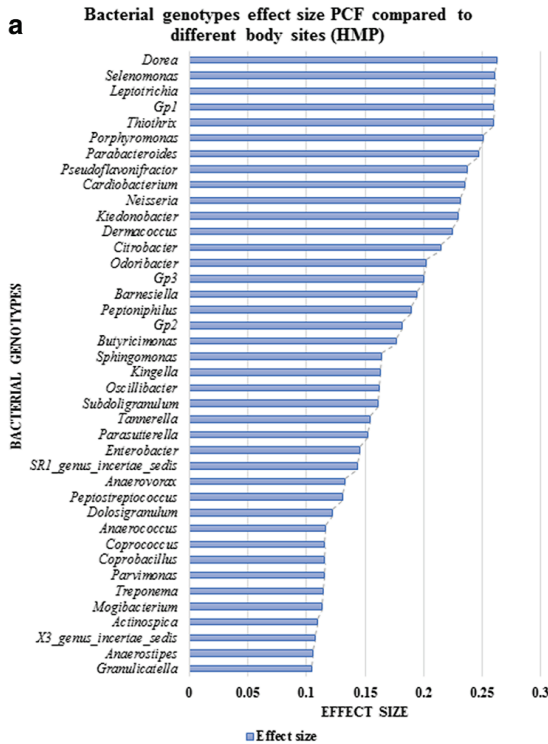
### Microbial composition of pancreatic cyst fluid suggests the presence of a unique ecosystem

A comparison of 13 different body sites (Human Microbiome Project) with the PCF bacterial population showed 26 unique bacterial genotypes in PCF when compared with stool, 27 genotypes compared with throat, 28 genotypes compared with tongue dorsum, and 36 unique bacterial genotypes in the PCF which are not present within any of the other body sites selected (**Fig. 4** and **Additional file 4**). ANOVA analysis,

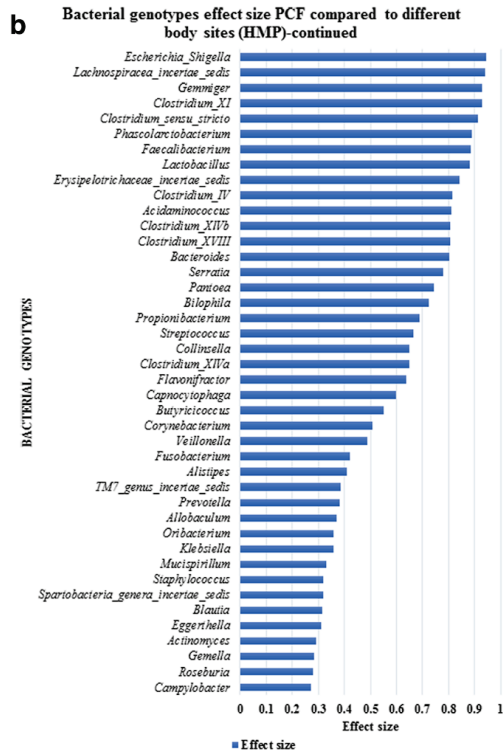
which is performed to identify the effect size of the bacterial population, reveals that 314 bacterial microbiota show an effect size variation between the bacterial commensals in PCF versus selected body sites, high effect size (0.1 to 0.9 and  $P < 0.01$ ) bacterial populations are shown in **Fig. 5**, and total effect size variation is shown in **Additional file 5**. The PCA of different body sites and PCF are shown in **Additional file 6: Figure S1**. The Welch's  $t$  test was performed to identify the specific bacterial populations having high and low abundance in the PCF when compared to HMP selected body sites. A total of 136 bacteria (**Additional file 6: Figure S2** and **Additional file 7**) are identified, of which 17 abundant bacteria in PCF were highly unique and potentially pro-cancerous. They are, in ascending order of abundance in PCF, *Coprococcus* spp., *Collinsella* spp., *Butyricicoccus* spp., *Ruminococcus* spp., *Parabacteroides* spp., *Alistipes* spp., *Clostridium* XI spp., *Gemmiger* spp., *Dorea* spp., *Lachnospiracea incertae sedis* spp., *Blautia* spp., *Bifidobacterium* spp., *Sphingomonas* spp., *Acidaminococcus* spp., *Faecalibacterium* spp., *Escherichia/Shigella*, and *Bacteroides* spp. ( $P < 0.0001$ ). A role for many of these bacteria in initiation and progression of colon, lung, and liver cancer has been suggested (26). The following 15 bacterial genera were higher in all the 13 body sites used in this analysis: *Streptococcus* spp., *Propionibacterium* spp., *Lactobacillus* spp., *Fusobacterium* spp., *Corynebacterium* spp., *Veillonella* spp., *Neisseria* spp., *Staphylococcus* spp., *Porphyromonas* spp., *Prevotella* spp., *Leptotrichia* spp., *Actinomyces* spp., *Capnocytophaga* spp., *Gemella* spp., and *Selenomonas* spp. ( $P < 0.0001$ ). These were also the least abundant bacteria present in the PCF. This comparison suggests that the PCF fluid bacterial colonization is unique and characterized by high abundant genera which have a tendency to propagate in the cancerous microenvironment, feeding the tumor (**Fig. 6**).



**Fig. 4** Pancreatic cyst microbiome is unique among other human body microbiomes. Publicly accessible databases were mined for composition of microbiomes at different human body sites and compared to those observed in PCF. It appears that there are 27 to 314 bacterial genotypes differently present in the PCF when compared to the selected body sites when analyzed via pairwise binomial test with high abundance PCF bacteria ( $P < 0.0001$ ) and ANOVA test, respectively.



**Fig. 5 a, b** Contribution of individual bacterial genera to the unique aspect of PCF fluid. For 82 different genera the relative in abundance in PCF was compared to that 13 other body sites and the relative contribution to the PCF-specific nature of the microbiome spectrum was calculated. In additional files information on a further 232 genera can be found. The results show that PCF contains a microbiome that is characterized by an overall uniqueness that cannot be attributed to a single genus.





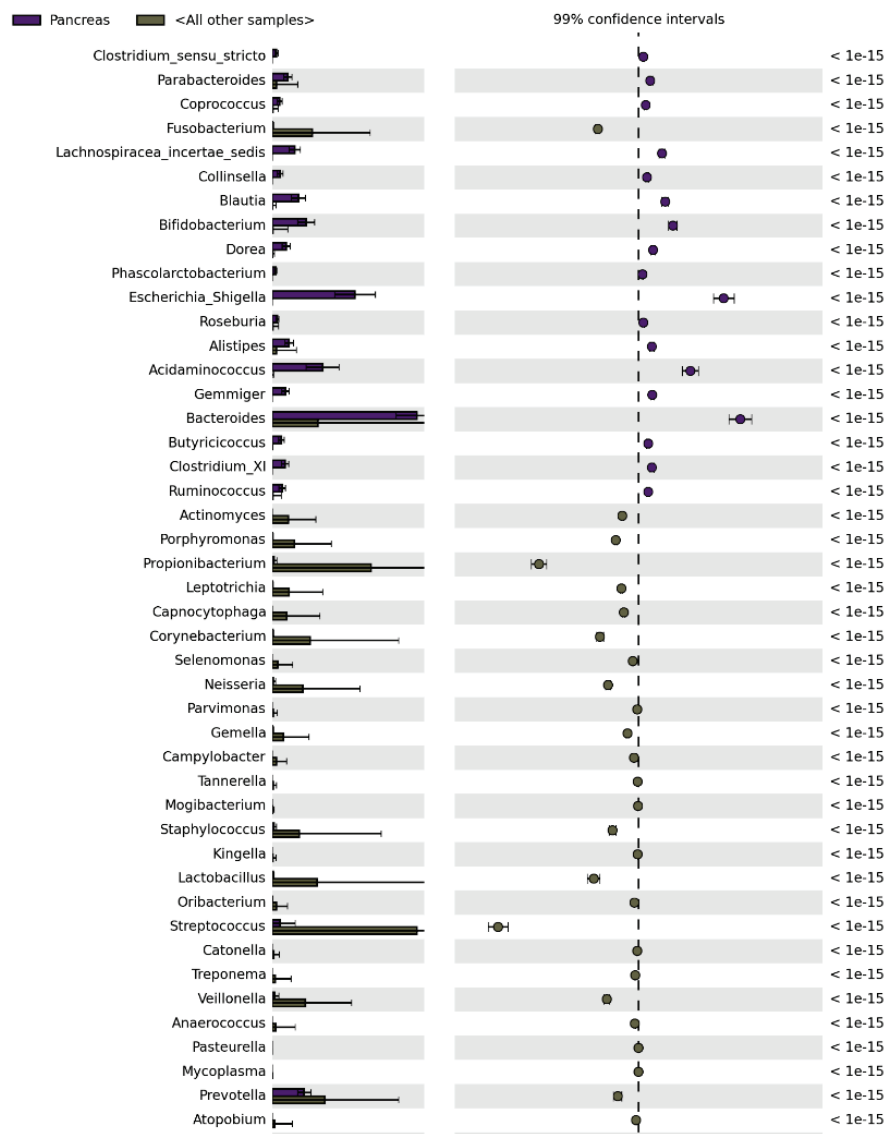


Fig. 6 (figure continued on next page)

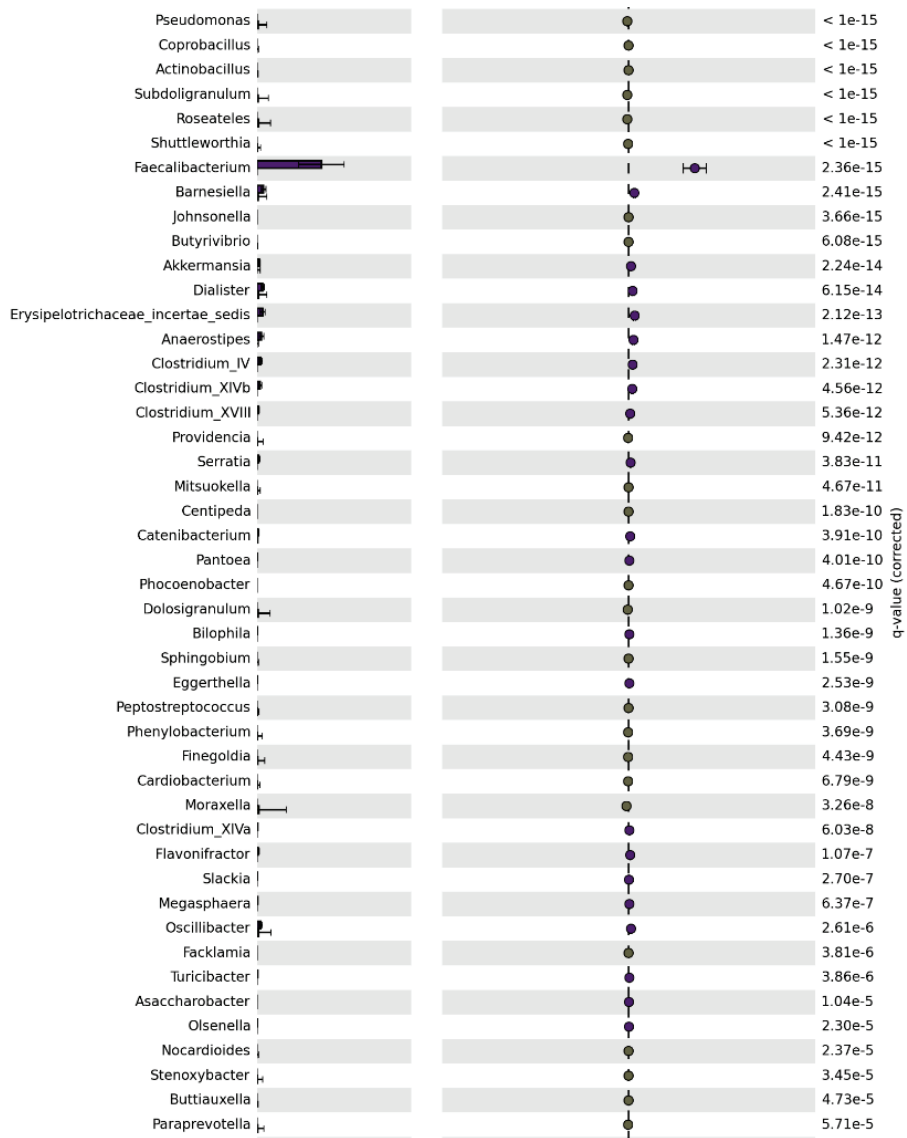
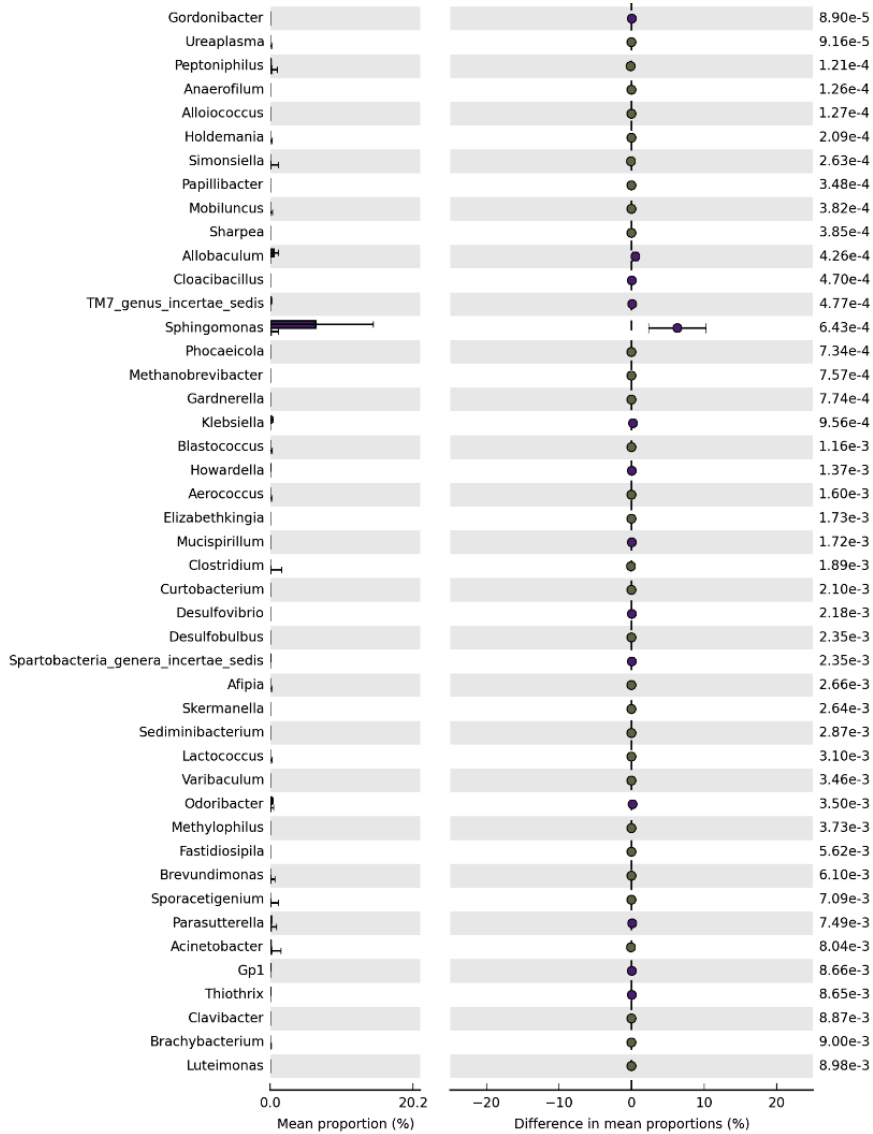


Fig. 6 (figure continued on next page)



**Fig. 6** (figure continued from previous page) Bacterial commensals of PCF and selected different body sites comparison (HMP database) shows the difference in the distribution uniqueness in the PCF than the body sites. One hundred thirty-six bacterial genus of high and low abundance with  $P < 0.0001$ , and their difference of mean is plotted with 99% confidence interval, obtained via the Welch's  $t$  test comparison between the groups.

## Discussion

In the current study, we have demonstrated that 92.8% of a large and clinically representative collection of pancreatic cyst fluids (PCF) is host to a complicated ecosystem of bacteria. The distinct bacterial community had a rich species diversity which is very different from that observed elsewhere in the proximal human digestive tract. Although the EUS-FNA is not a sterile procedure, 7.8% of the PCF samples were negative for bacterial DNA arguing that the procedure did not have a major impact on our cultivation-independent bacteriological findings. Furthermore, the ecosystem observed by deep sequencing of pancreatic cyst material is inconsistent with that expected from contamination by the oral, nasal, pharyngeal, oesophageal, gastric, or duodenal flora. Members of the genus *Streptococcus* constitute by far the major component of the bacterial ecosystem in the esophagus (27), but while *Streptococcus* was detected in all samples sequenced, it constitutes less than 1% of the total microbiota in any sample in our series, with the exception of one sample obtained from an IPMN. Likewise, the stomach is relatively sterile and dominated by *H. pylori*. The genus *Helicobacter*, however, was rare in our analysis, this genus the 67th genus in our ranking of our samples. Other bacteria found regularly in the stomach include *Streptococcus*, *Neisseria*, and *Lactobacillus*, and these species are not dominant in our analysis (being ranked as the 18th, 129th, and 54th most prevalent genus) (20). Thus, the esophagus and stomach are an unlikely source of major contamination in our results. The duodenum can be home to substantial bacterial content, as also evident from conditions such as small intestinal bacterial overgrowth or SIBO. Valeria D'Argenio et al. recently reported on the bacterial composition of the duodenum in a cohort including 15 healthy individuals (28). The most prevalent genus emerging from this analysis was *Propionibacterium*, which, although fairly often detected, was relatively rare in our analysis of pancreatic cyst fluids, with its abundance not even approaching 1% of the entire flora in any of the samples tested. Other genera dominating the duodenal microbiota include *Porphyromonas* (ranked 76th in cyst fluid), *Streptococcus*, *Neisseria*, and *Heamophilus* (ranked 21st, 58th, and 57th in cyst fluid). Hence, the bacteria we observed in cyst fluid widely diverges from that expected

if duodenal contamination was a major factor. Indeed in an analysis of microbiome of the healthy proximal tracts in our own institution (manuscript in preparation), we observed that *Streptococcus*, *Veillonella*, *Prevotella*, and *Pseudomonas* have the highest relative abundance from distal esophagus to the jejunum and proximal ileum. This situation is markedly different from that observed in pancreatic cyst fluids (although *Prevotella* was the 7th most abundant genus in pancreatic cyst fluid, this is still markedly lower as that observed in the proximal gastrointestinal tract of volunteers). In apparent agreement, in a preliminary series of experiments employing fluorescent in situ hybridization on surgically obtained pancreatic cyst material, bacteria were apparent (not shown). Thus, the most straightforward interpretation of results is that the pancreatic cyst is home to a previously unsuspected and also unique bacterial ecosystem.

The presence of such a bacterial ecosystem in pancreatic cysts raises important questions as to the role of the bacteria present in such cysts in the development of such structures. In general bacteria have been linked to transformation and trans-differentiation of endodermal epithelia, with *H. pylori* being the most important example. The genus *Helicobacter* was only marginally detected in pancreatic cyst fluid; other bacterial species linked to transformation in the intestine were, however, more prevalent. In particular, *F. nucleatum*, which is relatively predominant in a fraction of the samples and has been convincingly linked to neoplasm formation in the colon, excites interest in this respect. Emerging evidence suggests that *Fusobacterium* species detected in pancreatic cancer tissues is associated with a worse clinical outcome in pancreatic cancer patients (29). But also other bacteria seen in pancreatic cyst fluids, even if it is at low abundance, may have a role here. Using NGS technology, multiple studies have compared oral microbiota between healthy individuals and those with pancreatic cancer (4,30). Although no correlation between known pro-oncogenic oral pathogens and pancreatic cancer was detected, Lin et al. showed that pancreatic cancer patients had significantly higher levels of *Bacteroides* genus compared to control subjects (30), which is in agreement with our finding that *Bacteroides* spp. were the most predominant genus in PCF. In addition, a recent pilot study using 16S rRNA gene sequencing on saliva specimens showed that while there was no difference

in diversity of oral microbiota between patients with pancreatic ductal adenocarcinoma (PDAC), IMPN, or healthy controls, pancreatic cancer patients had higher levels of members of the phylum *Firmicutes* while healthy individuals were associated with higher relative abundances of *Proteobacteria* (31). Evidence gathered from those studies supposes that pancreatic malignancy may be associated with changes in abundances of some groups of bacteria in the human digestive tract. Therefore, dysbiotic microbiota in the upper digestive tract including oral cavity may interrupt the unique ecosystem in pancreatic cyst fluids along the neoplastic process in pancreas. At bay with the notion of a causative role for the bacterial flora in pancreatic cyst formation, however, is the observation that both detection of bacteria *per se* or the composition of the cyst flora did not show correlation to clinical parameters and thus the role of the flora in cystogenesis, if any, awaits further study.

Pancreatic infections mainly arise from translocation of bacteria from the small bowel, and rarely from the colon and oropharyngeal route as demonstrated by study on *Veillonella* and *Bifidobacterium* spp. which were identified in pancreatic abscesses (32). A study of Brook et al. has identified 158 bacterial species from pancreatic abscesses, of which 77 isolates were aerobic and the remaining 81 were anaerobic bacteria (33). The most commonly detected microorganisms in infected pancreatic pseudocysts include often not only opportunistic bacteria like *E. coli*, *Enterobacter* spp., *Klebsiella* spp., and *Staphylococcus* spp. but also fungal isolates including *Candida albicans* (15 case studies) (34). Importantly, EUS FNAB procedure caused serious *Clostridium perfringens* infections in five patients leading to pancreatitis and pancreatic cyst formation, which required surgical interventions (35). Such studies illustrate the nature of the bacterial transfer from the early to mid-gut commensal bacteria to the pancreas. Yet none of these earlier studies have directly proved the presence of bacteria in the pancreatic cyst and its fluid. Its apparent divergence from the flora in the duodenum is probably a reflection of the specific conditions in pancreatic cysts which include an absence of contact with the digestive nutrient stream, an exposure to high levels of pancreatic secretates and the presence of abundant mucus. At present, clinical behavior of pancreatic cysts is very difficult to predict and adequate management of cysts represents one of the largest challenges in clinical

gastroenterology. It is tempting to speculate, however, that changing conditions in cysts related to transformation of the structure to full-blown malignant cancer would also influence the *milieu interieur* formed by cyst fluid and hence have a profound effect on the bacterial composition. As such, changes in the bacterial communities may serve to detect such transformation in cyst-forming structures and may become useful for guiding clinical management.

Irrespective, however, of its potential as future diagnostic and prognostic marker, the present study shows an as yet unknown bacterial ecosystem in pancreatic cyst fluid. As it is evident that bacteria influence physiology and pathophysiology of their interacting epithelia everywhere in the gut, it is well possible that such interactions also exist in the pancreas and its cysts, and that the biology of the pancreatic cysts is in important ways shaped by this ecosystem. Studies addressing this possibility are currently in progress.

## Conclusions

The study reveals previously undescribed bacterial diversity present in human pancreas and its cyst fluids. As specific bacteria are associated with this body site, we propose that such bacteria may carry the potential to influence the development of pathophysiological processes in the pancreas. The study points out to the need to further explore the microbiome in this specific niche for diagnostic and therapeutic purposes.

## **Acknowledgements**

PCF fluid material kindly provided by Gastroenterologists, Dr. Wesley K Utomo, Dr. Henri Braat, Dr. J.W. Poley, and Dr. A.D. Koch, Erasmus MC, The Netherlands. The authors acknowledge Vilvapathy Narayanan for his work related to the manuscript.

## **Authors' contributions**

SL participated in the molecular analyses and drafted the manuscript. NB and TJ performed the statistical analysis of the NGS data. MPP, GF, and MB participated in the design of the study and reviewed the manuscript. SK conceived of the study, participated in its design and coordination and drafted the manuscript. All authors read and approved the final manuscript.

## **Ethics approval and consent to participate**

This study received approval of the Institutional Review Board (MEC-2008-233 and MEC-2012-107). A Declaration of Helsinki protocols was followed, and all patients provided written informed consent.

## **Availability of data and materials**

The Sanger sequences and NGS data have been deposited at NCBI under gene accession identifiers MF061964–MF061990 (Sanger) and identifies in Additional file 1 (NGS).



## **Additional files**

**Additional file 1:** Detailed patient characteristics, samples characteristic of samples taken for NGS, gene accession numbers of Sanger sequenced samples and SRA numbers for the NGS selected samples, NGS sequenced V3-V4 variable regions of 16S rRNA. Excel file containing data is available online at:

[https://static-content.springer.com/esm/art%3A10.1186%2Fs40168-017-0363-6/MediaObjects/40168\\_2017\\_363\\_MOESM1\\_ESM.xlsx](https://static-content.springer.com/esm/art%3A10.1186%2Fs40168-017-0363-6/MediaObjects/40168_2017_363_MOESM1_ESM.xlsx)

**Additional file 2:** Detailed quality of reads, taxonomic profiles, taxonomic distribution and taxonomic distribution rank. Excel file containing data is available online at:

[https://static-content.springer.com/esm/art%3A10.1186%2Fs40168-017-0363-6/MediaObjects/40168\\_2017\\_363\\_MOESM2\\_ESM.xlsx](https://static-content.springer.com/esm/art%3A10.1186%2Fs40168-017-0363-6/MediaObjects/40168_2017_363_MOESM2_ESM.xlsx)

**Additional file 3:** SRA numbers of Human Microbiome Project used for comparisons. Excel file containing data is available online at:

[https://static-content.springer.com/esm/art%3A10.1186%2Fs40168-017-0363-6/MediaObjects/40168\\_2017\\_363\\_MOESM3\\_ESM.xlsx](https://static-content.springer.com/esm/art%3A10.1186%2Fs40168-017-0363-6/MediaObjects/40168_2017_363_MOESM3_ESM.xlsx)

**Additional file 4:** Selected bacterial counts of the PCF and selected body sites used for the binomial pairwise comparison. Excel file containing data is available online at:

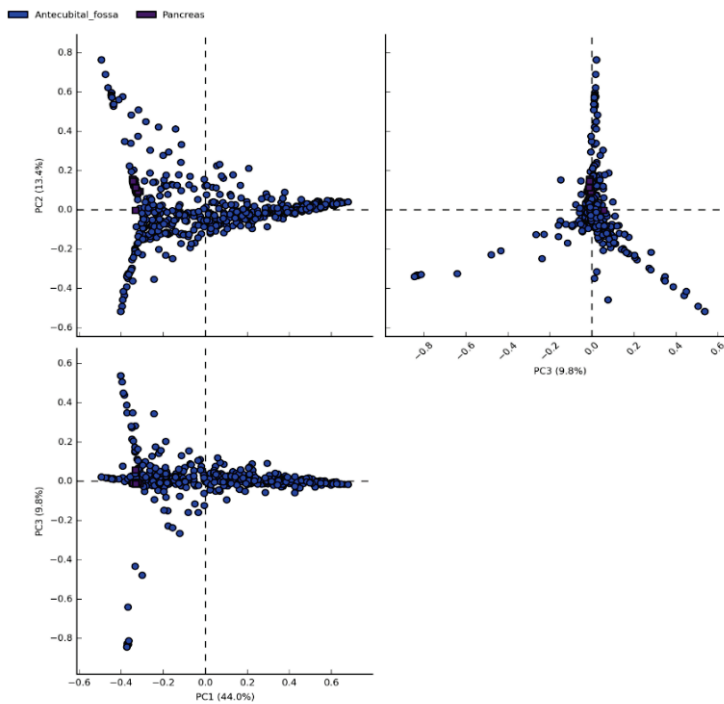
[https://static-content.springer.com/esm/art%3A10.1186%2Fs40168-017-0363-6/MediaObjects/40168\\_2017\\_363\\_MOESM4\\_ESM.xlsx](https://static-content.springer.com/esm/art%3A10.1186%2Fs40168-017-0363-6/MediaObjects/40168_2017_363_MOESM4_ESM.xlsx)

**Additional file 5:** ANOVA analysis statistics table for the PCF and selected body sites of bacterial distribution. Excel file containing data is available online at:

[https://static-content.springer.com/esm/art%3A10.1186%2Fs40168-017-0363-6/MediaObjects/40168\\_2017\\_363\\_MOESM5\\_ESM.xlsx](https://static-content.springer.com/esm/art%3A10.1186%2Fs40168-017-0363-6/MediaObjects/40168_2017_363_MOESM5_ESM.xlsx)

**Additional file 6:**

**Figure S1.** PCA of pancreatic cyst fluid (PCF) and 13 body site microbiome comparisons. PCA showing the difference between pancreatic cyst fluid and 13 different body site microbiome selected from Human Microbiome Project database. When compared 136 bacterial genus with  $p < 0.01$  showing high (54) and low (82) abundance distribution between the PCF and 13 body site selected. This image constitutes the comparison between the PCF and 13 body site microbiomes (principal component analysis).



**Figure S1. A.** PCA of antecubital fossa and pancreatic cyst fluids microbiome.

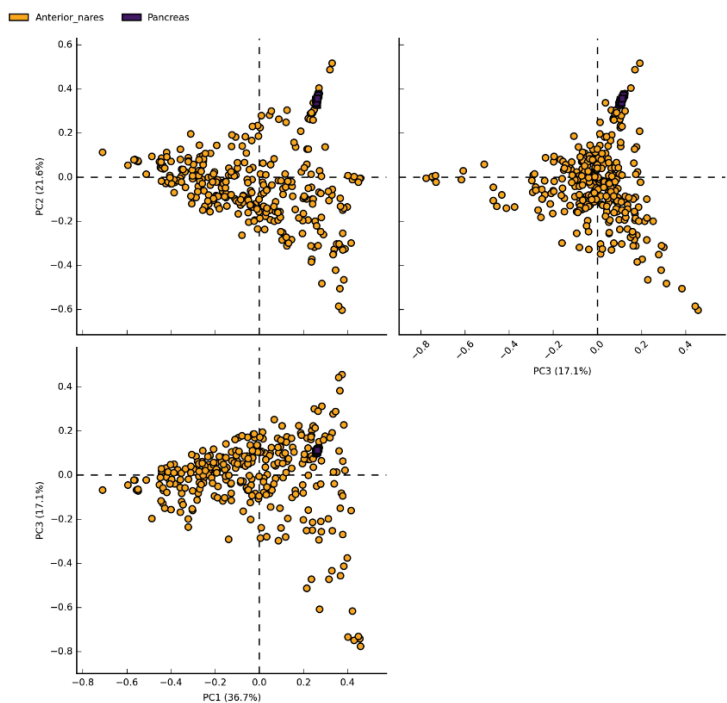


Figure S1. B. PCA of anterior nares and pancreatic cyst fluids microbiome.

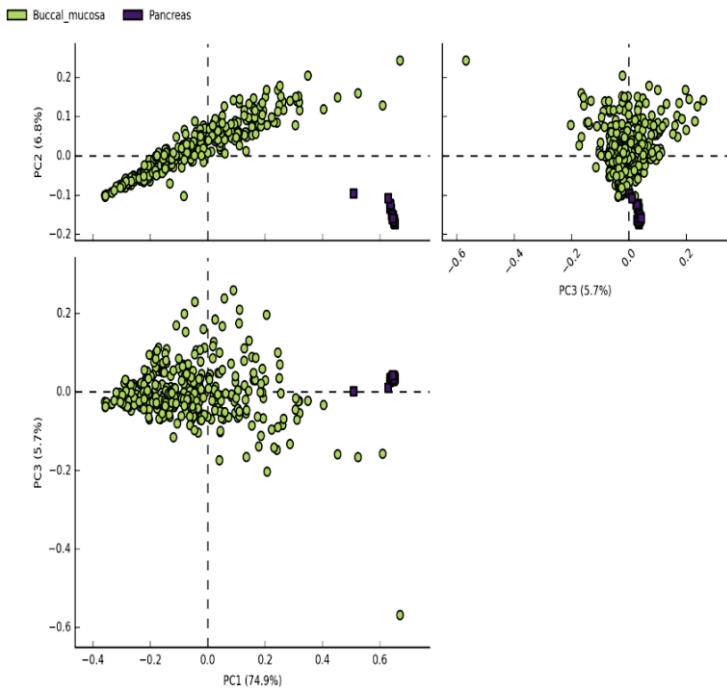


Figure S1. C. PCA of buccal mucosa and pancreatic cyst fluids microbiome.

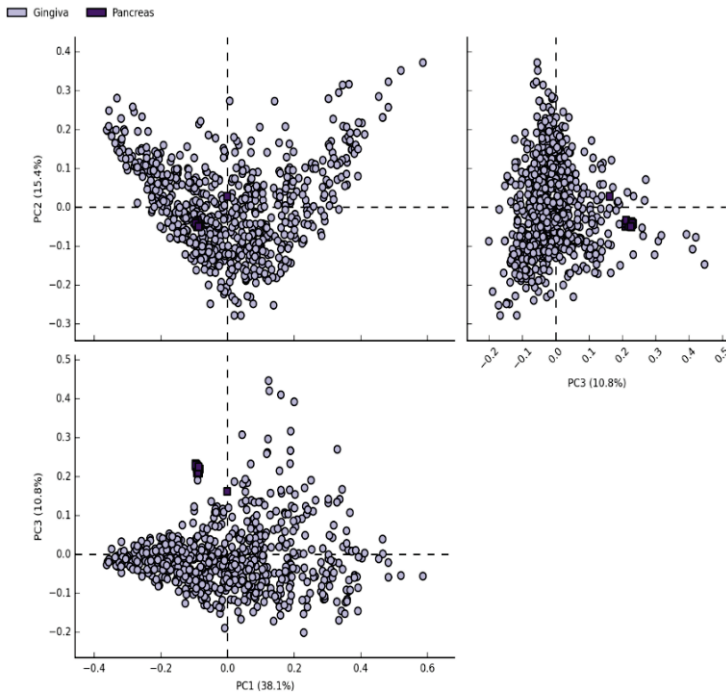


Figure S1. D. PCA of gingiva and pancreatic cyst fluids microbiome.

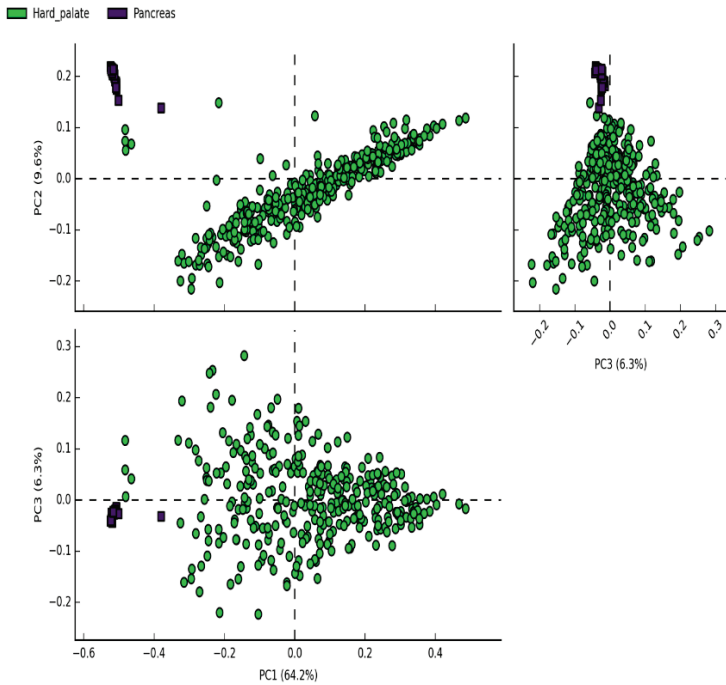


Figure S1. E. PCA of hard palate and pancreatic cyst fluids microbiome.

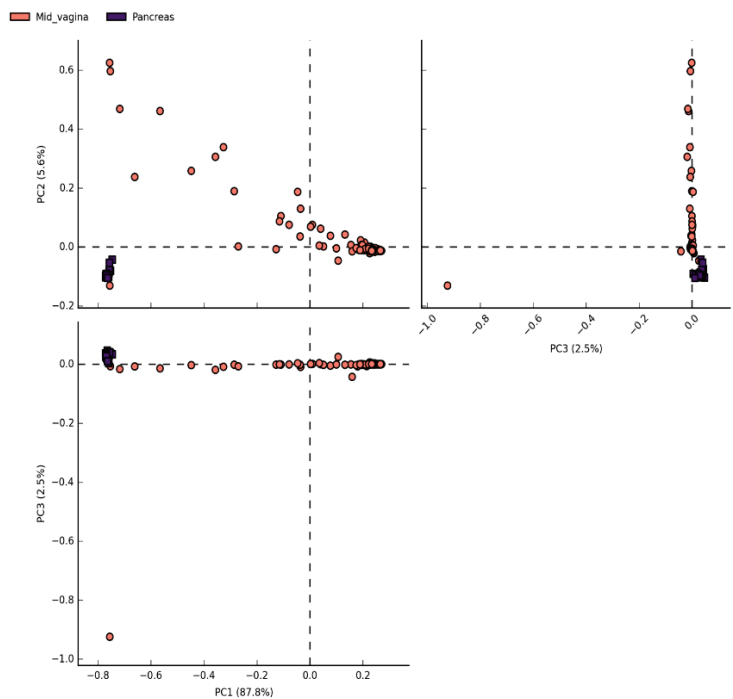


Figure S1. F. PCA of mid-vagina and pancreatic cyst fluids microbiome.

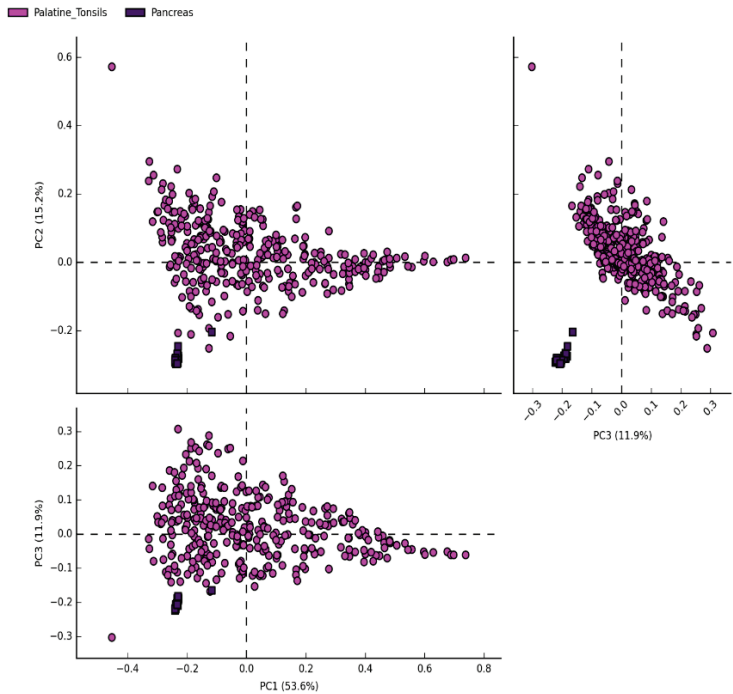


Figure S1. G. PCA of posterior fornix and pancreatic cyst fluids microbiome.

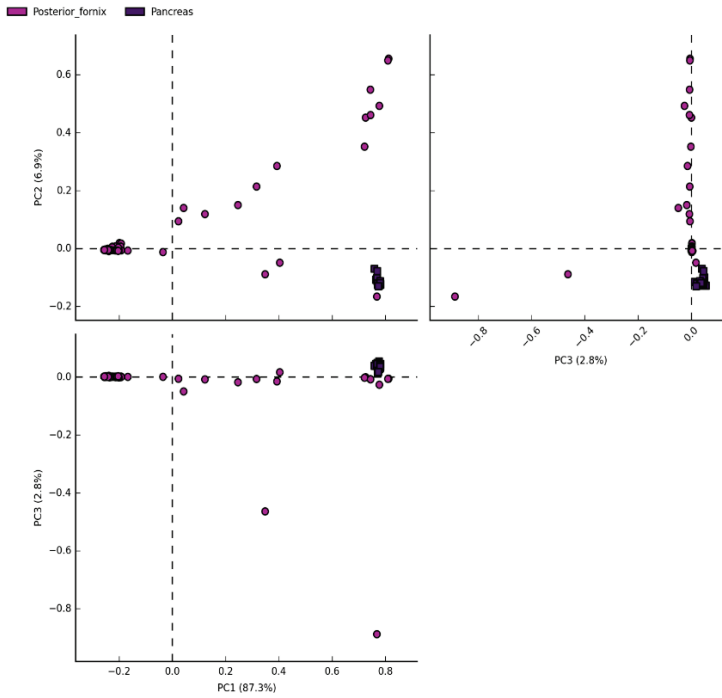


Figure S1. H. PCA of palatine tonsils and pancreatic cyst fluids microbiome.

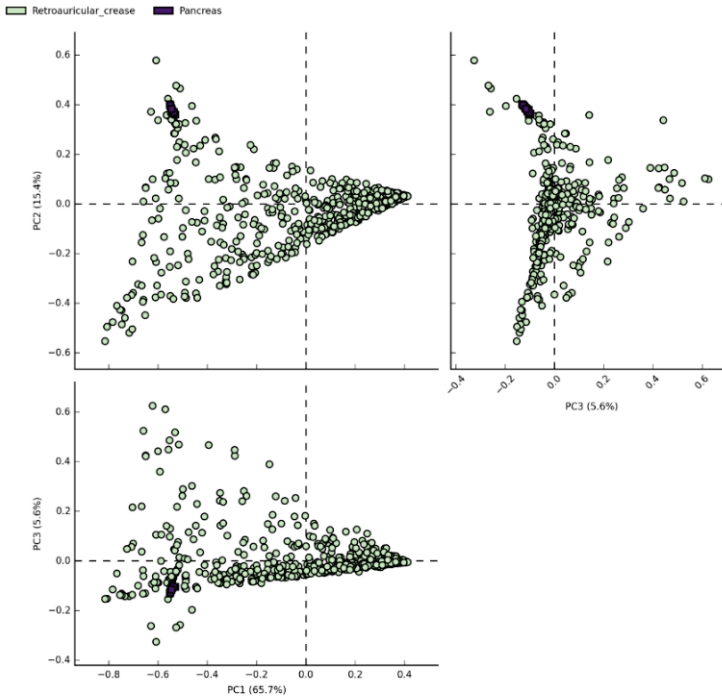


Figure S1. I. PCA of retroauricular crease and pancreatic cyst fluids microbiome.

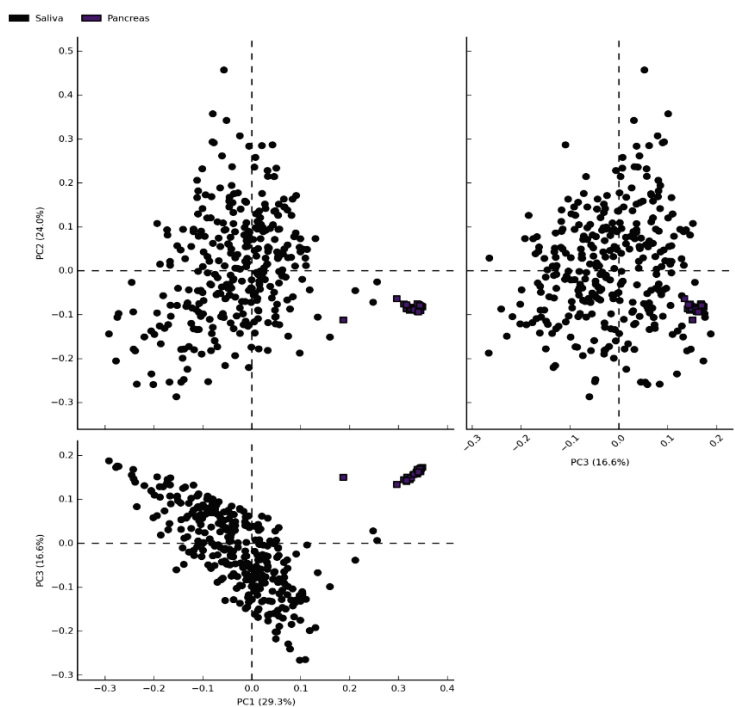


Figure S1. J. PCA of stool and pancreatic cyst fluids microbiome.

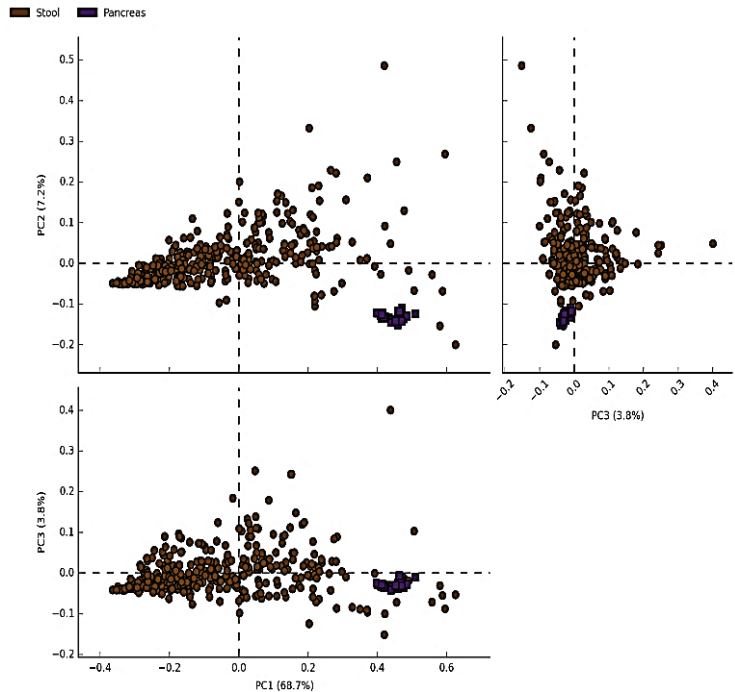
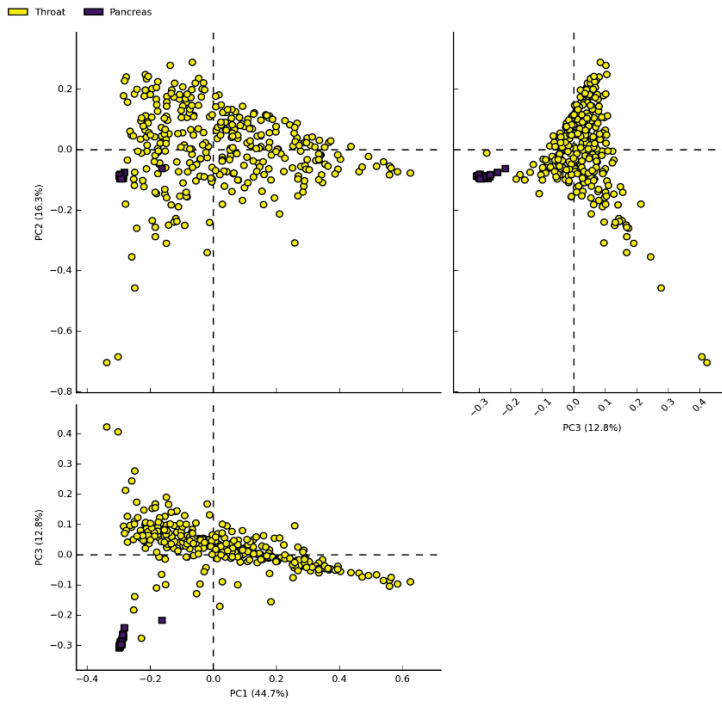
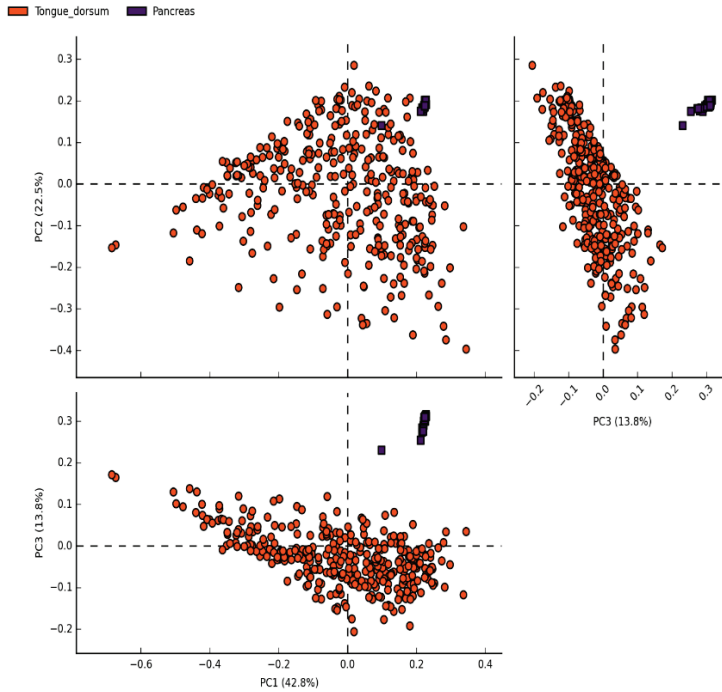


Figure S1. K. PCA of saliva and pancreatic cyst fluids microbiome.

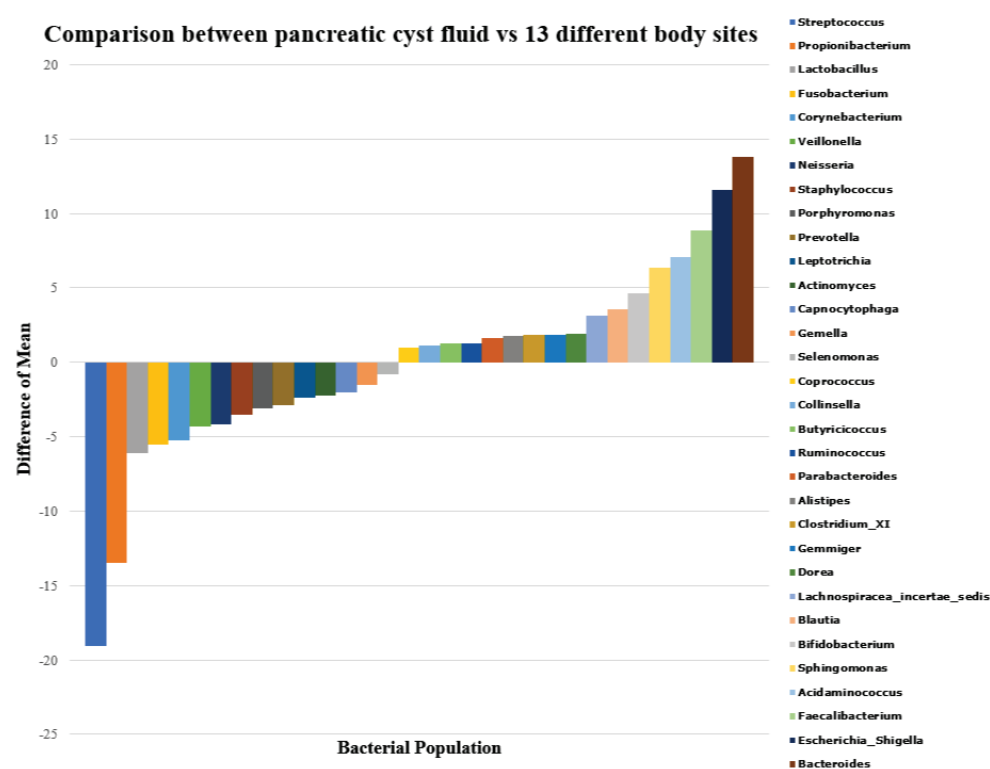


**Figure S1. L.** PCA of tongue dorsum and pancreatic cyst fluids microbiome.



**Figure S1. M.** PCA of throat and pancreatic cyst fluids microbiome.





**Additional file 6: Figure S2.** Difference of mean of selected high and low abundance bacterial microbiome in PCF and 13 body sites together, respectively. Difference of mean between the bacterial genus distribution of pancreatic cyst fluid and 13 different body site microbiome selected from Human Microbiome Project database. When compared 17 bacterial genus (with  $p < 0.01$ ) showing high abundance in PCF and 15 bacterial genus (with  $p < 0.01$ ) showing high abundance in 13 body sites selected.

**Additional file 7:** Welch's  $t$  test statistics run for the bacterial genotypes comparison calculation of the PCF and 13 different body sites microbiota. Excel file containing data is available online at:

[https://static-content.springer.com/esm/art%3A10.1186%2F40168-017-0363-6/MediaObjects/40168\\_2017\\_363\\_MOESM7\\_ESM.xlsx](https://static-content.springer.com/esm/art%3A10.1186%2F40168-017-0363-6/MediaObjects/40168_2017_363_MOESM7_ESM.xlsx)

## References

1. H. Matthaei, R. D. Schulick, R. H. Hruban, A. Maitra, Cystic precursors to invasive pancreatic cancer. *Nat. Rev. Gastroenterol. Hepatol.* **8**, 141–150 (2011).
2. B. U. Wu *et al.*, Prediction of malignancy in cystic neoplasms of the pancreas: a population-based cohort study. *Am. J. Gastroenterol.* **109**, 121–129 (2014).
3. R. Grutzmann, M. Niedergethmann, C. Pilarsky, G. Kloppel, H. D. Saeger, Intraductal papillary mucinous tumors of the pancreas: biology, diagnosis, and treatment. *Oncologist* **15**, 1294–1309 (2010).
4. J. J. Farrell *et al.*, Variations of oral microbiota are associated with pancreatic diseases including pancreatic cancer. *Gut* **61**, 582–588 (2012).
5. J. J. Farrell, C. Fernandez-del Castillo, Pancreatic cystic neoplasms: management and unanswered questions. *Gastroenterology* **144**, 1303–1315 (2013).
6. H. Matthaei *et al.*, miRNA biomarkers in cyst fluid augment the diagnosis and management of pancreatic cysts. *Clin. Cancer Res.* **18**, 4713–4724 (2012).
7. M. S. Sawhney *et al.*, International consensus guidelines for surgical resection of mucinous neoplasms cannot be applied to all cystic lesions of the pancreas. *Clin. Gastroenterol. Hepatol.* **7**, 1373–1376 (2009).
8. B. K. Goh *et al.*, Evaluation of the Sendai and 2012 International Consensus Guidelines based on cross-sectional imaging findings performed for the initial triage of mucinous cystic lesions of the pancreas: a single institution experience with 114 surgically treated patients. *Am. J. Surg.* **208**, 202–209 (2014).
9. M. Tanaka *et al.*, International consensus guidelines for management of intraductal papillary mucinous neoplasms and mucinous cystic neoplasms of the pancreas. *Pancreatology* **6**, 17–32 (2006).
10. K. de Jong *et al.*, Accuracy of preoperative workup in a prospective series of surgically resected cystic pancreatic lesions. *Scand. J. Gastroenterol.* **47**, 1056–1063 (2012).
11. A. Khalid, W. Brugge, ACG practice guidelines for the diagnosis and management of neoplastic pancreatic cysts. *Am. J. Gastroenterol.* **102**, 2339–2349 (2007).
12. K. E. Monkemuller *et al.*, Biochemical analysis of pancreatic fluid collections predicts bacterial infection. *J. Gastroenterol. Hepatol.* **20**, 1667–1673 (2005).
13. J. H. Kim *et al.*, Intraductal papillary mucinous neoplasms with associated invasive carcinoma of the pancreas: imaging findings and diagnostic performance of MDCT for prediction of prognostic factors. *AJR Am. J. Roentgenol.* **201**, 565–572 (2013).
14. A. Khalid *et al.*, Pancreatic cyst fluid DNA analysis in evaluating pancreatic cysts: a report of the PANDA study. *Gastrointest. Endosc.* **69**, 1095–1102 (2009).
15. J. Qin *et al.*, A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).
16. M. Arumugam *et al.*, Enterotypes of the human gut microbiome. *Nature* **473**, 174–180 (2011).
17. S. R. Konstantinov, E. J. Kuipers, M. P. Peppelenbosch, Functional genomic analyses of the gut microbiota for CRC screening. *Nat. Rev. Gastroenterol. Hepatol.* **10**, 741–745 (2013).
18. O. Franzén, J. Hu, X. Bao, S. H. Itzkowitz, I. Peter, A. Bashir, Improved OUT-picking using long-read 16S rRNA gene amplicon sequencing and generic hierarchical clustering. *Microbiome* **3**, 43 (2015).
19. R. C. Edgar, Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
20. Z. K. Wang, Y. S. Yang, Upper gastrointestinal microbiota and digestive diseases. *World J. Gastroenterol.* **19**, 1541–1550 (2013).
21. C. Quast *et al.*, The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
22. D. H. Parks, R. G. Beiko, Identifying biologically relevant differences between metagenomic communities. *Bioinformatics* **26**, 715–721 (2010).
23. J. Peterson *et al.*, The NIH Human Microbiome Project. *Genome Res.* **19**, 2317–2323 (2009).
24. J. R. White, N. Nagarajan, M. Pop, Statistical methods for detecting differentially abundant features in clinical metagenomic samples. *PLoS Comput. Biol.* **5**, e1000352 (2009).

25. R. Talar-Wojnarowska *et al.*, Pancreatic cyst fluid analysis for differential diagnosis between benign and malignant lesions. *Oncol. Lett.* **5**, 613–616 (2013).
26. R. Schwabe, C. Jobin, The microbiome and cancer. *Nat. Rev. Cancer* **13**, 800–812 (2013).
27. V. Di Pilato, G. Freschi, M. N. Ringressi, L. Pallecchi, G. M. Rossolini, P. Bechi, The esophageal microbiota in health and disease. *Ann. N. Y. Acad. Sci.* **1381**, 21–33 (2016).
28. V. D'Argenio *et al.*, Metagenomics reveals dysbiosis and a potentially pathogenic *N. flavescens* strain in duodenum of adult celiac patients. *Am. J. Gastroenterol.* **111**, 879–890 (2016).
29. K. Mitsuhashi *et al.*, Association of *Fusobacterium* species in pancreatic cancer tissues with molecular features and prognosis. *Oncotarget* **6**, 7209–7220 (2015).
30. I. Lin *et al.*, Pilot study of oral microbiome and risk of pancreatic cancer. *Cancer Res.* **73**(8 Suppl), Abstract nr 101 (2013).
31. S. Olson *et al.*, The oral microbiota in patients with pancreatic cancer, patients with IPMNs, and controls: a pilot study. *Cancer Causes Control* **28**, 959–969 (2017).
32. R. Verma, R. Dhamija, S. C. Ross, D. H. Batts, M. E. Loehrke, Symbiotic bacteria induced necrotizing pancreatitis. *JOP.* **11**, 474–476 (2010).
33. I. Brook, E. H. Frazier, Microbiological analysis of pancreatic abscess. *Clin Infect Dis.* **22**, 384–385 (1996).
34. B. Zulfikaroglu, M. Koc, N. Ozalp, *Candida albicans*-infected pancreatic pseudocyst: report of a case. *Surg. Today* **34**, 466–469 (2004).
35. J. Cherenfant, M. Nikfarjam, A. Mathew, E. T. Kimchi, K. F. Staveley-O'Carroll, Completion pancreatectomy for treatment of a *Clostridium perfringens* pancreatic infection. *Arch. Surg.* **144**, 368–370 (2009).

# Chapter 3

## **Bacterial Biofilms in Colorectal Cancer Initiation and Progression**

Shan Li<sup>1</sup>, Sergey R. Konstantinov<sup>1</sup>, Ron Smits<sup>1</sup>, Maikel P. Peppelenbosch<sup>1</sup>

<sup>1</sup>*Department of Gastroenterology and Hepatology, Erasmus MC-University Medical Center, Rotterdam, The Netherlands.*

***Trends in Molecular Medicine.*** 2017, 23(1):18-30.



## Abstract

Intestinal microbiota have emerged as an important factor in colorectal cancer (CRC) initiation and progression. The currently prominent view on bacterial tumorigenesis is that CRC initiation is triggered by local mucosal colonization with specific pathogens (drivers), and that subsequent changes in the peritumoral environment allow colonization by opportunistic (passenger) microbes, further facilitating disease progression. Screening for CRC ‘driver-passenger’ microorganisms might thus allow early CRC diagnosis or preventive intervention. Such efforts are now being revolutionized by the notion that CRC initiation and progression require organization of bacterial communities into higher-order structures termed biofilms. We explore here the concept that a polymicrobial biofilm promotes pro-carcinogenic activities that may partially underlie progression along the adenoma–CRC axis.

**Keywords:** Biofilm; Colorectal cancer; Microbiome; Tumor initiation; Progression.

## Trends

- The organization of bacterial communities into biofilms (higher-order spatial structures of bacterial species) may be necessary for bacteria-induced CRC initiation.
- The interaction of the intestinal epithelium with the microbiota is highly dependent on the nature of the spatial organization of bacterial communities.
- Bacterial biofilms might act as direct triggering factors contributing to colorectal cancer.
- The biofilm confers highly-invasive properties to opportunistic bacteria, and a putative tumor-promoting potential.
- In experimental models, biofilm microbial populations can significantly impair the intestinal epithelial barrier function, alter polyamine metabolism affecting cellular proliferation, enhance pro-inflammatory/pro-oncogenic responses, and exacerbate intestinal dysbiosis.
- The invasive and co-aggregation capacity of microbiota may be essential for biofilm-promoted colon tumorigenesis.

## **The intestinal microbiome: a new window to studying colon carcinogenesis**

Nowhere in the human body are interactions between the **microbiome** (see Glossary at the end of this chapter) and host physiology as pronounced as in the gastrointestinal (GI) tract. It hosts an estimated 40 trillion microbes composed of at least 1000 species of which the vast majority reside in the colon (1). It is thus to be expected that, if the microbiome and human pathophysiology are interlinked, this should be especially pronounced in the colon. A principal role for the microbiome in the pathophysiology of acute and chronic inflammatory diseases of the colon is indeed well established (2,3). Intriguingly, despite the relatively negligible bacterial colonization of the stomach, the link between gastric cancer and *Helicobacter pylori* infection is beyond discussion. Emerging evidence suggests that other microbiota colonizing the stomach might also be involved in human gastric cancer progression (4,5). However, a causal link between colorectal cancer (CRC) and the microbiome has been less evident. Recently, however, this field has moved forward by linking colonic intestinal microbiota to CRC progression (6–13). Such studies have generated high expectations that screening for microbiological constituents might provide early diagnosis of CRC, or that disease might be potentially prevented through dietary or other interventions that could modulate colonic microbiome composition. In view of the substantial challenge that CRC poses to society, such efforts are considered to be of high importance (14,15).

It is now becoming clear that CRC is not attributable to a single pathogenic microorganism, and instead that it requires a complex intestinal bacterial community. Studies comparing fecal matter from patients with CRC relative to healthy controls have demonstrated substantial differences in human gut microbiome composition (16,17). These studies also show that CRC is characterized by **microbial dysbiosis** (12,18). Furthermore, recent work has shown that mice deficient in the immune sensor **absent in melanoma 2 (Aim2)**, and which were colonized with dysbiotic gut microbiota, are highly susceptible to tumorigenesis in the colon compared to the same mice colonized by healthy microbiota (19). Important questions in the field involve

the extent to which these changes are a cause or a consequence of CRC, in addition to identifying the mechanisms mediating such changes. Currently, various models for bacteria-induced carcinogenesis have been postulated, suggesting how intestinal microbiota as well as microbe–microbe and microbe–host interactions contribute to CRC (20–22). Nevertheless, the mechanisms by which the intestinal microbiota interact with themselves and the human host to induce CRC initiation and progression remain largely obscure. It is becoming clear that these CRC-eliciting interactions are highly dependent on the nature and spatial organization of multispecies bacterial communities in higher-order structures (termed **biofilms**) (11,13). In this opinion article we argue that polymicrobial biofilms promote pro-carcinogenic activities, and that invasive biofilm appears to be indispensable for CRC initiation.

### **From single pathogenic microorganism to polymicrobial infections and cancer**

As shown in the quintessential example of *H. pylori* in gastric cancer, specific microorganisms *per se* are capable of driving carcinogenic and other cancerous processes in the human GI tract (23–26). In addition, other types of human cancer can also be provoked by infection with a specific pathogen, examples being liver cancer (chronic hepatitis B or C virus), cervical cancer (human papilloma virus), Burkitt's lymphoma (Epstein–Barr virus), and bladder cancer (induced by *Schistosoma haematobium*) (20,21,27,28).

Regarding human CRC, two North American studies in 2012 showed over-representation of *Fusobacterium nucleatum* in CRC tumors compared to the surrounding normal tissue (7,9). This bacterium was linked to CRC development, as evidenced by its capacity to invade the colonic mucosa, induce local inflammation and increased expression of cytokines, such as interleukin 6 (IL-6), IL-8, IL-12, transforming growth factor  $\beta$  (TGF- $\beta$ ), and tumor necrosis factor  $\alpha$  (TNF- $\alpha$ ), thereby potentially exacerbating CRC (10,29,30). *F. nucleatum* bacterial infection directly contributes to colorectal carcinogenesis, as evidenced by two recent reports using the



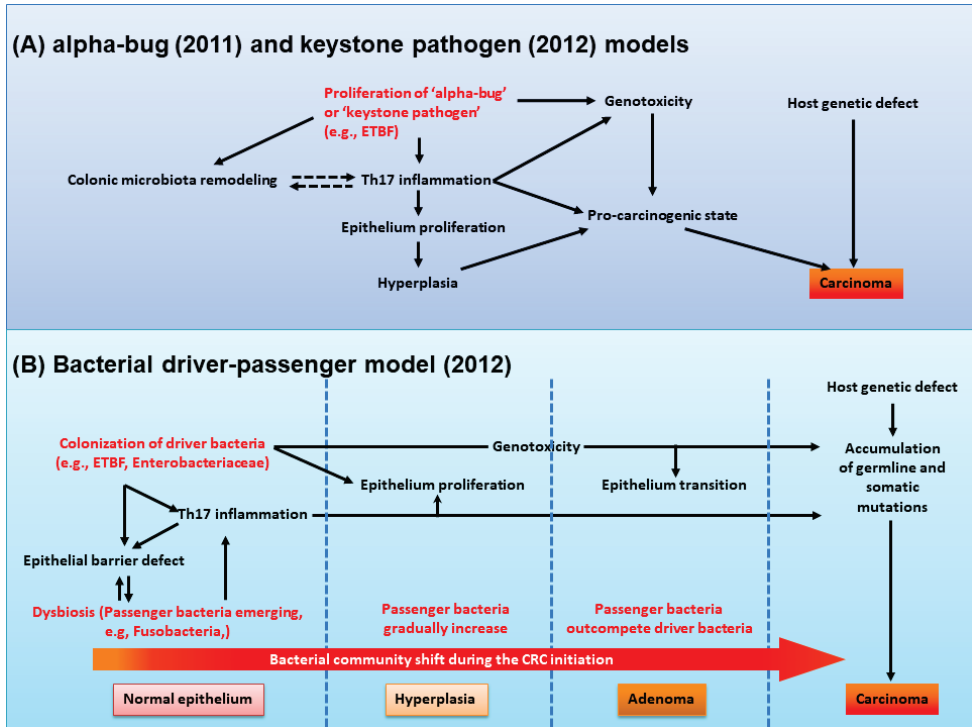
*Apc*<sup>Min/+</sup> **mouse** model and human CRC cell lines; the data show that *F. nucleatum* invasion results in the recruitment of tumor-infiltrating CD11b<sup>+</sup> (integrin subunit  $\alpha$ M, ITGAM) immune cells in the colon and establishing an oncogenic/pro-inflammatory microenvironment (10,31). Moreover, this recruitment process seemed to depend on bacterial FadA (adhesin fatty acid degradation A)-mediated adhesion (10,31). FadA has been previously shown to bind to **epithelial cadherin** (**E-cadherin**; also known as cadherin 1, CDH1), leading to activation of  $\beta$ -catenin signaling (31). With regard to *F. nucleatum*, a recent human study reported that a higher abundance of *F. nucleatum* was correlated with fewer CD3<sup>+</sup> T cells in CRC biopsies, suggesting a putative immunosuppressive effect linked to CRC progression (32). Thus, it is reasonable to suppose that *F. nucleatum* may play a relevant role in CRC initiation/progression.

Of note, several other pathogenic bacteria of relatively low abundance in the colonic microbiota may exhibit pro-oncogenic activity in CRC development via the action of unique virulence factors (33). As an example, enterotoxigenic *Bacteroides fragilis* (ETBF) has been proposed to be a **keystone pathogen** in CRC initiation; although ETBF typically comprises a small proportion (~1 to 2%) of the human fecal bacterial community it causes significant pro-carcinogenic effects linked to its abundance (20,21). The *B. fragilis* toxin, a metalloprotease toxin secreted by ETBF, can lead to the recruitment of **type 17 T helper (Th17) cells**, which in turn elicit rapid and robust inflammatory responses characterized by the production of **genotoxic** oxygen radicals, while concomitantly depressing T cell-mediated **tumor immune surveillance** through the selective activation of the signal transducer activator of transcription 3 (STAT3)-dependent pathway in the human colon (20,21,34,35). It is thus possible that some bacteria exhibiting invasive behavior similar to that of ETBF could be a major cause of CRC. In this context, the ‘alpha-bug’ hypothesis should be mentioned. This hypothesis – postulated by Sears and Pardoll (21) – together with a similar hypothesis concerning the ‘keystone pathogen’ proposed by Hajishengallis and colleagues (20), posit that the ‘alpha-bug’ or ‘keystone pathogen’, such as ETBF, not only induces direct genotoxicity in intestinal epithelial cells but also provokes remodeling in the composition of the colonic microbiota. By outcompeting anti-

oncogenic **symbionts**, such as *Lactobacilli* and *Bifidobacteria*, pro-oncogenic Th17 immune responses are evoked (**Figure 1**). In this line of reasoning particular ‘alpha-bugs’ or ‘keystone pathogens’ could be considered to be a major cause of CRC carcinogenesis, not unlike the role of *H. pylori* in gastric carcinogenesis.

Intriguingly, these keystone pathogens potentially disappear during oncogenic transformation from normal epithelium to a neoplastic lesion during CRC progression because they are outcompeted by opportunistic bacteria better adapted to the tumor microenvironment of human CRC (7,9,22,33). Tjalsma and colleagues (22) denominated this process the **bacterial driver-passenger** model (**Figure 1**). In this model, indigenous pathogenic bacteria colonizing the large intestine produce genotoxins termed ‘bacterial drivers’. Examples are *B. fragilis* toxin (BFT, also called fragilysin), colibactin, and cytolethal distending toxin (CDT). Alterations in the tumor microenvironment caused by a driver pathogen presumably induce pro-inflammatory responses that are able to contribute to selective pressure and induce subsequent changes in the abundance of intrinsic pathogenic members (ETBF or other driver pathogens) in the colonic microbial community. This in turn could lead to an increased number of opportunistic pathogens (the passenger pathogens), such as *Fusobacterium* spp. and *Streptococcus* spp., that gradually outcompete the driver pathogen within the entire microbial community, resulting in further progression towards CRC (7,9,22,33). The ‘driver pathogen’ ETBF, for instance, might initially colonize the intestinal mucosa and subsequently induce a Th17-dependent immune response. This would result in increased proliferation in the intestinal epithelium, for instance by activating K-Ras proto-oncogene (*KRAS*) and B-Raf proto-oncogene (*BRAF*), or through loss-of-function mutations in tumor-suppressor genes such as adenomatous polyposis coli (*APC*). Concomitantly, the altered microenvironment might then allow ‘passenger pathogens’, such as *Fusobacterium* spp. or *Streptococcus* spp., to colonize the mucosa, promoting CRC progression (22,36) and outcompeting ETBF. In the end, the composition of human and murine mucosa-associated microbiota may be dramatically altered (22,33,34). In support, studies have reported spatial variation of microbial diversity between luminal and mucosal microbiota in both normal and inflamed human and murine intestine (37,38). The spatial

heterogeneity of gut microbiota in lumen and mucosa might favor alterations in microbial diversity and abundance in the inflamed colon, which might then proceed to CRC.



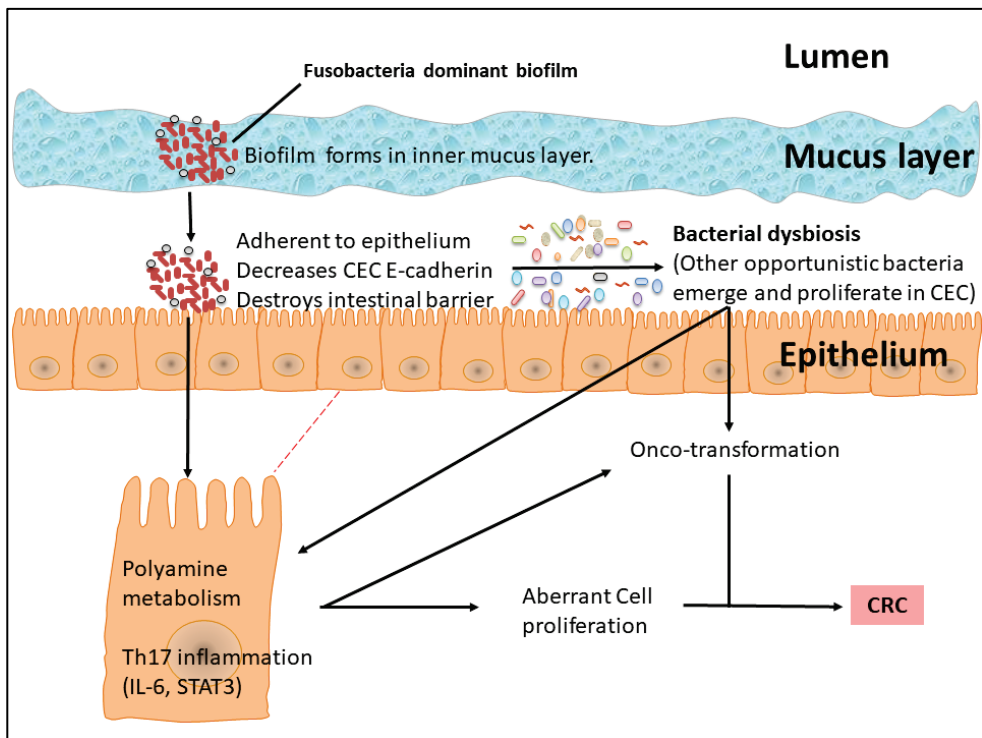
**Figure 1.** Schematic Depiction of Prominent Hypotheses for Bacteria-Induced Colon Carcinogenesis. **(A)** The pro-carcinogenic state of intestinal epithelial cells is driven by the so-called ‘alpha-bug’ or ‘keystone pathogen’, such as enterotoxigenic *Bacteroides fragilis* (ETBF), and is not merely related to the production of genotoxins that directly cause DNA damage, type 17 T helper cell (Th17)-dependent inflammatory responses, and epithelial proliferation (among other characteristics). The hypothesis posits that remodeling of the microbiota composition in the mucosa is the main driving force that enhances and sustains inflammation and epithelial proliferation. Collectively, all these factors contribute to the accumulation of genetic mutations and epithelial hyperplasia. The hyperplastic epithelium together with host genetic factors linked to colorectal cancer (CRC) susceptibility and other pro-oncogenic determinants may cooperatively initiate colon cancer. **(B)** The ‘bacterial driver-passenger’ hypothesis states that colonization of the colon by putative driver bacteria (e.g., ETBF) that possess invasive characteristics can trigger CRC initiation. The bacterial driver also causes an epithelial barrier defect, inducing the subsequent colonization of opportunistic (passenger) pathogens in the colonic mucosa, and leading to intestinal microbial dysbiosis. Th17-dependent inflammation induced by the driver pathogen may also result in alterations in the tumor microenvironment that might contribute to selective stress and subsequent changes to the colonic microbial community. In this model, driver pathogens are outcompeted by passenger pathogens, causing in turn further CRC progression.

## Biofilm organization may be necessary for bacteria-induced CRC initiation

Ecological constraints in the human intestine favor a higher-order level of spatial organization of multi-organism structures in mucosal microbial communities, of which biofilm formation is the best-known (11,13,39). Such biofilms are associated with chronic bacterial infection that is not always easy to eradicate. Biofilms appear to be an important etiological factor in human infectious disease, especially endocarditis and cystic fibrosis lung disease (40,41). With respect to the former, the adherence of *Streptococcus* spp. to extracellular matrix proteins of human endothelial cells and subsequent biofilm formation have been associated with endocarditis (41). Recently, biofilms have been linked to CRC initiation and development – especially in the human right colon (here defined as proximal colon to the hepatic flexure) (11). Biofilms are much more prevalent in colonic tissue samples of CRC patients than of healthy controls (11). An emerging possibility is that the biofilms comprise different bacterial species, rather than solely a solitary invading organism, and might result in enhanced inflammatory responses and the production of genotoxic bacterium-derived compounds. Accordingly, *Fusobacteria*-dominant polymicrobial biofilms were shown to be abundantly present in tumor samples from patients with colorectal **adenoma** and cancer but not in paired tumor-free tissue (11). Biofilms of driver bacteria may create novel ecological niches for passenger bacteria in CRC development, eventually outcompeting the driver organism. According to the **adenoma-carcinoma sequence model** proposed by Fearon and Vogelstein (42), microbial biofilm may be regarded as an independent ‘driver’ at an early stage of CRC carcinogenesis, before the malignant transformation from adenoma to carcinoma occurs (**Box 1**). Various processes mediating the action of biofilms in driving the CRC process have been proposed, and these include the presence of reduced levels of E-cadherin in human **intestinal crypts**, increased intestinal permeability, the production of **polyamine metabolites** and their subsequent acetylation, and elevated induction of IL-6/ STAT3 signaling (**Figure 2, Key Figure**) (11,13). To promote CRC carcinogenesis, microbial pathogens may thus require biofilms.

**Box 1. Biofilm-Driven Bacterial Tumorigenesis in the Colon and Rectum**

Invasive polymicrobial biofilms appear to be essential for bacteria-associated tumorigenesis (Figure 2) (11). During biofilm-related CRC, initially, invasive bacteria such as *Fusobacteria* may form biofilms that directly and dramatically invade the intestinal epithelium. This biofilm invasion may result in increased intestinal epithelial permeability, enhanced bacterial dysbiosis, induced inflammatory responses, and alterations in the host/bacterial metabolome. Collectively, these pro-oncogenic events might provide a genotoxic and CRC-permissive microenvironment, while simultaneously local ecological niches would be opened that allow colonization by other microorganisms (11,13). As such, the initial biofilm-forming organism could be considered the driver (or ‘alpha-bug’) in colorectal carcinogenesis. New organisms exploiting the ecological niches created might aggravate the oncogenic process through their specific interactions with the host.

**Key Figure: The Biofilm-Driven Colorectal Cancer (CRC) Carcinogenesis Model**

**Figure 2.** A biofilm composed of multiple-species forms in the inner colonic mucus layer, leading to redistribution of E-cadherin in colonic epithelial cells (CECs), increased gut permeability, and loss of intestinal barrier function, in turn enhancing intestinal dysbiosis. Dysbiosis may favor increased proliferation of other opportunistic invasive bacteria (i.e., passengers). The pro-oncogenic functions of the bacterial biofilm, and subsequently of passenger bacteria, in combination with bacterial polyamine metabolic changes and type 17 T helper cell (Th17)-mediated inflammation, results in cell proliferation, transformation, and aberrant tumor growth.

## **Biofilms require invasion and coaggregation properties to stimulate CRC**

Bacterial biofilms are not carcinogenic *per se* but only in the context of specific invasive bacteria, especially *Fusobacteria* (11). Accordingly, biofilms have been observed in the rat, baboon, and human non-tumorous gut by electron microscopy (43). One study also showed that biofilms could be detected in the colon of normal mice, and biofilms have been sampled from healthy individuals using colonoscopy (38). Moreover, colonoscopy biopsy specimens from healthy individuals have revealed that thin biofilms in the mucosa consist of relatively harmless bacteria, particularly *Bacteroidetes*, *Lachnospiraceae*, and *Enterobacteriaceae* in the right colon, and *Bacteroidetes* and *Lachnospiraceae* – but notably, no *Fusobacterium* – in the left colon (11). Most, if not all, of these bacterial communities detected in normal mucosa in this study are commensal species devoid of invasive capability. Thus, the invasive potential of biofilm-forming bacteria might be relevant to CRC pathogenesis. In support of this hypothesis, although *Fusobacterium* spp. is a relatively common and harmless opportunistic biofilm-forming pathogen in the oral cavity, in the intestine such biofilms can provoke severe inflammation (44,45). Furthermore, co-aggregation of *F. nucleatum* facilitates colonization with other bacterial species in the biofilm (46), suggesting that the formation of biofilms might indeed provide novel ecological niches. *F. nucleatum* may act as a robust agent in the process of biofilm formation. For example, biofilm cultures *in vitro* have shown that this bacterium supplies a niche for *Tannerella forsythia* adherence and growth because *F. nucleatum* produces a favorable environment for obligate anaerobes. Hence, *F. nucleatum* may recruit other types of bacteria into biofilms (47).

*Campylobacter* species such as *C. concisus*, *C. rectus*, and *C. curvus* are harmless colonizers of the human oral cavity, but are significantly associated with esophageal and colonic adenocarcinoma (48). Some oral *C. concisus* strains appear to produce zonula occludens toxin (ZOT) to induce cytoskeletal remodeling and to disassemble tight junctions of intestinal epithelial cells, promoting bacterial translocation and inflammation (49). *Campylobacter* spp. are also capable of forming biofilm (48,50).

For example, *Campylobacter* spp., especially *C. showae*, colocalize with *Fusobacterium* and *Leptotrichia* species in human CRC tissues (50) and may increase the risk of **inflammatory bowel disease (IBD)** (51). *Leptotrichia* spp. have been found to be considerably more abundant in the stomach of individuals from populations with high risk of gastric cancer compared to individuals from low-risk populations, suggesting that *Leptotrichia* spp. may also have the potential to induce human intestinal cancer (52). Thus, a polymicrobial signature of Gram-negative anaerobic bacteria including *Fusobacterium*, *Campylobacter*, and *Leptotrichia* appears to be significantly associated with CRC (50). In addition, there is an association between *Streptococcus gallolyticus* (formerly known as *Streptococcus bovis*) and CRC. This bacterium has been observed in 20–50% of human colon tumor samples, but fewer than 5% of samples collected from healthy individuals presented this bacterium (12). Recent studies have shown that *Streptococcus* spp. have adherence potential and the ability to form biofilms (41,53–55). Collectively, these observations suggest that bacteria possessing invasion and co-aggregation properties might be required for the formation of tumor-promoting biofilms.

### **Biofilms enhance host–microbe interactions in CRC**

Bacterial biofilms can contribute to increases in intestinal permeability and enhanced bacterium-induced barrier function loss which, in turn, is one of the most important early pathophysiologic alterations in colorectal carcinogenesis (56). There are two lines of evidence supporting the biofilm-promoted barrier-loss concept. First, bacterial invasion is present in all biofilm-positive human colorectal tumors, including CRCs and adenomas, but invasive characteristics are absent from biofilm-negative colon tumors (11). Second, scanning electron microscope (SEM) imaging and fluorescence in situ hybridization (FISH) have unveiled that dense polybacterial biofilms can be detected in all right-sided human colon tumors, with bacterial communities being attached directly to the intestinal epithelial cell surface, but few biofilms have been found in left-sided colorectal tumors (11). Right-sided CRC patients typically present worse clinical outcomes than patients with left-sided CRC, which

may be related to this phenomenon. Biofilm-mediated growth allows bacteria to grow in close proximity to the **intestinal epithelial barrier**, an essential condition for bacterial invasion and for triggering subsequent inflammatory responses (11,36,57). Intestinal barrier loss also appears to exacerbate bacterial dysbiosis because failure to function might facilitate microbial adherence to epithelial cells; increased entry of bacterial products into epithelial cells might then contribute to creating a tumor-promoting environment, such as by activating Th17 immune responses, and favoring CRC initiation (11,36).

### **Biofilm formation and enhanced bacteria-mediated genotoxicity in CRC**

Of the possible mechanisms by which bacterial biofilms can promote oncological disease, genotoxic stress resulting from bacterial toxins appears to bear the most evident link to transformation *per se*. For instance, different bacteria produce various toxins, of which BFT and CDT contribute to genotoxicity and human CRC initiation (58–63). Indeed, BFT (generated from ETBF) is a genotoxin that indirectly causes DNA damage (34,60,64). For example, the BFT, in HT29/C1 and T84 colonic epithelial cell lines *in vitro*, has been reported to cause upregulation of spermine oxidase (SMO), leading to induced production of **reactive oxygen species (ROS)** (60). It has been suggested that increased ROS production activates the NLRP3 inflammasome [nucleotidebinding oligomerization domain (NOD)-like receptor containing pyrin domain 3], an important activator of innate immune responses that may further trigger DNA damage (65).

It has been proposed that BFT cleaves the intercellular adhesion molecule E-cadherin in human colonic epithelial cells (e.g., HT29/C1), thus compromising intestinal barrier function (66,67). The intestinal barrier deterioration may cause increased leakage of microbial products that could contribute to colonic pre-malignant lesions (adenoma) (36). Then, as shown in **CPC-APC mice**, this could lead to an IL-23/IL-17-involved inflammatory response, cause DNA damage, and ultimately induce tumor formation (36). Interestingly, *B. fragilis* biofilm is considered to be a key



characteristic of IBD, of which the tumor-promoting effects of persistent intestinal inflammation have now been recognized and are well understood (68). Indeed, more than 60% of the bacterial biofilm mass obtained from samples of IBD patients have been attributed to this *Bacteroides* species (68). Moreover, the *in vitro* formation rate of biofilm communities identified from IBD samples is significantly higher than that of samples taken from individuals either with self-limiting colitis or under non-inflamed conditions (68). It is thus reasonable to propose that biofilms comprising *B. fragilis* are especially potent with respect to promoting intestinal inflammation and CRC. We posit that polymicrobial biofilms might theoretically modulate the intestinal tumor microenvironment in multiple ways (including genotoxicity), and facilitate the development of colon cancer.

### **The biofilm influences host metabolism in CRC**

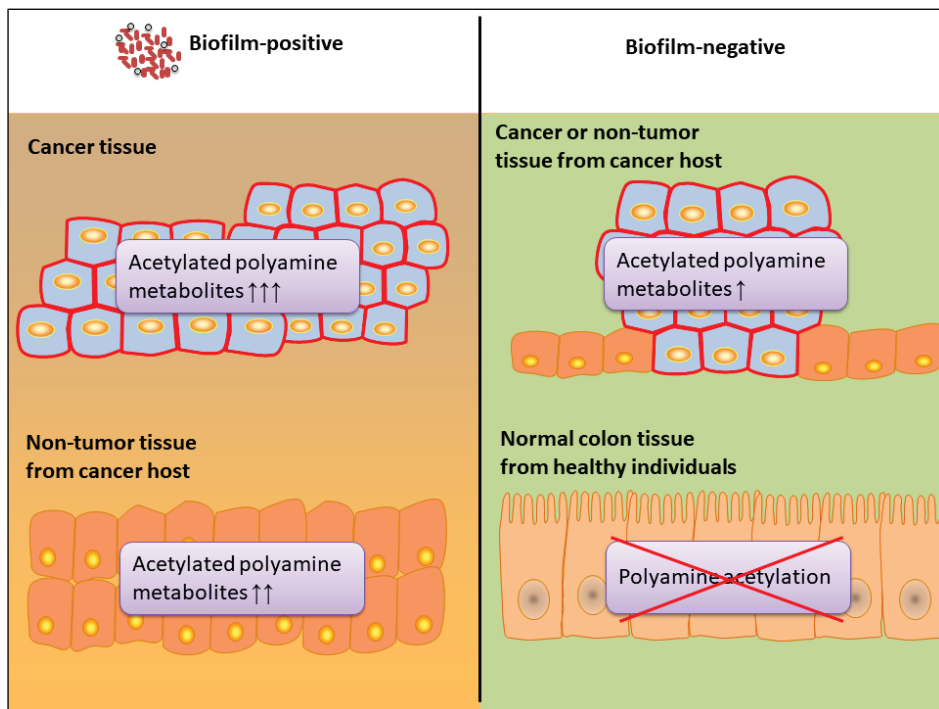
Microbe–host metabolic interactions may directly or indirectly provoke bacteria-induced CRC progression. Accumulating evidence has linked the intestinal microbiota to the regulation of multiple metabolic pathways of endogenous and exogenous substrates, such as secondary bile acid biosynthesis, polyamine catabolism, and activation of carcinogens, which in turn are associated with increased risk for various human cancers such as CRC, esophageal cancer, and liver cancer (13,60,69–75). As a good example, deoxycholic acid (DCA, also known as deoxycholate) appears to be the most important endogenous metabolite associated with CRC carcinogenesis in *Apc*<sup>Min/+</sup> mice and also in human colon biopsies (70,73). DCA can act as a naturally-occurring carcinogen that induces oncogenic transformation in digestive tract epithelium because DCA has been shown to induce **oxidative stress**, DNA damage, and **genomic instability** in human esophageal biopsy tissues, human colonic epithelial cancer cells, and in mouse models (70,74,76). Furthermore, DCA promotes tumor cell proliferation and inhibits apoptosis by activating Wnt signaling in the *Apc*<sup>Min/+</sup> mouse model (73). Because biofilms might conceivably provide a highly-efficient interface for bile acid deconjugation and dehydroxylation, cells of an epithelium covered by biofilm are likely to be exposed to very high concentrations of

secondary bile acids. Furthermore, such biofilms may also be sources of hydrogen sulfide ( $\text{H}_2\text{S}$ ) and nitrosamine. These compounds have been documented to be genotoxic and carcinogenic in rat and mouse models, as well as in human colon cancer cell lines (HT29, HCT116, and SW480), by inducing DNA damage and genomic instability (77–80). Thus, we hypothesize that the metabolism of biofilm constituents is likely to contribute to the carcinogenic potential of the underlying epithelium.

When non-adherent and biofilm bacteria are compared, higher levels of butyrate are produced by non-adherent bacterial communities, indicating that distinct bacterial communities displaying higher-order organization may display altered production of **short-chain fatty acids (SCFAs)** such as butyrate (81). *Fusobacterium*-dominant biofilm may decrease SCFA production and might contribute to promoting CRC tumorigenesis. However, microbe-derived SCFAs including butyrate may also promote hyperproliferation and sensitize colon epithelial cells to transformation, at least in **mismatch repair (MMR)**-deficient *Apc*<sup>Min/+</sup> mice (82). *Fusobacterium* is a major source of butyrate in the gut. Interestingly, *Fusobacterium* was not detected in this mouse model (82). Knowledge of the role of SCFAs in maintaining intestinal epithelium or inducing CRC remains limited, but altered production of these metabolites by biofilms certainly requires further investigation.

Emerging evidence suggests that bacterial biofilms might contribute to human CRC through polyamine synthesis and acetylation (**Figure 3**) (13). Multispecies biofilms express bacterial spermidine/spermine *N*<sup>1</sup>-acetyltransferase (SSAT) which is indispensable for polyamine acetylation (13). The importance of this observation is highlighted by data showing that expression of human SSAT is not important for cancer formation, and thus bacterial SSAT expression might make a difference in this context (13). Indeed, acetylated polyamines are substantially upregulated in both biofilm-covered human colon cancer and in paired normal tissues compared to colonic tissue devoid of biofilm (13), suggesting that biofilm enhances polyamine catabolism and acetylation, in turn inducing unwanted cell proliferation and cancer growth (60,83–85). Moreover, elevated levels of acetylated polyamine metabolites, including *N*<sup>1</sup>-acetylspermidine, *N*<sup>1</sup>-acetylspermine, and especially *N*<sup>1</sup>, *N*<sup>12</sup>-

diacetylspermine, are observed in human CRC compared to matched non-tumor tissues from the same patients (13). In agreement, polyamine metabolites at the mucosal edge (where the biofilm starts) of colonic cancer tissue samples have been found to emit stronger signals, as detected by nanostructure imaging mass spectrometry (NIMS), relative to the center of the **cancer cell nest** (13). This result further suggests that the biofilm may be the most important source of acetylated polyamine (13). In addition, targeted metabolomic analysis showed that no acetylated polyamine metabolites were present in biofilm-negative normal colonic biopsies from healthy individuals (13). The data further support the hypothesis that the presence of polymicrobial biofilm might be putatively linked to tumor growth in the colon through a mechanism involving elevated levels of acetylated polyamine metabolites.



**Figure 3.** Role of Acetylated Polyamine Metabolites in Colorectal Cancer (CRC). Schematic representation of the different levels of acetylated polyamine metabolites in (A) biofilm-positive colorectal cancer (CRC) patients and (B) biofilm-negative CRC patients or healthy individuals. Increased levels of acetylated polyamines in tissues from colon cancer patients, especially in cancer tissue from biofilm-covered CRC, are suggestive of a positive correlation between CRC pathogenesis and biofilm involvement. Acetylated polyamine metabolites have not been detected in biofilm-negative normal colonic tissue from healthy individuals, suggesting an important role of biofilms in polyamine catabolism. Bacterial biofilm may thus be a significant contributor to the elevated polyamine acetylation associated with colorectal cancer.

## Polymicrobial biofilm as a trigger of pro-carcinogenic inflammatory responses

The contribution of bacterium-induced inflammation to tumor growth is well-accepted (34,86,87). Specifically, the IL-23/IL-17 signaling axis appears to be important in bacterium-induced inflammation (36,88). This pathway has been shown to promote granulocyte accumulation with antibacterial activity, but it can also lead to DNA damage in humans and mice (22,36,87). Moreover, granulocytes are especially prone to releasing pro-inflammatory cytokines such as IL-1, IL-6, and IL-21 (22,36). A substantial body of literature also supports a role for IL-17 in human CRC development and progression (36,88,89). Accordingly, bacterial toxins, such as BFT, selectively promote IL-23/IL-17-dependent pro-inflammatory responses, inducing mouse colitis (20,21,34). For example, blockade of IL-17 and IL-23 receptor (IL23R), by using blocking antibody, inhibits ETBF-induced colitis in the *Apc*<sup>Min/+</sup> mouse model, and only Stat3 is activated after ETBF colonization (34). In another example, depleting the intestinal microbial flora in CPCAPC mice by short-term (3 weeks) cocktail treatment with broad-spectrum antibiotics (in drinking water) significantly reduced IL17A mRNA expression and decreased Stat3 activation in colonic tumor cells (36). Moreover, long-term (3.5 months) antibiotic treatment markedly reduced colon tumor size in CPC-APC mice compared to IL23R-deficient CPC-APC mice, further suggesting that tumor-eliciting gut pathogens might be able to promote IL-23-dependent pro-oncogenic signaling (36). Indeed, recent studies have shown a significant positive association between the abundance of intra-tumor *Fusobacterium* spp. and the protein levels of several inflammatory cytokines (such as TNF- $\alpha$  and IL-6) in human CRC, although no such correlation was found in the colonic mucosa of normal controls, implying that *Fusobacteria*-induced carcinogenesis may depend on the potential of biofilm to provoke mucosal inflammation (29,30). Consequently, the induction of IL-23/IL-17-mediated immune responses in the intestinal epithelium could be linked to the putative pro-oncogenic potential of biofilm bacteria.

Accumulating evidence also supports the idea that mucosal biofilms enhance the pro-oncogenic potential of microbial dysbiosis by eliciting inflammatory responses

(11,68,90). For instance, the concentration (mean density) of mucosa-adherent bacteria (biofilm formation) was found to be significantly higher in intestinal biopsy specimens from patients with **Crohn's disease (CD)** compared to non-inflammatory controls, including patients with **irritable bowel syndrome (IBS)** and healthy individuals; almost all of the microbiota detected in the IBD patient intestinal tract were adherent to the epithelium (68). Intestinal microbiota are controlled by interactions between microbial activities and host immunity, and involve highly organized spatial structures of microorganisms linked to specific immune responses (11,38,68). For instance, highly-virulent *Enterococcus* spp., which colonize the intestinal epithelium of the colon and rectum, form a biofilm that protects bacteria from reactive oxygen and promotes detoxification of hydrogen peroxide in IBD patients (91). Conversely, microbial biofilms also promote specific immune responses, which in turn may stimulate cancer development. As previously stated, biofilms may directly trigger the activation of IL-6/STAT3 pro-inflammatory signaling in human intestinal epithelial cells, a well-established tumor-promoting pathway. In support, a recent study showed that IL-6 and downstream phosphorylated STAT3 levels were markedly elevated in non-tumor biopsy samples from patients with biofilm-positive CRC compared to normal mucosal tissues from biofilm-negative CRC patients (11). Furthermore, IL-6 was found to be highly expressed in biofilm-covered intestinal mucosa with STAT3 activation, even in healthy individuals, suggesting that biofilm *per se* may be capable of promoting IL-6/STAT3 inflammatory signaling (11).

Another example relates to the role of **amyloid fibrils** and inflammation. These structures constitute a common and important component of mucosal biofilm (92,93), and are produced by dominant intestinal bacterial species belonging to the phyla *Firmicutes*, *Bacteroidetes*, and *Proteobacteria* (94). Amyloid fibrils, such as curli, provide relevant **pathogen-associated molecular patterns (PAMPs)** that can induce **Toll-like receptor (TLR) signaling** in immune cells (93–96). Curli, a potent TLR2/TLR1 ligand, has been shown to activate a TLR2-dependent host response in murine colitis models (mice pretreated with streptomycin), contributing to IL17A/IL-22-involved inflammatory responses (93–95). By incubating mouse naïve CD4<sup>+</sup> T cells

with supernatants from **bone marrow-derived dendritic cells (BMDCs)** pretreated with curli fibrils *in vitro*, the expression and production of IL-17A and IL-22 can induce the differentiation of naïve CD4<sup>+</sup> T cells into Th17 cells (94). Collectively, it is thus reasonable to assume that pro-oncogenic Th17-dependent signaling might be promoted by bacterial biofilm, leading to intestinal inflammation that, when severe, might constitute a way by which biofilms contribute to colonic neoplasia.

## Concluding remarks

Intestinal bacterial biofilms and their composition appear to play an important role in triggering and sustaining CRC progression. The molecular mechanisms underlying the crosstalk between bacteria-produced carcinogenic factors, bacterial biofilms, and host responses in CRC initiation and progression are only now emerging and, evidently, further research will be necessary to provide mechanistic insight into their precise involvement in the cancer process. The relative importance of these events remains unclear, but with the advent of population-wide screening programs for CRC, and the associated possibilities for prospective studies, rapid progress in this field seems likely (see **Outstanding Questions** and **Box 2**). When considering the distinct features of the spatial organization of microbiota in the human proximal and distal colon tissues (including cancer and normal tissues), important questions should be raised on whether invasive biofilm formation is preferentially present in human right colon cancer (11). In line with the unfavorable prognosis of right-sided CRC (97), we postulate that biofilm-positive CRC tumors may result in worse clinical outcomes relative to biofilm-negative tumors given that biofilm communities may cause additional serious epithelial tissue injury and intestinal inflammation. Moreover, recent emerging evidence supports this hypothesis; acetylated spermine can be significantly reduced in resected tumor tissues from CRC patients receiving oral antibiotics 24 h before surgery compared to untreated colonic biofilm-covered cancer tissue samples (13). We therefore posit that CRC initiation and progression arise as a consequence of the pro-oncogenic activities of biofilms composed of invasive bacteria. Preventive strategies aimed at the early detection and inhibition of such biofilm

deposition might prove to be useful for individuals at risk for CRC. In addition, specific antibiotics targeting bacterial biofilms might be considered among various potential therapeutic approaches. However, such strategies should be tested cautiously when attempting to target cancer cells (selectively or otherwise) because modifying the intestinal milieu and its microbiota could result in metabolic imbalances and deleterious consequences for the host.

### Outstanding Questions

- What are the precise mechanisms of biofilm-mediated carcinogenesis, and do these differ in proximal versus distal colon?
- What precise role does the mucosal biofilm play in host–microbe and trans-kingdom microbial interactions?
- What might drive the biofilm formation that has been linked to CRC initiation?
- Are specific bacterial species necessary to form the bacterial biofilm? Or, do non-specific microbes contribute to CRC-associated biofilm formation?
- What types of bacterial species are implicated in this process? Are they opportunistic pathogens or are they intrinsically pathogenic microorganisms?
- Are there any differences in the pro-carcinogenic potential of different bacterial species constituting intestinal biofilms?
- Can bacterial biofilm be effectively used to screen for early colon cancer?

### Box 2. Clinician's Corner

In the past decade scientists have looked with interest at the action of a single causative pathogen that triggers CRC initiation, and promotes tumor progression, with the presumption that early detection of a specific microorganism might benefit CRC prevention and treatment. Recently, the concept of polymicrobial synergy together with host–microbe interactions underlying bacterial oncogenesis has challenged this notion. Polybacterial biofilm (a high-order bacterial community structure) is not only a microbial block mass that shows resistance to antibiotics and to host defenses but also a dynamically complex ecosystem associated with progression and severity of disease, such as in periodontitis, colitis, and CRC. Biofilm has increasingly been implicated in various intestinal diseases, especially in IBD and CRC. The multispecies biofilm could potentially lead to severe intestinal inflammation and more-aggressive colon tumors. Future studies in prospective CRC patient cohorts are necessary to confirm the importance of bacterial biofilm in human CRC. If successful, assaying for invasive biofilm might become useful for improving the specificity and sensitivity of population-based screening for CRC. New antibiotics specifically targeting eradication of invasive biofilm might be contemplated as potentially promising preventive/therapeutic strategies against CRC.

## Glossary

**Absent in melanoma 2 (Aim2):** an innate immune sensor encoded by the *AIM2* gene that is frequently mutated in patients with CRC. The mouse homolog is *Aim2*.

**Adenoma:** a benign tumor originating in epithelial tissue within glandular structures; it may affect various organs such as stomach, colon, and lung.

**Adenoma–carcinoma sequence model:** an experimental model described as a stepwise progression from normal colorectal epithelium to adenoma, and eventually to invasive carcinoma as a result of the accumulation of genetic and epigenetic mutations.

**Amyloid fibrils:** highly-ordered peptide or protein aggregates formed in human organs and tissues under abnormal conditions that are associated with various human diseases, such as peptide aggregates in brain tissue of Alzheimer patients. Bacteria can also assemble amyloid fibrils to contribute to bacterial biofilm formation.

**Apc<sup>Min/+</sup> mice:** a well-established murine model for studying human colon cancer. Mice carry a truncating mutation of the mouse homolog of the *APC* gene and develop multiple intestinal neoplasia (Min).

**Bacterial driver-passenger:** bacteria-induced CRC carcinogenesis model; it defines driver bacteria as inducers of DNA damage to epithelial cells, initiating mutagenesis, allowing colonization of the epithelium by opportunistic bacteria (passenger pathogens) which outcompete driver bacteria, and ultimately resulting in cell proliferation/neoplasia.

**Biofilm:** a higher-order structure generated by an assembly of microbial pathogens that may display enhanced pathogenicity to the host. When microbes attach and reproduce on a living or non-living substrate surface, bacterial communities will be influenced by so-called extracellular quorum-sensing signals. These signals can modify the microbiota into a specific structure encased in a polymeric matrix including polysaccharides, proteins, and other components.

**Bone marrow-derived dendritic cells (BMDCs):** potent antigenpresenting cells that induce antitumor immunity; they can be exploited to assess immunomodulatory and anti-inflammatory responses in various experimental models.

**Cancer cell nest:** an accumulation of proliferating cancer cells separate from the stroma and necrotic lesions.

**CPC-APC mice:** *Apc<sup>fl/+</sup>* mice harboring a Cdx2–Cre transgene which can induce genetic deficiency in *Apc* gene expression, resulting in tumorigenesis, primarily in the distal colon.

**Crohn's disease (CD):** a type of IBD characterized by chronic relapsing transmural inflammation that may affect any segment of the GI tract.

**Epithelial cadherin (E-cadherin):** a cell–cell adhesion receptor that is the main component of epithelial adherent junctions, and has also been implicated in modulating tumor progression by acting as a tumor suppressor in the  $\beta$ -catenin signaling pathway.

**Genomic instability:** a typical characteristic of many cancers; refers to the accumulation and high frequency of chromosomal aberrations. These mutations include deletions, duplications, aneuploidy, and chromosomal translocations.

**Genotoxic:** a term to defined the ability of radiation and chemical agents to damage the genetic materials of cells, resulting in mutations, and potentially leading to oncogenesis.

**Inflammatory bowel disease (IBD):** a chronic inflammatory condition, including CD and ulcerative colitis, that involves part or the entirety of the intestinal tract, and that is linked to host immune responses against the intestinal microbiome. IBD is well-established as an important risk factor for the initiation, development, and progression of colorectal cancer.

**Intestinal crypts:** branched tubular structures surrounding the base of villi, composed of cells that can differentiate into proliferating epithelial cells, and contributing to self-renewal of the intestinal epithelium.

**Intestinal epithelial barrier:** composed of an outer loose mucus layer, an inner firm mucus layer, and tight epithelial cell–cell junctions. The thick inner mucus layer is nearly impenetrable to bacterial invasion and it affords protection to the intestinal epithelium by physically separating bacteria from the cells.



**Irritable bowel syndrome (IBS):** a common functional disorder that chronically affects the large intestine but which occurs in the absence of pathological lesions.

**Keystone pathogen:** a species present in low abundance in the intestine but that is of particular functional significance because it interferes with host immunity and remodels the composition of the entire microbial community.

**Microbial dysbiosis:** bacterial imbalance caused by alterations in the number or identity of bacteria making up the microbiota in the host. Bacterial dysbiosis has been linked not only to intestinal diseases such as IBD and IBS but also to extra-intestinal disorders including cardiovascular disease and diabetes.

**Microbiome:** in general this refers to the microbial community that inhabits a host organ or part of it, the microbial habitat, and the entire genome of its constituent microorganisms.

**Mismatch repair (MMR):** a system for recognizing, removing, and correcting errors in wrongly paired DNA bases, replacing them with correct bases during DNA replication; it plays an important role in maintaining genome stability.

**Oxidative stress:** a condition of disproportionate amounts of oxidants and relatively low levels of antioxidants, leading to excessive oxidation and potential cell damage.

**Pathogen-associated molecular patterns (PAMPs):** small molecular motifs, which are derived from groups of microbes, that can be recognized by host immune cells to induce innate immunity in response to microbial infection.

**Polyamine metabolites:** substances produced by polyamine metabolism, such as *N*<sup>1</sup>-acetylspermidine, *N*<sup>1</sup>-acetylspermine, and *N*<sup>1</sup>, *N*<sup>12</sup>-diacetylspermine.

**Reactive oxygen species (ROS):** highly reactive oxygen-containing molecules that may damage DNA, RNA, and protein.

**Short-chain fatty acids (SCFAs):** fatty acids with a chain length shorter than six carbon atoms, such as butyrate, acetate, and propionate, which can be produced by bacterial fermentation of dietary fibers in the colon.

**Symbionts:** microorganisms that live together and may benefit their host.

**Toll-like receptor (TLR) signaling:** a pathway that exerts important roles in the innate immune system. Toll-like receptors bind and recognize microbes or molecules derived from pathogens (e.g., PAMPs), triggering innate immunity.

**Tumor immune surveillance:** protective immune cell recognition and elimination of cancerous or precancerous cells by the host.

**Type 17 T helper (Th17) cells:** a subgroup of T helper CD4<sup>+</sup> lymphocytes characterized by the production of interleukin 17 (IL-17), which induces pro-inflammatory activities.

## Acknowledgments

S.L. is supported by a PhD fellowship from the China Scholarship Council (CSC, 201408060053).

## References

1. R. Sender *et al.*, Are we really vastly outnumbered? Revisiting the ratio of bacterial to host cells in humans. *Cell* **164**, 337-340 (2016).
2. J. R. Thiagarajah *et al.*, Secretory diarrhoea: mechanisms and emerging therapies. *Nat. Rev. Gastroenterol. Hepatol.* **12**, 446-457 (2015).
3. A. W. DuPont, H. L. DuPont, The intestinal microbiota and chronic disorders of the gut. *Nat. Rev. Gastroenterol. Hepatol.* **8**, 523-531 (2011).
4. F. Aviles-Jimenez *et al.*, Stomach microbiota composition varies between patients with non-atrophic gastritis and patients with intestinal type of gastric cancer. *Sci. Rep.* **4**, 4202 (2014).
5. C. S. Eun *et al.*, Differences in gastric mucosal microbiota profiling in patients with chronic gastritis, intestinal metaplasia, and gastric cancer using pyrosequencing methods. *Helicobacter* **19**, 407-416 (2014).
6. A. Machida-Montani *et al.*, Atrophic gastritis, *Helicobacter pylori*, and colorectal cancer risk: a case-control study. *Helicobacter* **12**, 328-332 (2007).
7. M. Castellarin *et al.*, *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome Res.* **22**, 299-306 (2012).
8. J. C. Arthur *et al.*, Microbial genomic analysis reveals the essential role of inflammation in bacteria-induced colorectal cancer. *Nat. Commun.* **5**, 4724 (2014).
9. A. D. Kostic *et al.*, Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome Res.* **22**, 292-298 (2012).
10. A. D. Kostic *et al.*, *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe* **14**, 207-215 (2013).
11. C. M. Dejea *et al.*, Microbiota organization is a distinct feature of proximal colorectal cancers. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18321-18326 (2014).
12. Z. G. Gao *et al.*, Microbiota disbiosis is associated with colorectal cancer. *Front. Microbiol.* **6**, 20 (2015).
13. C. H. Johnson *et al.*, Metabolism links bacterial biofilms and colon carcinogenesis. *Cell. Metab.* **21**, 891-897 (2015).
14. R. Siegel *et al.*, Colorectal cancer statistics, 2014. *CA Cancer J. Clin.* **64**, 104-117 (2014).
15. L. A. Torre *et al.*, Global cancer statistics, 2012. *CA Cancer J. Clin.* **65**, 87-108 (2015).
16. J. P. Zackular *et al.*, The human gut microbiome as a screening tool for colorectal cancer. *Cancer Prev. Res.* **7**, 1112-1121 (2014).
17. N. T. Baxter *et al.*, Microbiota-based model improves the sensitivity of fecal immunochemical test for detecting colonic lesions. *Genome Med.* **8**, 37 (2016).
18. N. Wu *et al.*, Dysbiosis signature of fecal microbiota in colorectal cancer patients. *Microb. Ecol.* **66**, 462-470 (2013).
19. S. M. Man *et al.*, Critical role for the DNA sensor AIM2 in stem cell proliferation and cancer. *Cell* **162**, 45-58 (2015).
20. G. Hajishengallis *et al.*, The keystone-pathogen hypothesis. *Nat. Rev. Microbiol.* **10**, 717-725 (2012).
21. C. L. Sears, D. M. Pardoll, Perspective: alpha-bugs, their microbial partners, and the link to colon cancer. *J. Infect. Dis.* **203**, 306-311 (2011).
22. H. Tjalsma *et al.*, A bacterial driver-passenger model for colorectal cancer: beyond the usual suspects. *Nat. Rev. Microbiol.* **10**, 575-582 (2012).
23. A. Nomura *et al.*, *Helicobacter pylori* infection and gastric carcinoma among Japanese Americans in Hawaii. *N. Engl. J. Med.* **325**, 1132-1136 (1991).
24. J. Parsonnet *et al.*, *Helicobacter pylori* infection and the risk of gastric carcinoma. *N. Engl. J. Med.* **325**, 1127-1131 (1991).
25. N. J. Talley *et al.*, Gastric adenocarcinoma and *Helicobacter pylori* infection. *J. Natl. Cancer Inst.* **83**, 1734-1739 (1991).

26. J. Parsonnet *et al.*, *Helicobacter pylori* infection in intestinal- and diffuse-type gastric adenocarcinomas. *J. Natl. Cancer Inst.* **83**, 640-643 (1991).
27. C. de Martel *et al.*, Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. *Lancet Oncol.* **13**, 607-615 (2012).
28. G. Zeller *et al.*, Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol. Syst. Biol.* **10**, 766 (2014).
29. T. Saito *et al.*, Two FOXP3<sup>+</sup>CD4<sup>+</sup> T cell subpopulations distinctly control the prognosis of colorectal cancers. *Nat. Med.* **22**, 679-684 (2016).
30. A. N. McCoy *et al.*, *Fusobacterium* is associated with colorectal adenomas. *PLoS One* **8**, e53653 (2013).
31. M. R. Rubinstein *et al.*, *Fusobacterium nucleatum* promotes colorectal carcinogenesis by modulating E-cadherin/ $\beta$ -catenin signaling via its FadA adhesin. *Cell Host Microbe* **14**, 195-206 (2013).
32. K. Mima *et al.*, *Fusobacterium nucleatum* and T cells in colorectal carcinoma. *JAMA Oncol.* **1**, 653-661 (2015).
33. J. R. Marchesi, *et al.*, Towards the human colorectal cancer microbiome. *PLoS One* **6**, e20447 (2011).
34. S. Wu *et al.*, A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nat. Med.* **15**, 1016-1022 (2009).
35. A. L. Geis *et al.*, Regulatory T-cell response to Enterotoxigenic *Bacteroides fragilis* colonization triggers IL17-dependent colon carcinogenesis. *Cancer Discov.* **5**, 1098-1109 (2015).
36. S. I. Grivennikov *et al.*, Adenoma-linked barrier defects and microbial products drive IL-23/IL-17-mediated tumour growth. *Nature* **491**, 254-258 (2012).
37. A. Lavelle *et al.*, Spatial variation of the colonic microbiota in patients with ulcerative colitis and control volunteers. *Gut* **64**, 1553-1561 (2015).
38. A. Swidsinski *et al.*, Spatial organization of bacterial flora in normal and inflamed intestine: A fluorescence in situ hybridization study in mice. *World J. Gastroenterol.* **11**, 1131-1140 (2005).
39. J. W. Costerton *et al.*, Colonization of particulates, mucous, and intestinal tissue. *Prog. Food Nutr. Sci.* **7**, 91-105 (1983).
40. E. Drenkard, F. M. Ausubel, *Pseudomonas* biofilm formation and antibiotic resistance are linked to phenotypic variation. *Nature* **416**, 740-743 (2002).
41. T. Vollmeret *et al.*, Interactions between endocarditis-derived *Streptococcus gallolyticus* subsp. *gallolyticus* isolates and human endothelial cells. *BMC Microbiol.* **10**, 78 (2010).
42. E. R. Fearon, B. Vogelstein, A genetic model for colorectal tumorigenesis. *Cell* **61**, 759-767 (1990).
43. D. Palestrant *et al.*, Microbial biofilms in the gut: visualization by electron microscopy and by acridine orange staining. *Ultrastruct. Pathol.* **28**, 23-27 (2004).
44. A. Swidsinski *et al.*, Acute appendicitis is characterised by local invasion with *Fusobacterium nucleatum/necrophorum*. *Gut* **60**, 34-40 (2011).
45. J. Strauss *et al.*, Invasive potential of gut mucosa-derived *Fusobacterium nucleatum* positively correlates with IBD status of the host. *Inflamm. Bowel Dis.* **17**, 1971-1978 (2011).
46. D. J. Bradshaw *et al.*, Role of *Fusobacterium nucleatum* and coaggregation in anaerobe survival in planktonic and biofilm oral microbial communities during aeration. *Infect. Immun.* **66**, 4729-4732 (1998).
47. A. Sharma *et al.*, Synergy between *Tannerella forsythia* and *Fusobacterium nucleatum* in biofilm formation. *Oral Microbiol. Immunol.* **20**, 39-42 (2005).
48. N. O. Kaakoush *et al.*, Is *Campylobacter* to esophageal adenocarcinoma as *Helicobacter* is to gastric adenocarcinoma? *Trends Microbiol.* **23**, 455-462 (2015).
49. N. P. Deshpande *et al.*, *Campylobacter concisus* pathotypes induce distinct global responses in intestinal epithelial cells. *Sci. Rep.* **6**, 34288 (2016).
50. R. L. Warren *et al.*, Co-occurrence of anaerobic bacteria in colorectal carcinomas. *Microbiome* **1**, 16 (2013).
51. N. Castaño-Rodríguez *et al.*, Dual role of *Helicobacter* and *Campylobacter* species in IBD: a systematic review and meta-analysis. *Gut* **66**, 235-249 (2015).

52. I. Yang *et al.*, Different gastric microbiota compositions in two human populations with high and low gastric cancer risk in Colombia. *Sci. Rep.* **6**, 18594 (2016).
53. J. Sillanpaa *et al.*, A collagen-binding adhesin, Acb, and ten other putative MSCRAMM and pilus family proteins of *Streptococcus gallolyticus* subsp. *gallolyticus* (*Streptococcus bovis* Group, Biotype I). *J. Bacteriol.* **191**, 6643-6653 (2009).
54. J. Sillanpaa *et al.*, Adherence characteristics of endocarditis-derived *Streptococcus gallolyticus* ssp. *gallolyticus* (*Streptococcus bovis* biotype I) isolates to host extracellular matrix proteins. *FEMS Microbiol. Lett.* **289**, 104-109 (2008).
55. A. Boleij *et al.*, Surface-exposed histone-like protein A modulates adherence of *Streptococcus gallolyticus* to colon adenocarcinoma cells. *Infect. Immun.* **77**, 5519-5527 (2009).
56. A. P. Soler *et al.*, Increased tight junctional permeability is associated with the development of colon cancer. *Carcinogenesis* **20**, 1425-1431 (1999).
57. M. E. V. Johansson *et al.*, Bacteria penetrate the normally impenetrable inner colon mucus layer in both murine colitis models and patients with ulcerative colitis. *Gut* **63**, 281-291 (2014).
58. J. P. Nougayrede *et al.*, *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science* **313**, 848-851 (2006).
59. G. Cuevas-Ramos *et al.*, *Escherichia coli* induces DNA damage *in vivo* and triggers genomic instability in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11537-11542 (2010).
60. A. C. Goodwin *et al.*, Polyamine catabolism contributes to enterotoxigenic *Bacteroides fragilis*-induced colon tumorigenesis. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15354-15359 (2011).
61. E. Buc *et al.*, High prevalence of mucosa-associated E-coli producing cyclomodulin and genotoxin in colon cancer. *PLoS One* **8**, e56964 (2013).
62. R. Guidi *et al.*, Chronic exposure to the cytolethal distending toxins of Gram-negative bacteria promotes genomic instability and altered DNA damage response. *Cell. Microbiol.* **15**, 98-113 (2013).
63. N. U. Toprak *et al.*, A possible role of *Bacteroides fragilis* enterotoxin in the aetiology of colorectal cancer. *Clin. Microbiol. Infect.* **12**, 782-786 (2006).
64. J. M. Kim *et al.*, Inhibition of apoptosis in *Bacteroides fragilis* enterotoxin-stimulated intestinal epithelial cells through the induction of c-IAP-2. *Eur. J. Immunol.* **38**, 2190-2199 (2008).
65. J. M. Abais *et al.*, Redox regulation of NLRP3 inflammasomes: ROS as trigger or effector? *Antioxid. Redox Signal.* **22**, 1111-1129 (2015).
66. S. Wu *et al.*, *Bacteroides fragilis* enterotoxin induces c-Myc expression and cellular proliferation. *Gastroenterology* **124**, 392-400 (2003).
67. S. Wu *et al.*, *Bacteroides fragilis* toxin stimulates intestinal epithelial cell shedding and gamma-secretase-dependent E-cadherin cleavage. *J. Cell Sci.* **120**, 1944-1952 (2007).
68. A. Swidsinski *et al.*, Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. *J. Clin. Microbiol.* **43**, 3380-3389 (2005).
69. S. R. Gill *et al.*, Metagenomic analysis of the human distal gut microbiome. *Science* **312**, 1355-1359 (2006).
70. C. Bernstein *et al.*, Carcinogenicity of deoxycholate, a secondary bile acid. *Arch. Toxicol.* **85**, 863-871 (2011).
71. S. Yoshimoto *et al.*, Obesity-induced gut microbial metabolite promotes liver cancer through senescence secretome. *Nature* **499**, 97-101 (2013).
72. M. Quante *et al.*, Bile acid and inflammation activate gastric cardia stem cells in a mouse model of Barrett-like metaplasia. *Cancer Cell* **21**, 36-51 (2012).
73. H. Cao *et al.*, The secondary bile acid, deoxycholate accelerates intestinal adenoma-adenocarcinoma sequence in *Apc<sup>min/+</sup>* mice through enhancing Wnt signaling. *Fam. Cancer* **13**, 563-571 (2014).
74. K. Dvorak *et al.*, Bile acids in combination with low pH induce oxidative stress and oxidative DNA damage: relevance to the pathogenesis of Barrett's oesophagus. *Gut* **56**, 763-771 (2007).
75. W. R. Russell *et al.*, Major phenylpropanoid-derived metabolites in the human gut can arise from microbial fermentation of protein. *Mol. Nutr. Food Res.* **57**, 523-535 (2013).

76. F. Májer *et al.*, New highly toxic bile acids derived from deoxycholic acid, chenodeoxycholic acid and lithocholic acid. *Bioorg. Med. Chem.* **22**, 256-268 (2014).
77. W. J. Cai *et al.*, Hydrogen sulfide induces human colon cancer cell proliferation: role of Akt, ERK and p21. *Cell Biol. Int.* **34**, 565-572 (2010).
78. M. S. Attene-Ramos *et al.*, Evidence that hydrogen sulfide is a genotoxic agent. *Mol. Cancer Res.* **4**, 9-14 (2006).
79. N. Ijssennagger *et al.*, Gut microbiota facilitates dietary heme-induced epithelial hyperproliferation by opening the mucus barrier in colon. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 10038-10043 (2015).
80. B. S. Alam *et al.*, Synthesis of nitrosopiperidine from nitrate and piperidine in the gastro-intestinal tract of the rat. *Nature* **232**, 199-200 (1971).
81. S. Macfarlane, G. T. Macfarlane, Composition and metabolic activities of bacterial biofilms colonizing food residues in the human gut. *Appl. Environ. Microbiol.* **72**, 6204-6211 (2006).
82. A. Belcheva *et al.*, Gut microbial metabolism drives transformation of MSH2-deficient colon epithelial cells. *Cell* **158**, 288-299 (2014).
83. R. Chaturvedi *et al.*, Spermine oxidase mediates the gastric cancer risk associated with *Helicobacter pylori* CagA. *Gastroenterology* **141**, 1696-1708 (2011).
84. N. Babbar, R. A. Casero, Tumor necrosis factor- $\alpha$  increases reactive oxygen species by inducing spermine oxidase in human lung epithelial cells: a potential mechanism for inflammation-induced carcinogenesis. *Cancer Res.* **66**, 11125-11130 (2006).
85. J. M. Tucker *et al.*, Potent modulation of intestinal tumorigenesis in *Apc<sup>min/+</sup>* mice by the polyamine catabolic enzyme spermidine/spermine N1-acetyltransferase. *Cancer Res.* **65**, 5390-5398 (2005).
86. Y. Li *et al.*, IL-6-induced DNMT1 activity mediates SOCS3 promoter hypermethylation in ulcerative colitis-related colorectal cancer. *Carcinogenesis* **33**, 1889-1896 (2012).
87. L. B. Meira *et al.*, DNA damage induced by chronic inflammation contributes to colon carcinogenesis in mice. *J. Clin. Invest.* **118**, 2516-2525 (2008).
88. L. Wang *et al.*, IL-17 can promote tumor growth through an IL-6-Stat3 signaling pathway. *J. Exp. Med.* **206**, 1457-1464 (2009).
89. S. Grivennikov *et al.*, IL-6 and Stat3 are required for survival of intestinal epithelial cells and development of colitis-associated cancer. *Cancer Cell* **15**, 103-113 (2009).
90. A. Swidsinski *et al.*, Mucosal flora in inflammatory bowel disease. *Gastroenterology* **122**, 44-54 (2002).
91. E. Golinska *et al.*, Virulence factors of *Enterococcus* strains isolated from patients with inflammatory bowel disease. *World J. Gastroenterol.* **19**, 3562-3572 (2013).
92. P. M. Gallo *et al.*, Amyloid-DNA composites of bacterial biofilms stimulate autoimmunity. *Immunity* **42**, 1171-1184 (2015).
93. G. O. Oppong *et al.*, Biofilm-associated bacterial amyloids dampen inflammation in the gut: oral treatment with curli fibres reduces the severity of hapten-induced colitis in mice. *NPJ Biofilms Microbiomes* **1**, 15019 (2015).
94. J. H. Nishimori *et al.*, Microbial amyloids induce interleukin 17A (IL-17A) and IL-22 responses via Toll-like receptor 2 activation in the intestinal mucosa. *Infect. Immun.* **80**, 4398-4408 (2012).
95. G. O. Oppong *et al.*, Epithelial cells augment barrier function via activation of the Toll-like receptor 2/phosphatidylinositol 3-kinase pathway upon recognition of *Salmonella enterica* serovar Typhimurium curli fibrils in the gut. *Infect. Immun.* **81**, 478-486 (2013).
96. G. J. Rapsinski *et al.*, Toll-like receptor 2 and NLRP3 cooperate to recognize a functional bacterial amyloid, curli. *Infect. Immun.* **83**, 693-701 (2015).
97. U. Nitsche *et al.*, Right sided colon cancer as a distinct histopathological subtype with reduced prognosis. *Dig. Surg.* **33**, 157-163 (2016).

# Chapter 4

## **Bacterial Biofilms as a Potential Contributor to Mucinous Colorectal Cancer Formation**

Shan Li<sup>1,2</sup>, Maikel P. Peppelenbosch<sup>1</sup>, Ron Smits<sup>1</sup>

<sup>1</sup>*Department of Gastroenterology and Hepatology, Erasmus MC-University Medical Center, Rotterdam, the Netherlands*

<sup>2</sup>*Department of Hepatobiliary Surgery, Daping Hospital, Army Medical University, Chongqing, China*

**BBA-Reviews on Cancer.** 2019, 1872(1):74-79.



## Abstract

A prominent mucinous phenotype is observed in 10–15% of all colorectal cancers (CRCs). They are associated with a proximal location, and more commonly observed among tumors with mismatch repair defects and a promoter CpG methylator phenotype. However, none of these features has been clearly linked mechanistically to this mucinous subtype. Here, we propose that bacterial biofilms could represent a currently unappreciated contributor to mucinous CRC formation. The colonic microbiome and biofilms in particular, are emerging as important factors in tumor initiation and progression. Intriguingly, biofilms preferentially accompany proximal tumors, suggesting that there may be a direct mechanistic link with mucinous CRCs.

**Keywords:** Biofilm; Microbiome; Colorectal cancer; Mucinous neoplasm; Mucus.



## **1. Introduction**

Colorectal cancer (CRC) is the third most frequent malignancy in the world, and is the second most common cause of cancer-related mortality (1). According to recent global cancer statistics, about 1.7 million people in the world were diagnosed with CRC, which resulted in approximately 832,000 deaths in 2015 (1). CRC has been classified in different subtypes according to criteria such as their histological appearance. Mucinous colorectal cancer (MCC) represents an important histological subset of CRC that is observed in 10–15% of cancers, and is arbitrarily defined as a tumor with more than 50% extracellular mucin on histologic examination (2,3). They are more commonly observed in the proximal colon (4). Mucinous histology by itself is associated with an increase in mortality compared with their non-mucinous counterparts, even when corrected for stage (2,5). Currently, the etiology of this subset of tumors is not yet fully understood, while they nevertheless are observed in one out of every 6–10 colorectal cancer patients. In this mini-review, we first introduce the dual character of mucus in initially preventing CRC development, while at later stages contributing to their progression. Next, we briefly describe the forms of genetic instability observed in CRC and their link to a mucinous phenotype. We then focus on the interactions between mucus, bacteria, and biofilms, and discuss probable reasons for the preferential occurrence of cancer-related biofilms in the proximal colon. Finally, we discuss the biofilm-associated mechanisms leading to enhanced mucus production during CRC initiation and development that may explain the emergence of mucinous CRCs. Many papers used for our review arbitrarily define mucinous tumors when showing more than 50% mucus and arbitrarily divide the colon in a proximal and distal part. For convenience we will adhere to these distinctions as well, but it should be realized that in reality these processes will follow a more gradual continuous model along the colonic tract (6).

## **2. The dual character of mucus in cancer formation**

Mucins are secreted by various organs to protect the epithelium against the external environment. The colon represents a prime example, as a thick mucus layer is formed

shielding the colonic epithelium from physical and chemical injury induced by food and microbes (7). Improper functioning of the mucus layer is observed in patients with cystic fibrosis and inflammatory bowel disease (IBD), in both cases strongly contributing to the etiology of the disease (7,8). Proper functioning of the mucus layer also decreases the chance that tumor growth is initiated, which among others is evidenced by the increased intestinal tumor predisposition of mice defective in MUC2 or ATOH1 (9–13), respectively, resulting in a strongly impaired mucus layer or complete absence of the mucin-producing goblet cells. This beneficial tumor suppressive effect is however reversed when tumors progress to malignancy. Elevated mucin levels have been associated with worse prognosis for various tumor types including those of the colon, and can contribute to tumor growth in various ways (14,15). In mucinous cancer cells, the characteristic apical secretion of mucins typical for normal cells is lost, and the secreted mucus completely surrounds the cell surface. This has been shown to protect cancer cells from the adverse external environment and to assist cancer cells in evading immune responses (14,15). Tumor cells also use the adhesive properties of mucins on one hand to detach from the primary tumor mass and on the other hand to attach to endothelia and invade distant structures (14). High mucin levels have also been shown to reduce effectiveness of anti-cancer agents by acting as a mechanical barrier (16–18). Thus, the mucus that initially protected the epithelial cells against tumor initiation, now supports tumor cells in their survival and growth. As such, it is important to acquire a better understanding of the mechanisms underlying mucin production in cancers.

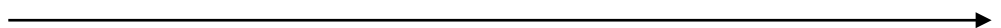
### **3. Forms of genetic instability in CRC and their link to the mucinous subtype**

Colorectal cancers are also categorized based on their underlying genetic instability. Chromosomal instability (CIN) accounts for more than 75% of all CRCs (19), which show a slight preference for the distal (left-sided) colorectal tract (**Fig. 1**). However, a prominent mucinous phenotype might be observed in only a small proportion of these tumors (20). A second form of genetic instability observed in CRCs is DNA mismatch

repair (MMR) deficiency, characterized by the accumulation of numerous mutations at the nucleotide level, especially in mono- or dinucleotide repeats. This high mutational load leads to the (in)activation of tumor-associated genes and the formation of many neo-antigens, ultimately resulting in the recruitment of abundant immune cells, a characteristic feature of this subset of tumors. These tumors account for 15% of all CRCs and predominantly arise in the proximal (right-sided) colon (70–80%) (**Fig. 1**) (21). Importantly, mucinous cancers are much more prominent among this subgroup and are observed in about 30–35% of MMR-deficient lesions (22). Lastly, a CpG island methylator phenotype (CIMP) is present in a significant subset of CRCs, resulting in hypermethylation and inactivation of promoters, some of which may be tumor suppressor genes. The mechanisms leading to CIMP are still not fully understood. On the proximal site about 30–40% of tumors are CIMP-high, whereas this is only 3–12% among distal tumors. The majority of CIMP-high tumors are characterized by a serrated histology, and are nowadays considered to represent a precursor lesion for a subset of mucinous cancers (23–26). An extensive overlap exists between the CIMP-phenotype and CRCs with sporadic MMR inactivation due to hypermethylation of the *MLH1* promoter, one of the mismatch repair genes.

Thus, mucinous CRC is associated with a proximal location, and more commonly observed among tumors with defects in the MMR machinery and/or CIMP-phenotype (**Fig. 1**). In addition, they show a higher incidence among IBD patients, suggesting a link with inflammation (27). Mucinous CRCs also show higher levels of *BRAF* and *PIK3CA* mutation than their non-mucinous counterparts (22). However, none of these features has been clearly linked mechanistically to the mucinous subtype. Here, we propose that bacterial biofilms could represent a currently unappreciated contributor to the mucinous subtype of CRCs.

---



**Fig. 1** Occurrence of colorectal tumors with CIN, MMR-deficiency, CIMP, and prominent mucinous phenotype in the proximal and distal colon. Tumors with this mucinous phenotype are more commonly observed in the proximal colon, and are observed in about 30–35% of lesions with MMR-deficiency and/or CIMP. They also show high levels of activating mutations in the *BRAF* and *PIK3CA* genes. Size and overlap of ellipses is in proportion to frequencies reported in the literature. CIN, Chromosomal Instability; MMR, mismatch repair; CIMP, CpG island methylator phenotype; MUC, mucinous subtype.

(see figure on next page)

#### 4. Mucus, bacteria, and biofilms

In the healthy colon, the secreted mucus organizes itself in a firm mucus layer directly attached to the epithelial cells, followed by a more loose layer (28). The firm and loose mucus layers are interacted with each other and in a dynamic situation, caused by continuous mucin-degradation by microbiota and constant replenishment from goblet cells, resulting in an ascending gradient of mucus viscosity from lumen to intestinal epithelium. The firm layer is mostly reported to be devoid of bacteria, while the loose layer is inhabited by commensal bacteria that in a symbiotic fashion aid in the digestion of luminal content and exclusion of potential pathogens (28). However, this homeostatic situation can be changed when the intestine is inflamed or temporary damaged by other insults. Under such circumstances the mucus barrier can become disrupted, allowing bacteria to come into direct contact with the epithelial cells. Bacterial species otherwise rarely observed in the healthy colon thus can find a niche to grow and possibly flourish. The chances of successful establishment are greatly increased by the formation of so-called bacterial biofilms. These are loosely defined as bacterial communities aggregating in a matrix such as mucus, which allows bacteria that normally would be rapidly purged from the colon to adhere to structures

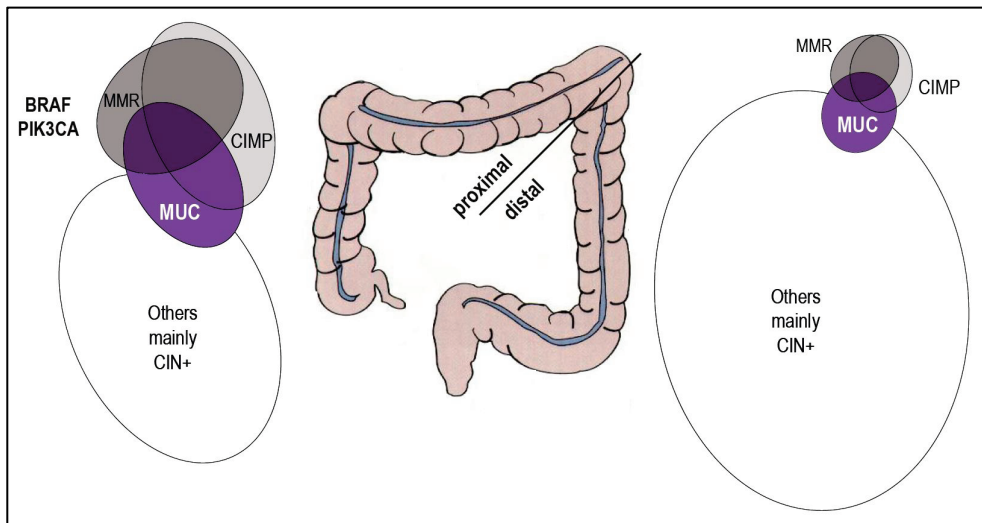


Fig. 1 (see legend on previous page)

such as the colonic epithelium or tumors thereof. In these biofilms, bacterial species cooperate in various ways, as outlined in more detail in several recent reviews (29–31). Some bacterial species are better adapted to adhering, invading or digesting the mucus layer, thereby helping others to remain in the intestine and get into closer contact with the underlying epithelium. One example of such a cooperation is represented by fecal co-colonization of pks-positive *Escherichia coli* and enterotoxigenic *Bacteroides fragilis* (32). The latter can reduce mucus depth allowing the pks+ *E. coli* and its associated colibactin genotoxin to get into closer contact with the intestinal epithelium.

These intestinal biofilms have emerged as an important contributing factor to CRC (30,32–36). Among others they will locally exacerbate intestinal inflammation, resulting in the production of reactive oxygen and nitrogen species that combined with genotoxic bacterial compounds, will increase the mutation rate within epithelial cells (32,33,37–39). Other consequences are a compromised epithelial barrier function, modulation of host metabolism, and promotion of epithelial cell proliferation (30,32–34). Combined these effects can increase the chance to trigger and promote colorectal tumorigenesis.

## **5. Preferential occurrence of cancer-related biofilms in the proximal colon**

Initially, one bacterial strain gained special interest for CRC formation, that is *Fusobacterium nucleatum* (Fn). It is rarely observed in the healthy colon, but possibly from an oral source, may find a niche in the diseased colon, often in consortium with other oral bacterial species (36,40). Nowadays, it is considered a causative agent for colorectal cancer initiation and/or progression. For example, *Fusobacterium* can increase the number of colonic tumors in the *Apc*<sup>Min</sup> mouse model, a mouse strain that spontaneously develops intestinal tumors(41). A potential interesting link with mucinous tumors is that several studies reported a proximal preference for *Fusobacterium* associated CRCs (35,42,43). The same holds true for biofilms in general as they were nearly always (around 93%) detected on proximal colonic tumors, but

much less frequently (about 27%) on distal tumors (33,36). Biofilms were mostly of polymicrobial nature and frequently enriched for *B. fragilis* and oral pathogens including *Fusobacterium*. Interestingly, when a biofilm is detected on a tumor, its flanking normal tissue is mostly also covered by biofilm, suggesting that it expands over long distances in the tumor environment.

The underlying mechanisms for the preferential proximal presence of tumor-associated biofilms are still unclear. Both sides of the colon differ in various aspects such as embryonic origin and luminal content (44). In mouse and rat the proximal mucus layer is much thinner than the distal one (45–47). In humans the difference is less pronounced but appears to double in thickness towards the distal end (48). A thinner proximal mucus layer possibly may be easier breached by bacteria, bringing them more readily in direct contact with the epithelial surface to form a biofilm. Secondly, the specific combination of bacteria in the proximal colon might be more efficient in forming biofilms (30,33). It has been shown that the microbiome differs along the colorectal tract (49,50). There is also a large degree of discordance in the microbial community compositions of intestinal mucosal samples (including tumor and paired non-tumor tissue) between biofilm-positive and biofilm-negative CRC patients (33). A third explanation may reside in the consistency of the luminal content, which is fluidic in the proximal colon and becomes more firm towards the distal end where stool is formed. The abrasive forces of this stool may prevent an efficient formation of biofilms. In addition, during the formation of feces in the distal mouse and rat colon, most bacteria apparently become entrapped in pellets encapsulated by mucus that is captured from the epithelium (47). Away from these pellets, the intestinal lumen and epithelium are mostly sterile. Whether bacteria also become entrapped during human stool formation remains to be determined, but in support it was shown that also human fecal pellets are enclosed by a continuous mucus gel layer (51). Thus, the combined effects of shear force and entrapment of bacteria within the stool may prevent distal biofilm formation.

Biofilms could also be secondary to tumor formation. The emergence of tumors by itself has been shown to damage the normal mucus barrier (48), which may already facilitate biofilm formation. In the proximal colon an additional mechanism may be

at play, that is about 25–30% of all proximal tumors are MMR-deficient tumors (Fig. 1). The strong local immune response that accompanies these tumors may further disrupt local tissue architecture, possibly making it easier for bacteria to find a niche. This can however not explain why virtually all proximal tumors show biofilms.

## **6. Biofilm-associated mechanisms leading to enhanced mucus production**

Although it is not entirely clear why biofilms preferentially accompany proximal tumors, it is intriguing that this is also the side where most mucinous tumors are formed, suggesting that there may be a direct mechanistic link. So what evidence exists to support such a hypothesis? Mucus production in the colon is dynamic and can be influenced by various factors. One important direct contributor to the amount and composition of mucus secreted by the colonic cells, are bacteria (52). For example, germ-free mice show significantly lower amounts of MUC2 protein in their colonic mucus layer, making the mucus more penetrable compared with conventionally raised mice, while gavage with cecal microbiota increases MUC2 expression and restores the impenetrable mucus in a matter of weeks (53). Similar observations have been made for colorectal tumor cells. Direct exposure of the human mucinous CRC cell line LS174T to highly invasive *Fn* strains strongly promoted *MUC2* expression and increased expression of the pro-inflammatory cytokine tumor necrosis factor alpha (TNF- $\alpha$ ) (54), which by itself can also enhance mucus production (see below). Likewise is mucin production increased in HT-29 colonic tumor cells by incubating them with a pathogenic *E. coli* strain or *Vibrio cholera* bacteria (55,56). Various other reports have presented similar observations (29).

A second more indirect link with mucus production is the exacerbation of inflammation induced by biofilms. The enhanced inflammatory response leads to the generation of large amounts of cytokines, such as TNF- $\alpha$ , IL-22 and others, for which several reports have shown that they can increase mucus production by colonic tumor cells (54,55,57–62). For example, prolonged TNF- $\alpha$  treatment of colonic tumor cells strongly increased the stability of ATOH1 protein, one of the main transcription

factors regulating mucinous differentiation, thereby increasing mucus production (62).

Taken together, it seems that inflammation and bacterial biofilms in a concerted action can induce more mucus production by tumor cells. The secreted mucus in turn provides the building blocks for an efficient biofilm formation, in a way leading to a vicious circle of biofilm formation, inflammation and enhanced mucus production.


An important prerequisite however is that the genetic alterations present in CRCs still allow for sufficient differentiation towards the mucinous direction. Several reports have shown increased *MUC2* and *ATOH1* promoter methylation and inactivation in a subset of colorectal tumors, which associated with low mucus production (9,63,64). Obviously, in tumors where this occurred, the mucus promoting features of biofilms and inflammation will have little effect on overall mucus production. Secondly, the great majority of CRCs acquire mutations in components of the Wnt/ $\beta$ -catenin signaling pathway resulting in aberrantly enhanced  $\beta$ -catenin signaling. This is mostly accomplished by inactivating mutations in the *APC* tumor suppressor gene, and to a lesser extent activating mutations of  $\beta$ -catenin itself or inactivating mutations in genes such as *AXIN1/AXIN2* or *RNF43* (21,65–70). The resulting enhanced  $\beta$ -catenin signal imposes a crypt progenitor phenotype onto the tumor cells (71), while simultaneously reducing but importantly not entirely blocking the possibilities for differentiation. Interestingly, we and others have shown that proximal CRCs select for mutations that lead to a moderate enhancement of  $\beta$ -catenin signaling, while distal tumors prefer a stronger signal (21,65–67). We have also outlined that this phenomenon likely explains the preferential proximal location of mismatch repair deficient tumors (21). In short, the defect in MMR leads mainly to *APC* or *CTNNB1* (encoding  $\beta$ -catenin protein) mutations resulting in moderate signaling levels ideal for the proximal colon, making their outgrowth on that side more likely. Whether a similar mechanism also explains the proximal preference of CIMP-high tumors remains to be determined. Anyway, the generally lower level of  $\beta$ -catenin signaling observed in CRCs on the proximal side likely allows for more differentiation of the tumor cells. In combination with the prevalent proximal biofilm formation



and/or the accompanying inflammation, this may more readily result in tumors that generate sufficient mucus to qualify them as mucinous CRCs (more than 50% mucus).

To sum up, the following scenarios linking bacterial biofilms and mucinous CRC can be envisaged. As depicted in **Fig. 2**, biofilms that are enabled to form in close contact with the intestinal epithelium, for example by inflammation or other insults, can contribute to tumor initiation through the various mechanisms described above. Once the tumor is formed the bacterial biofilms in concerted action with the exacerbated inflammation, enforces more mucus production within the tumor cells. This will however only occur when the underlying genetic alterations allow for sufficient differentiation of the tumor cells or mucin gene expression. Thus in this scenario, bacterial biofilms first contribute to more tumor formation and in a second phase to a specific differentiation pattern. In a second scenario, biofilm formation is secondary to tumor initiation. In that case, the biofilms mainly contribute to tumor progression and possibly increasing mucus production, likely again in concerted action with inflammation.

---



**Fig. 2** Possible scenario explaining how bacterial biofilms, inflammation and colonic tumor cells may interact to form a mucinous tumor. In the healthy colon a sterile firm mucus layer separates the epithelium from a more loose mucus layer inhabited by commensal bacteria. In instances of inflammation or other insults to the epithelium the mucus barrier may become breached, possibly resulting in biofilm formation. This biofilm in concerted action with an exacerbated inflammation increase the chances of tumor initiation. Once tumors are formed, these same features can induce mucus production by tumor cells, leading to a more likely diagnosis of a mucinous CRC. Alternatively, biofilm formation is secondary to tumor initiation, as the emergence of tumors by itself damages the normal mucus barrier, thereby providing a favorable niche for bacterial colonization and subsequent biofilm formation. This figure was created with Biorender (<https://biorender.com/>).

*(see figure on next page)*

## 7. Conclusions and perspectives

In the last decade, it is becoming increasingly clear that the colonic microbiome and bacterial biofilms play an important role in colorectal tumor development. Here, we have hypothesized that biofilms may also contribute to the specific mucinous phenotype observed in 10–15% of CRCs. This was inspired by the preferential proximal location of both mucinous CRCs as well as tumor-associated biofilms. There are however still several unanswered questions. For example, on histological examination the mucinous regions of cancers are often observed at their invasive front,

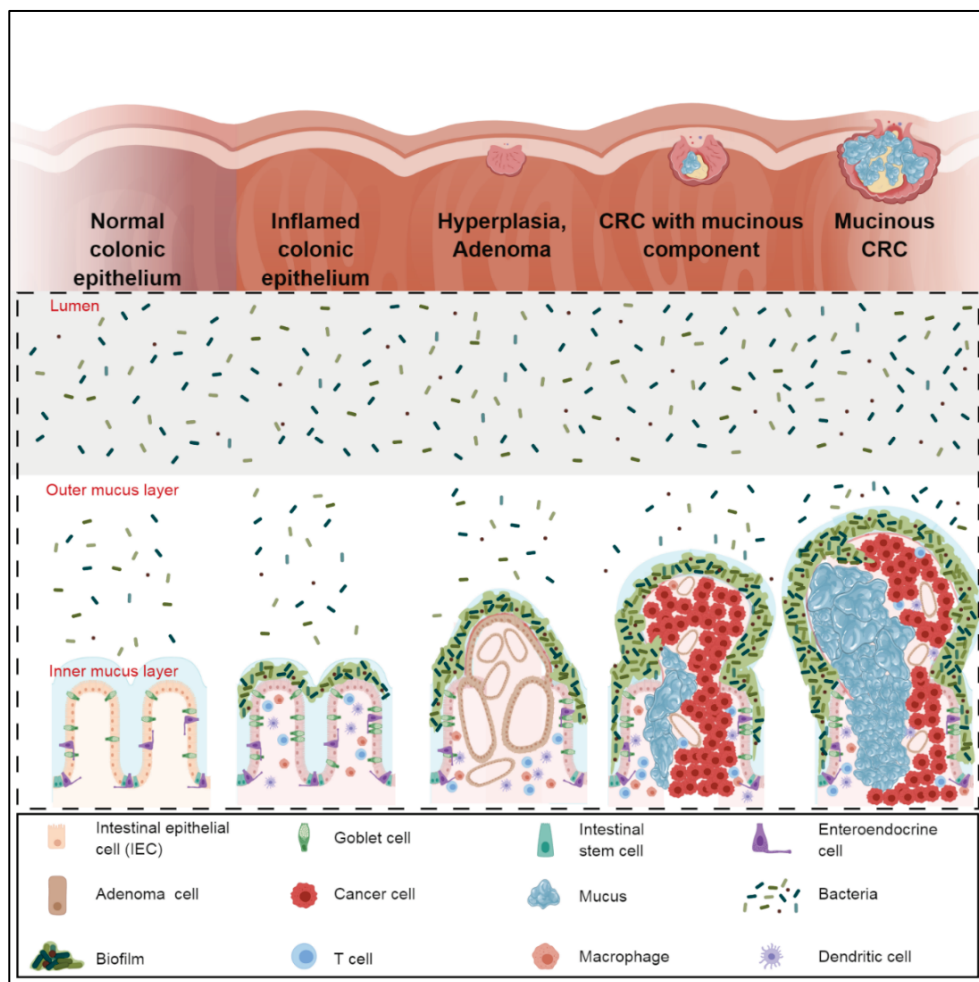


Fig. 2 (see legend on previous page)

so potentially at some distance from the luminal located biofilms. This may in part represent a technical artefact, that is the pre-operative procedures to clean the patient's bowel and the subsequent fixation and paraffin embedment are likely to remove mucus that is not entrapped within tissue sections. Nevertheless, our hypothesis needs confirmation by demonstrating bacterial aggregates within reasonable distance from the mucus producing tumor cells, or providing evidence that bacterial products can affect tumor cell behavior at some distance. For the latter indirect support is already provided by the altered mucus production of normal colonic cells not being in direct contact with the luminal bacteria (53). Furthermore, it is still unclear why biofilms mainly form in the proximal ascending colon. We have postulated some explanations, like the shear force and entrapment by stool preventing biofilms on the distal side, but whether this holds true remains to be shown. Likewise, it is unclear if the appearance and composition of biofilms associated with mucinous CRCs differs from other ones. Only few research groups have used the appropriate tools to look at biofilms and used fixation procedures that preserve mucus (e.g. Carnoy's fixative), but to our knowledge no reports have specifically looked at mucinous tumors. For the same reason it is also not known if specific bacterial strains are especially strong contributors to the mucinous subtype. Given the recent acknowledgement of biofilms contributing to colorectal tumor growth, obviously more detailed molecular and genetic analyses are needed. Moreover, as advocated by Hamada et al., this should ideally be complemented with a thorough epidemiologic analysis of lifestyle factors, dietary patterns, medications (e.g. antibiotics), and environmental exposures, which are all expected to interact with the microbiome, tumor cells and immune system (72). In the future, these analyses may uncover potential tailor-made therapies specifically aimed at the mucinous subtype of colorectal cancers.

### **Acknowledgements**

This work was financially supported by a China Scholarship Council PhD fellowship to Shan Li (File NO. 201408060053) and an Erasmus MC Grant. We also thank Dr. Michael Doukas for helpful discussions.

## References

1. C. Global Burden of Disease Cancer *et al.*, Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study. *JAMA Oncol.* **3**, 524-548 (2017).
2. J. Verhulst, L. Ferdinande, P. Demetter, W. Ceelen, Mucinous subtype as prognostic factor in colorectal cancer: a systematic review and meta-analysis. *J. Clin. Pathol.* **65**, 381-388 (2012).
3. N. Hugen, G. Brown, R. Glynne-Jones, J. H. de Wilt, I. D. Nagtegaal, Advances in the care of patients with mucinous colorectal cancer. *Nat. Rev. Clin. Oncol.* **13**, 361-369 (2016).
4. J. R. Hyngstrom *et al.*, Clinicopathology and outcomes for mucinous and signet ring colorectal adenocarcinoma: analysis from the National Cancer Data Base. *Ann. Surg. Oncol.* **19**, 2814-2821 (2012).
5. L. J. Mekenkamp *et al.*, Mucinous adenocarcinomas: poor prognosis in metastatic colorectal cancer. *Eur. J. Cancer* **48**, 501-509 (2012).
6. M. Yamauchi *et al.*, Colorectal cancer: a tale of two sides or a continuum? *Gut* **61**, 794-797 (2012).
7. Y. S. Kim, S. B. Ho, Intestinal goblet cells and mucins in health and disease: recent insights and progress. *Curr. Gastroenterol. Rep.* **12**, 319-330 (2010).
8. Y. H. Sheng, S. Z. Hasnain, T. H. Florin, M. A. McGuckin, Mucins in inflammatory bowel diseases and colorectal cancer. *J. Gastroenterol. Hepatol.* **27**, 28-38 (2012).
9. W. Bossuyt *et al.*, Atonal homolog 1 is a tumor suppressor gene. *PLoS Biol.* **7**, e39 (2009).
10. A. Velcich *et al.*, Colorectal cancer in mice genetically deficient in the mucin Muc2. *Science* **295**, 1726-1729 (2002).
11. K. Yang *et al.*, Interaction of Muc2 and Apc on Wnt signaling and in intestinal tumorigenesis: potential role of chronic inflammation. *Cancer Res.* **68**, 7313-7322 (2008).
12. G. Peignon *et al.*, Complex interplay between  $\beta$ -catenin signalling and Notch effectors in intestinal tumorigenesis. *Gut* **60**, 166-176 (2011).
13. A. Kazanjian, T. Noah, D. Brown, J. Burkart, N. F. Shroyer, Atonal homolog 1 is required for growth and differentiation effects of notch/gamma-secretase inhibitors on normal and cancerous intestinal epithelial cells. *Gastroenterology* **139**, 918-928 (2010).
14. M. A. Hollingsworth, B. J. Swanson, Mucins in cancer: protection and control of the cell surface. *Nat. Rev. Cancer* **4**, 45-60 (2004).
15. D. W. Kufe, Mucins in cancer: function, prognosis and therapy. *Nat. Rev. Cancer* **9**, 874-885 (2009).
16. A. V. Kalra, R. B. Campbell, Mucin overexpression limits the effectiveness of 5-FU by reducing intracellular drug uptake and antineoplastic drug effects in pancreatic tumours. *Eur. J. Cancer* **45**, 164-173 (2009).
17. X. Wang, A. A. Shah, R. B. Campbell, K. Wan, Glycoprotein mucin molecular brush on cancer cell surface acting as mechanical barrier against drug delivery. *Appl. Phys. Lett.* **97**, 263703 (2010).
18. N. Jonckheere, N. Skrypek, I. Van Seuningen, Mucins and tumor resistance to chemotherapeutic drugs. *Biochim. Biophys. Acta Rev. Cancer* **1846**, 142-151 (2014).
19. M. S. Pino, D. C. Chung, The chromosomal instability pathway in colon cancer. *Gastroenterology* **138**, 2059-2072 (2010).
20. N. Hugen *et al.*, Reduced rate of copy number aberrations in mucinous colorectal carcinoma. *Oncotarget* **6**, 25715-25725 (2015).
21. C. Albuquerque, E. R. Bakker, W. van Veelen, R. Smits, Colorectal cancers choosing sides. *Biochim. Biophys. Acta* **1816**, 219-231 (2011).
22. N. Hugen *et al.*, The molecular background of mucinous carcinoma beyond MUC2. *J. Pathol. Clin. Res.* **1**, 3-17 (2015).
23. C. Gallois, P. Laurent-Puig, J. Taieb, Methylator phenotype in colorectal cancer: A prognostic factor or not? *Crit. Rev. Oncol. Hematol.* **99**, 74-80 (2016).

24. Y. Y. Rhee, K. J. Kim, G. H. Kang, CpG island methylator phenotype-high colorectal cancers and their prognostic implications and relationships with the serrated neoplasia pathway. *Gut Liver* **11**, 38-46 (2017).
25. M. Bettington *et al.*, The serrated pathway to colorectal carcinoma: current concepts and challenges. *Histopathology*. **62**, 367-386 (2013).
26. C. T. Lee *et al.*, Serrated adenocarcinoma morphology in colorectal mucinous adenocarcinoma is associated with improved patient survival. *Oncotarget*. **8**, 35165-35175 (2017).
27. N. Hugen, J. J. van Beek, J. H. de Wilt, I. D. Nagtegaal, Insight into mucinous colorectal carcinoma: clues from etiology. *Ann. Surg. Oncol.* **21**, 2963-2970 (2014).
28. M. E. Johansson *et al.*, The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc Natl Acad Sci U S A* **105**, 15064-15069 (2008).
29. J. F. Sicard, G. Le Bihan, P. Voegelé, M. Jacques, J. Harel, Interactions of intestinal bacteria with components of the intestinal mucus. *Front. Cell. Infect. Microbiol.* **7**, 387 (2017).
30. S. Li, S. R. Konstantinov, R. Smits, M. P. Peppelenbosch, Bacterial biofilms in colorectal cancer initiation and progression. *Trends Mol. Med.* **23**, 18-30 (2017).
31. R. De Weirdt, T. Van de Wiele, Micromanagement in the gut: microenvironmental factors govern colon mucosal biofilm structure and functionality. *NPJ Biofilms Microbiomes*. **1**, 15026 (2015).
32. C. M. Dejea *et al.*, Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria. *Science* **359**, 592-597 (2018).
33. C. M. Dejea *et al.*, Microbiota organization is a distinct feature of proximal colorectal cancers. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18321-18326 (2014).
34. C. H. Johnson *et al.*, Metabolism links bacterial biofilms and colon carcinogenesis. *Cell Metab.* **21**, 891-897 (2015).
35. J. Yu *et al.*, Invasive *Fusobacterium nucleatum* may play a role in the carcinogenesis of proximal colon cancer through the serrated neoplasia pathway. *Int. J. Cancer* **139**, 1318-1326 (2016).
36. J. L. Drewes *et al.*, High-resolution bacterial 16S rRNA gene profile meta-analysis and biofilm status reveal common colorectal cancer consortia. *NPJ Biofilms Microbiomes*. **3**, 34 (2017).
37. L. R. Ferguson, Chronic inflammation and mutagenesis. *Mutat. Res.* **690**, 3-11 (2010).
38. A. Swidsinski, J. Weber, V. Loening-Baucke, L. P. Hale, H. Lochs, Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. *J. Clin. Microbiol.* **43**, 3380-3389 (2005).
39. A. C. Goodwin *et al.*, Polyamine catabolism contributes to enterotoxigenic *Bacteroides fragilis*-induced colon tumorigenesis. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15354-15359 (2011).
40. K. J. Flynn, N. T. Baxter, P. D. Schloss, Metabolic and community synergy of oral bacteria in colorectal cancer. *mSphere* **1**, e00102-00116 (2016).
41. A. D. Kostic *et al.*, *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe*. **14**, 207-215 (2013).
42. K. Mima *et al.*, *Fusobacterium nucleatum* in colorectal carcinoma tissue according to tumor location. *Clin. Transl. Gastroenterol.* **7**, e200 (2016).
43. K. Mima *et al.*, *Fusobacterium nucleatum* in colorectal carcinoma tissue and patient prognosis. *Gut* **65**, 1973-1980 (2016).
44. B. Iacopetta, Are there two sides to colorectal cancer? *Int. J. Cancer* **101**, 403-408 (2002).
45. R. Randal Bollinger, A. S. Barbas, E. L. Bush, S. S. Lin, W. Parker, Biofilms in the large bowel suggest an apparent function of the human vermiform appendix. *J. Theor. Biol.* **249**, 826-831 (2007).
46. A. Swidsinski, V. Loening-Baucke, H. Lochs, L. P. Hale, Spatial organization of bacterial flora in normal and inflamed intestine: a fluorescence in situ hybridization study in mice. *World J. Gastroenterol.* **11**, 1131-1140 (2005).

47. J. B. J. Kamphuis, M. Mercier-Bonin, H. Eutamene, V. Theodorou, Mucus organisation is shaped by colonic content; a new view. *Sci. Rep.* **7**, 8527 (2017).
48. K. Matsuo, H. Ota, T. Akamatsu, A. Sugiyama, T. Katsuyama, Histochemistry of the surface mucous gel layer of the human colon. *Gut* **40**, 782-789 (1997).
49. K. J. Flynn, M. T. Ruffin, D. K. Turgeon, P. D. Schloss, Spatial variation of the native colon microbiota in healthy adults. *Cancer Prev. Res. (Phila.)* **11**, 393-402 (2018).
50. B. Flemer *et al.*, Tumour-associated and non-tumour-associated microbiota in colorectal cancer. *Gut* **66**, 633-643 (2017).
51. A. Shimotoyodome, S. Meguro, I. Tokimitsu, T. Sakata, Histochemical structure of the mucus gel layer coating the fecal surface of rodents, rabbits and humans. *J. Nutr. Sci. Vitaminol. (Tokyo)* **51**, 287-291 (2005).
52. M. A. McGuckin, S. K. Linden, P. Sutton, T. H. Florin, Mucin dynamics and enteric pathogens. *Nat. Rev. Microbiol.* **9**, 265-278 (2011).
53. M. E. Johansson *et al.*, Normalization of host intestinal mucus layers requires long-term microbial colonization. *Cell Host Microbe* **18**, 582-592 (2015).
54. P. Dharmani, J. Strauss, C. Ambrose, E. Allen-Vercoe, K. Chadee, *Fusobacterium nucleatum* infection of colonic cells stimulates MUC2 mucin and tumor necrosis factor alpha. *Infect. Immun.* **79**, 2597-2607 (2011).
55. Y. Xue *et al.*, Host inflammatory response inhibits *Escherichia coli* O157:H7 adhesion to gut epithelium through augmentation of mucin expression. *Infect. Immun.* **82**, 1921-1930 (2014).
56. S. Almagro-Moreno, K. Pruss, R. K. Taylor, Intestinal colonization dynamics of *Vibrio cholerae*. *PLoS Pathog.* **11**, e1004787 (2015).
57. D. H. Ahn *et al.*, TNF-alpha activates MUC2 transcription via NF-kappaB but inhibits via JNK activation. *Cell. Physiol. Biochem.* **15**, 29-40 (2005).
58. F. Ishibashi *et al.*, Contribution of ATOH1<sup>+</sup> cells to the homeostasis, repair, and tumorigenesis of the colonic epithelium. *Stem Cell Reports* **10**, 27-42 (2018).
59. K. Sugimoto *et al.*, IL-22 ameliorates intestinal inflammation in a mouse model of ulcerative colitis. *J. Clin. Invest.* **118**, 534-544 (2008).
60. J. E. Turner, B. Stockinger, H. Helmby, IL-22 mediates goblet cell hyperplasia and worm expulsion in intestinal helminth infection. *PLoS Pathog.* **9**, e1003698 (2013).
61. S. Brand *et al.*, IL-22 is increased in active Crohn's disease and promotes proinflammatory gene expression and intestinal epithelial cell migration. *Am. J. Physiol. Gastrointest. Liver Physiol.* **290**, G827-838 (2006).
62. K. Fukushima *et al.*, Atonal homolog 1 protein stabilized by tumor necrosis factor alpha induces high malignant potential in colon cancer cell line. *Cancer Sci.* **106**, 1000-1007 (2015).
63. K. Okudaira *et al.*, MUC2 gene promoter methylation in mucinous and non-mucinous colorectal cancer tissues. *Int. J. Oncol.* **36**, 765-775 (2010).
64. A. Siedow *et al.*, De novo expression of the Muc2 gene in pancreas carcinoma cells is triggered by promoter demethylation. *Tumour Biol.* **23**, 54-60 (2002).
65. C. Albuquerque *et al.*, Colorectal cancers show distinct mutation spectra in members of the canonical WNT signaling pathway according to their anatomical location and type of genetic instability. *Genes Chromosomes Cancer* **49**, 746-759 (2010).
66. S. J. Leedham *et al.*, A basal gradient of Wnt and stem-cell number influences regional tumour distribution in human and mouse intestinal tracts. *Gut* **62**, 83-93 (2013).
67. M. Christie *et al.*, Different APC genotypes in proximal and distal sporadic colorectal cancers suggest distinct WNT/ $\beta$ -catenin signalling thresholds for tumourigenesis. *Oncogene* **32**, 4675-4682 (2013).
68. H. X. Hao, X. Jiang, F. Cong, Control of Wnt receptor turnover by R-spondin-ZNRF3/RNF43 signaling module and its dysregulation in cancer. *Cancers (Basel)* **8**, 54 (2016).

69. M. Giannakis *et al.*, *RNF43* is frequently mutated in colorectal and endometrial cancers. *Nat. Genet.* **46**, 1264-1266 (2014).
70. N. The Cancer Genome Atlas, Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330 (2012).
71. M. van de Wetering *et al.*, The  $\beta$ -catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* **111**, 241-250 (2002).
72. T. Hamada, J. A. Nowak, D. A. Milner, Jr., M. Song, S. Ogino, Integration of microbiology, molecular pathology, and epidemiology: a new paradigm to explore the pathogenesis of microbiome-driven neoplasms. *J. Pathol.* **247**, 615-628 (2019).

# Chapter 5

## Unravelling Posttranslational Modifications of ATOH1 Driving Colorectal Cancer Towards Mucinous Differentiation

Shan Li<sup>1,4</sup>, Werner Helvensteijn<sup>1</sup>, Raymond A. Poot<sup>2</sup>, Jeroen Demmers<sup>3</sup>, Wenhui Wang<sup>1,5</sup>,  
Zhijiang Miao<sup>1</sup>, Menggang Liu<sup>1,4</sup>, Bingting Yu, Marla Lavrijsen<sup>1</sup>, Buyun Ma,  
Shanshan Li, Pengyu Liu<sup>1</sup>, Maikel P. Peppelenbosch<sup>1</sup>, Ron Smits<sup>1</sup>

<sup>1</sup>Department of Gastroenterology and Hepatology, Erasmus MC-University Medical Center, Rotterdam, The Netherlands

<sup>2</sup>Department of Cell Biology, Erasmus MC-University Medical Center, Rotterdam, The Netherlands

<sup>3</sup>Proteomics Center, Erasmus MC-University Medical Center, Rotterdam, The Netherlands

<sup>4</sup>Department of Hepatobiliary Surgery, Daping Hospital, Army Medical University, Chongqing, China

<sup>5</sup>State Key Laboratory of Natural Medicines, Department of Pharmacology, China Pharmaceutical University, Nanjing, China

**Manuscript in Preparation.**





## Abstract

Atonal homolog 1 (ATOH1) is a transcription factor that drives intestinal progenitor cells towards secretory cell lineages to assure gut homeostasis. However, little is known about whether and how ATOH1 governs mucin production in an important histological subtype of colorectal cancer (CRC), termed mucinous CRC (MCC). Here we report that ATOH1 positively correlates with expression of secreted mucins. Moreover, we observed that a non-mucinous CRC xenograft expressing ATOH1, phenocopied mucinous histology. Using CRISPR/Cas9-mediated knockout and various biochemical studies, we demonstrate that the APC protein may be involved in facilitating ATOH1, however independent of either GSK3 or  $\beta$ -catenin, as was previously suggested. The previously reported JAK2-mediated phosphorylation of Y80 also moderately increased ATOH1 levels in CRC cell lines. Next, through mass-spectrometry analysis and co-immunoprecipitation, we found that the E3 ubiquitin ligase UBR5 can bind to and stabilize ATOH1 via its ubiquitin ligase activity in a subset of CRC cell lines. UBR5 exhibits this activity in a complex with or independent of DYRK proteins, for which we have observed to also be involved in ATOH1 stabilization using the DYRK2 and DYRK4 inhibitor ID-8. Lastly, we propose a potential biomarker signature in which *DYRK4* and *MUC2* jointly allow to molecularly distinguish mucinous CRC. Our findings may help to delineate mechanisms for controlling mucus production in CRC, thus providing new avenues for accurate diagnosis and treatment of CRC with mucinous phenotype.

## Introduction

Colorectal cancer (CRC) is the third most commonly diagnosed cancer in the world (1). Concerning histologic examination, of all the CRCs examined, 10–15% are arbitrarily defined as the so-called mucinous CRC (MCC) subtype, posing as a tumor composed of more than 50% extracellular mucin pools (2,3). This intra-tumoral mucus compartment serves as a barrier contributing to evasion of antitumor immune surveillance and resistance to chemotherapeutics, therefore, patients with MCC often have a poor cancer prognosis (3).

It is now well-accepted that the transcription factor Atonal homolog 1 (ATOH1) is a master regulator in triggering and determining intestinal progenitors towards a secretory cell fate, as exemplified by lack of secretory lineages (i.e., goblet, enteroendocrine, and Paneth) in mouse intestinal crypts upon ATOH1 depletion (4–7). Molecularly, ATOH1 directly regulates expression of mucin genes (e.g., *MUC2* and *MUC5AC*) and other secretory cell fate-specific genes such as *SPDEF* at the transcriptional level (4,8). Besides its functions in secretory differentiation for maintenance of intestinal homeostasis, ATOH1 is also deemed a tumor suppressor (9,10). Secretion of gel-forming mucins induced by sufficient ATOH1 can suppress tumor formation and progression in the mice colon (9,11–14). Of note, most human CRC tumor tissues and cell lines exhibited a low, even undetectable, *ATOH1* mRNA level and protein expression, attributable to frequent methylation of the ATOH1 CpG site or deletion in its genomic regions (9,10,15). However, a subset of CRCs expressed similar mRNA levels of *ATOH1* in both tumor and matched normal colon specimens. Nevertheless, these tumor tissues were characterized by a nearly complete absence of goblet cells, suggesting that ATOH1 protein is non-functional or proteolytically degraded (10). As such, posttranslational modifications (PTMs) of ATOH1 most likely contribute to its regulatory efforts to stimulate mucinous differentiation during colon tumor formation.

Cumulative evidence reveals that PTM modulations of ATOH1 and the resulting changes in its protein stability are indispensable for several homeostatic and physiopathological scenarios (7,15–19). However, most of these reports dealt with

ATOH1's role in brain development, while only few studies have investigated the PTM regulations of ATOH1 stability in colon cancer. One previous work argued that ATOH1 could be phosphorylated by glycogen synthase kinase 3 $\beta$  (GSK3 $\beta$ ) and subsequently targeted for proteasomal degradation in APC truncated CRC cells, proposing a reciprocal regulation of ATOH1 and  $\beta$ -catenin in intestinal cells (15). It is widely believed that APC inactivation is an early event that drives  $\beta$ -catenin-dependent carcinogenesis, with the highest frequencies (~80%) detected in sporadic CRC (20,21). APC serves as a scaffold, incorporating GSK3 $\beta$  and other components into the  $\beta$ -catenin destruction complex (22). Of note, the model of GSK3 $\beta$ -mediated reciprocal regulation of ATOH1 and  $\beta$ -catenin (15) builds upon an outmoded notion that APC truncation causes dissociation of the  $\beta$ -catenin destruction machinery. However, a strong body of evidence indicates that APC disruption hampers  $\beta$ -catenin ubiquitination and thereby promotes phosphorylated  $\beta$ -catenin saturating the APC/GSK3 $\beta$ / $\beta$ -catenin destruction complex, instead of directly disassembling this complex (22). Therefore, the scenario of a GSK3 $\beta$ -mediated reciprocal regulation may not fully explain the proteasomal degradation of ATOH1 in APC mutant CRC, in which the APC/GSK3 $\beta$  complex is compositionally intact. The exact molecular mechanisms underlying regulation of ATOH1 degradation/stabilization in mucinous differentiation within colorectal tumor mass remain obscure.

Here, we report that ATOH1 expression positively correlates with mucin expression, and show that ATOH1 overexpression confers a mucinous phenotype to non-mucinous CRC (NMC) cells. We demonstrate that APC contributes to ATOH1 stabilization, however independent of either GSK3 or  $\beta$ -catenin, as was previously suggested. We also show that the JAK2-mediated Y80 phosphorylation of ATOH1, previously identified in mouse brain tumors, may also lead to some ATOH1 stabilization in colorectal cancer cell lines. Next, using mass spectrometry we identify the E3 ubiquitin ligase UBR5 (aka EDD1) to interact with ATOH1. Together with the dual-specificity tyrosine phosphorylation-regulated kinase 2 and 4 (DYRK2/4), UBR5 protects ATOH1 from proteasomal degradation via its ubiquitin ligase activity in a subset of CRC cells. Finally, we propose a potential biomarker signature (i.e., DYRK4 and MUC2 in combination) for molecularly distinguishing MCC.

## Results

### ATOH1 drives CRC cells towards mucinous differentiation

To identify which genes are significantly altered in expression in MCC versus NMC, we performed a differential expression gene (DEG)-based analysis of CRC transcriptional profiles present in the TCGA cohort (**Fig. 1A**). Next, we focused our analysis on the mucin-related genes. Not surprisingly, among all the members of the mucin gene family only the expression of *MUC2*, *MUC5AC* and *MUC5B*, which encode gel-forming mucins, were highly upregulated in MCC tumor samples (**Fig. 1A, B**). Expression of *ATOH1* and its downstream target *SPDEF* were also markedly elevated in MCC tumor tissues (**Fig. 1A, B**). Similar trends were observed in the GSE2109 dataset (**Supplementary Fig. 1A**). Moreover, *ATOH1* expression was positively and strongly correlated with mRNA levels of *SPDEF* and the secreted mucin genes in tissues from large intestine, i.e. normal colon, colorectal cancers and cell lines thereof (**Fig. 1C–E; Supplementary Fig. 1B**). In addition, reduced mRNA levels of *SPDEF* (~4-fold) and *MUC2* (~2.8-fold) were observed in a MCC line LS174T with CRISPR/Cas9-mediated knockout of *ATOH1* (**Supplementary Fig. 1C and 2**). Conversely, forced expression of *ATOH1* imposed a goblet cell-like histology and strong PAS-positivity, indicative of the presence of mucins, onto the NMC cell line DLD1 in a subcutaneous xenograft model (**Fig. 1F**). These observations suggest that sufficient *ATOH1* may induce mucin-expressing cells in a tumor mass that phenocopies a mucinous tumor. Collectively, upregulation of *ATOH1* confers the ability to drive CRC cells towards mucinous differentiation.

**Fig. 1** *ATOH1* positively correlates with and regulates mucinous differentiation in colorectal cancer (CRC). (A) Volcano plot showing genes differentially expressed between mucinous CRC (MCC) and non-mucinous CRC (NMC) from the TCGA cohort. Vertical dashed lines denote log<sub>2</sub>-transformed fold change = ±1, and horizontal dashed line indicates adjusted *P*-value = 0.05. Dots representing *ATOH1*, *SPDEF*, and all members of the mucin gene family are highlighted in bold black. (B) Boxplots showing comparisons of mRNA levels of *ATOH1* and target genes thereof in tumor tissues between MCC and NMC from the TCGA cohort. (C–E) Association of mRNA expression of *ATOH1* with its downstream target genes across multiple datasets as indicated. *r* denotes Pearson correlation coefficient. (F) Representative images of hematoxylin and eosin (H&E) and periodic acid–Schiff (PAS) staining for xenografts of the non-mucinous CRC cell line DLD1 stably expressing Doxycyclin-induced *ATOH1* or a negative control (NC) vector. All data are presented as mean ± SD. Significance tested in (B) using Wilcoxon rank sum test.

(see figure on next page)

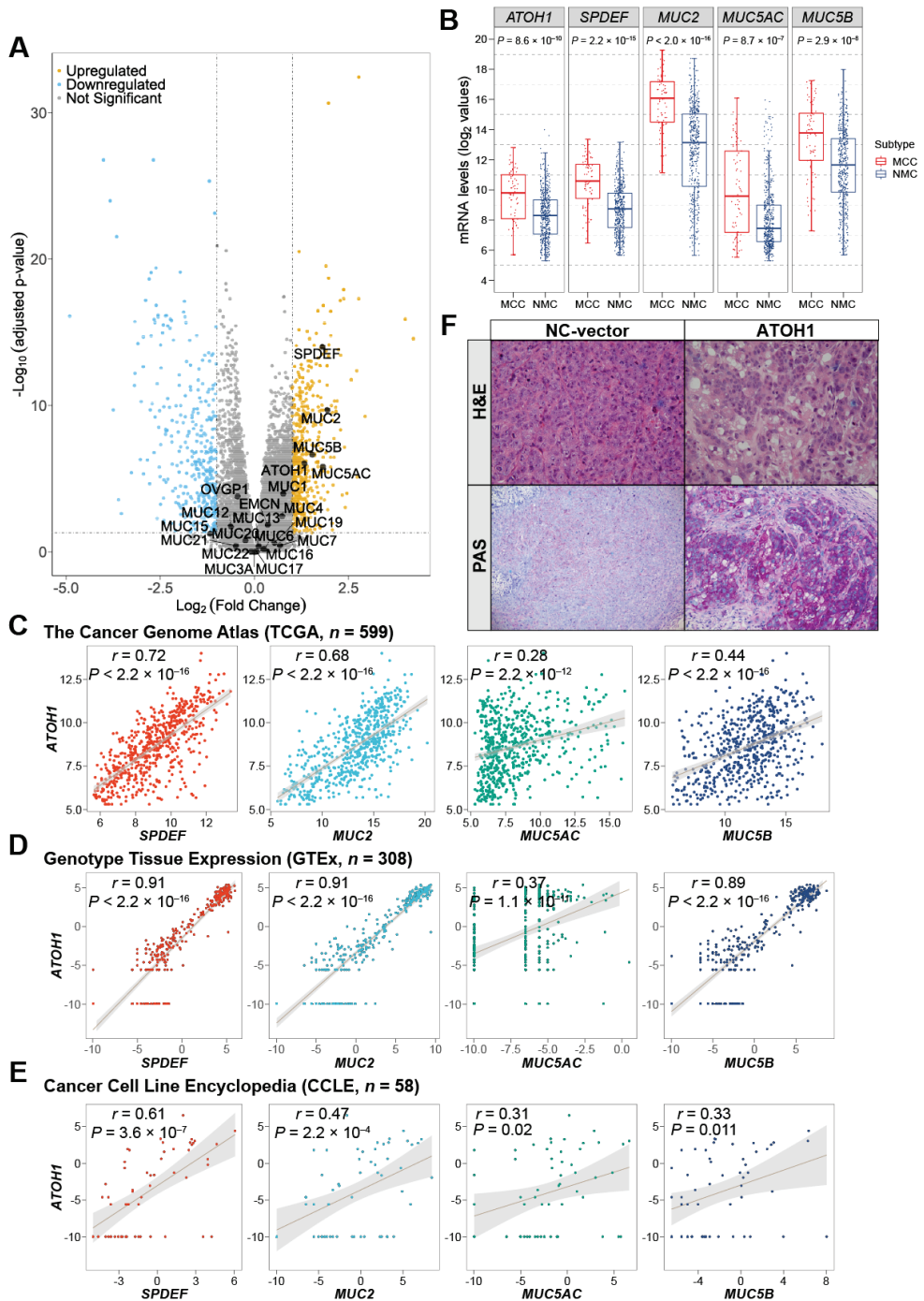
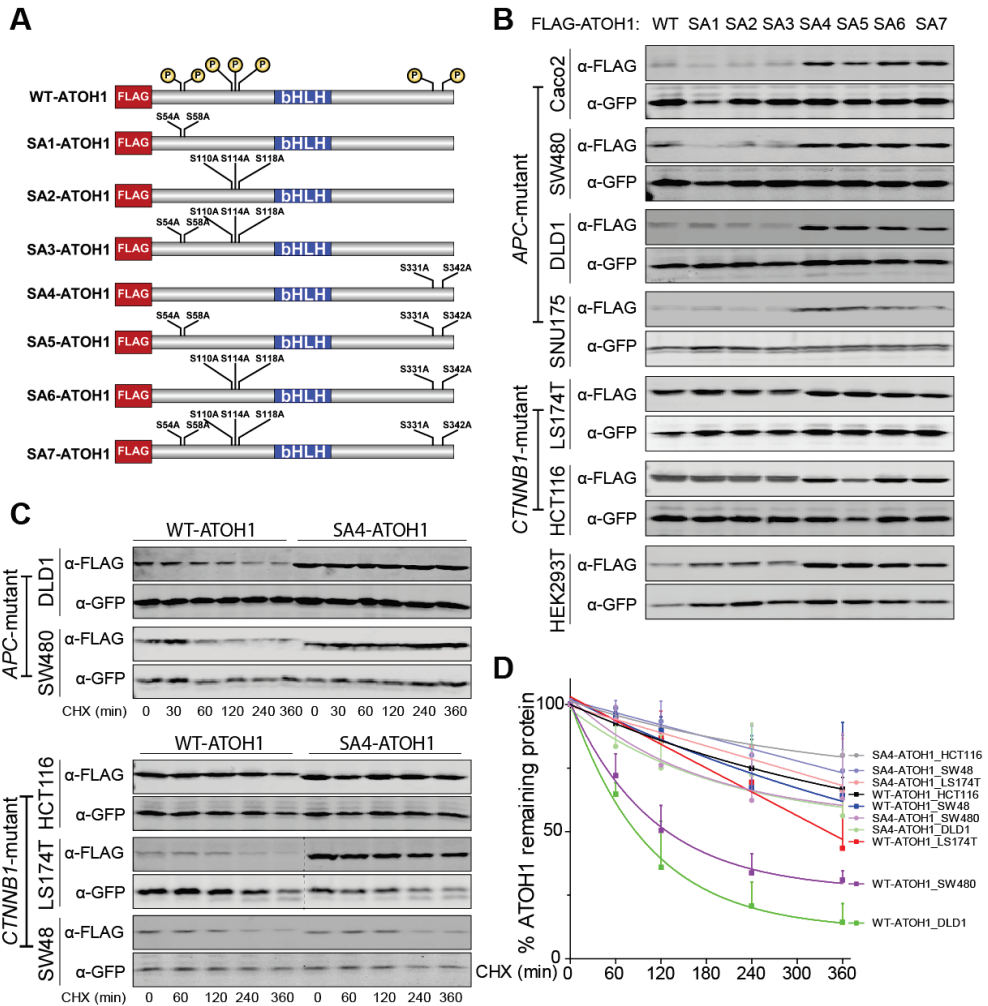


Fig. 1 (see legend on previous page)

### **Serine 331/342 phosphorylation sites are predominantly responsible for ATOH1 degradation**

The phosphorylation sites responsible for proteasomal degradation of ATOH1 remain controversial (15,16,23). To characterize and validate the reported phosphorylated residues, seven FLAG-tagged vectors expressing phospho-dead serine-to-alanine (SA) mutants of ATOH1 were constructed, representing the three previously reported phosphorylation clusters (S54/58A; S110/114/118A; S331/342A) alone or in combination with others (**Fig. 2A**). In all tested cell lines, the SA4, SA5, SA6 and SA7 variants, were clearly more stable than WT-ATOH1 and other variants, irrespective of whether the cell lines tested are NMC (Caco2, SW80, DLD1, and HCT116), MCC (SNU175 and LS174T), or embryo kidney (HEK293T) cell lines (**Fig. 2B**). Notably, these stable variants all harbor the S331/342A mutation, suggesting that phosphorylation at serines 331/342 is a key event in ATOH1 degradation. To confirm this finding, we next determined the turnover rate of WT and S331/S342A (SA4) ATOH1 variants using a cycloheximide chase assay (**Fig. 2C, D**). SA4-ATOH1 exhibited a greatly longer half-life compared with WT-ATOH1, especially in the APC-mutant SW480 and DLD1 cells (**Fig. 2C, D**). In further support, the ATOH1 serines 331/342 are highly conserved according to an evolutionary conservation analysis (consurf.tau.ac.il/2016) (24), in contrast to both other reported phosphorylation clusters (**Supplementary Fig. 3**). In accordance with previous reports (19,25), besides the S331/342 residues, the entire C-terminal serine-rich domain of ATOH1 represents a highly evolutionary conserved region (**Supplementary Fig. 3**). Overall, these analyses show that serines 331/342 located at C-terminal ATOH1 are critical residues responsible for its degradation.



**Fig. 2** Mutation of phosphorylation sites S331/S342 protects ATOH1 from degradation. **(A)** Schematic representation of FLAG-tagged vectors expressing wild-type (WT) ATOH1 and indicated serine-to-alanine (SA) mutants. **(B)** Protein expression of vectors expressing FLAG-tagged WT ATOH1 and SA-mutants thereof in the indicated cell lines, as determined by anti-FLAG immunoblotting. A GFP-expressing vector was co-transfected and served as transfection control. **(C, D)** The S331/342A (SA4) mutant exhibits increased stability with longer half-life than wild-type ATOH1. **(C)** Cycloheximide (CHX, 100  $\mu$ g/mL) was added at 40 h post-transfection for the indicated times. Immunoblotting showing the protein expressions of WT and SA4 ATOH1 vectors upon CHX treatment at indicated time points using anti-FLAG antibody. The vertical dashed line indicates the removal of irrelevant lanes from the same blot. **(D)** The band intensities of WT and SA4 ATOH1 in **(C)** were normalized to corresponding GFP values, then shown as a percentage to the value of CHX addition starting at time 0. Error bars represent  $\pm$  SD from three independent experiments.

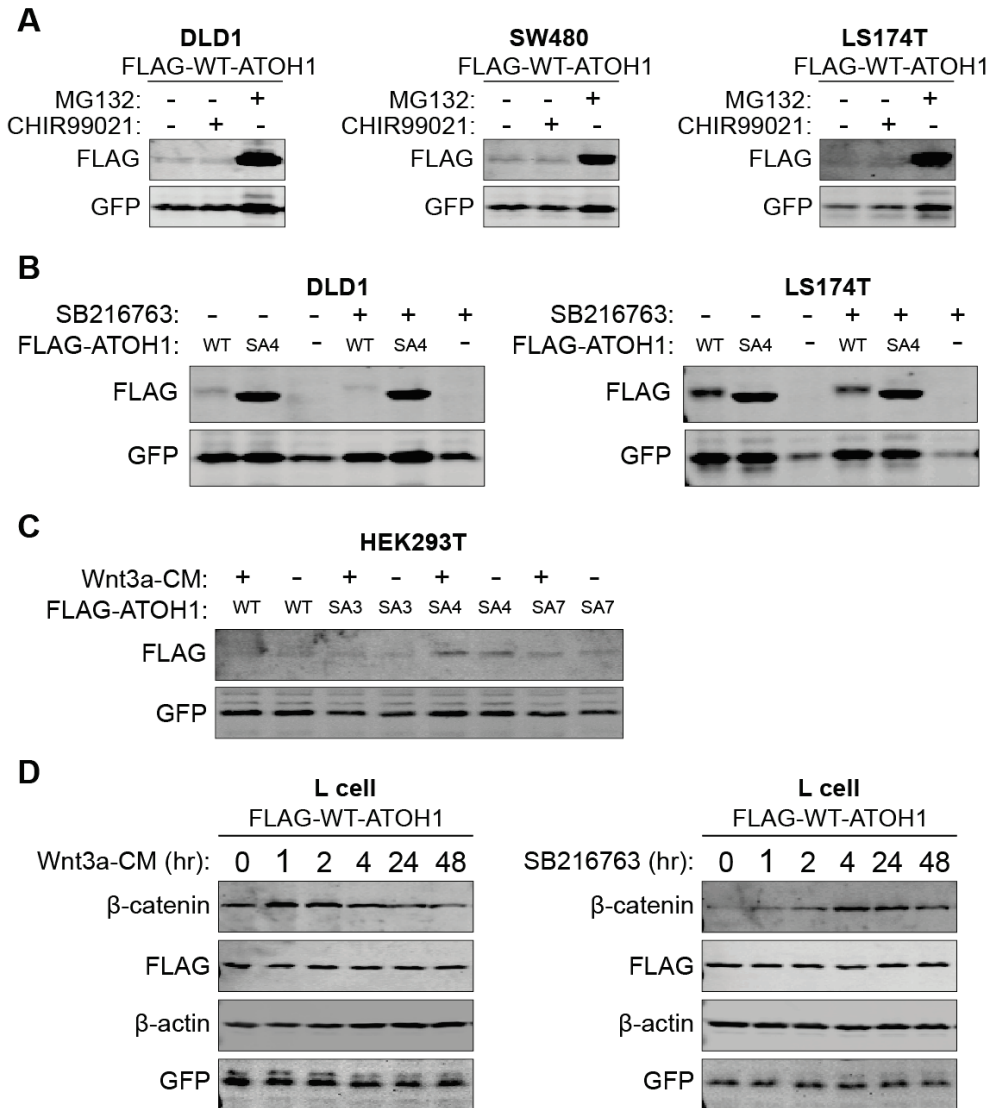


## **Evaluation of previously reported interactors of ATOH1 in its degradation**

We next sought to evaluate previously reported interactors of ATOH1 involved in its phosphorylation and degradation. Consistent with previous studies (15,16), the protein level of ATOH1 was remarkably increased following exposure to the commonly used proteasome inhibitor MG132, further confirming that ATOH1 degradation is proteasome-dependent (**Fig. 3A**).

Previously, GSK3 $\beta$  has been proposed to phosphorylate and thereby induce ATOH1 degradation, as the authors showed that a GSK3 inhibitor stabilized ATOH1 in APC-mutant CRC cells (15). However, we cannot reproduce this result in three CRC lines using the GSK3 inhibitors CHIR99021 (**Fig. 3A**) or SB216763 (**Fig. 3B**). Furthermore, levels of WT-ATOH1 were not reduced following Wnt3a-conditioned medium (Wnt3a-CM) treatment in HEK293T cells, which disagrees the result of Tsuchiya et al. that the protein level of wild-type ATOH1 was decreased in response to Wnt stimulation (**Fig. 3C**). To confirm these observations further, L-cells that harbor neither APC mutations nor mutated genes involved in the  $\beta$ -catenin pathway were employed. L-cells were treated with Wnt3a-CM or the GSK3 inhibitor SB216763 for several time periods, in which the endogenous  $\beta$ -catenin was probed for demonstrating the validity of treatments, not leading to any alteration in ATOH1 protein levels (**Fig. 3D**). Taken together, these data show that GSK3 is not involved in ATOH1 stabilization in CRC cells.

Sonic hedgehog (SHH) inhibition with vismodegib (GDC-0449) was also reported to increase phosphorylation levels of murine ATOH1 at both serine 328 and serine 339 (corresponding to serines 331/342 of human ATOH1), thereby reducing ATOH1 stability in mouse neural progenitor cells (16). Surprisingly, we observed the opposite result, that is WT-ATOH1 levels were increased in response to SHH inhibition in the human NMC (DLD1) and MCC (LS174T) cell lines (**Supplementary Fig. 4A, B**), indicating that there might be alternative mechanisms for SHH-mediated regulation of ATOH1 stability in colonic tissue.



**Fig. 3** ATOH1 proteasomal degradation is independent of GSK3 or Wnt3a stimulation. **(A)** DLD1, SW480, and LS174T cells were transfected with FLAG-tagged WT-ATOH1 for 48 h, treated with/wo the proteasomal inhibitor MG132, and in the presence/absence of the GSK3 inhibitor CHIR99021, and harvested for immunoblotting. **(B)** DLD1 and LS174T cells were transfected with FLAG-tagged ATOH1 variants (WT and SA4) for 48 h, treated with or without a second GSK3 inhibitor SB216763, and harvested for immunoblotting. **(C)** HEK293T cells were transfected with FLAG-tagged ATOH1 variants (WT, SA3, SA4, and SA7) for 48 h, treated with/wo Wnt3a-CM, and harvested for immunoblotting. **(D)** L-cells were transfected with FLAG-tagged WT-ATOH1 for 48 h, treated with Wnt3a-CM or SB216763 for the indicated time periods, and harvested for immunoblotting.

**APC, independent of  $\beta$ -catenin, stabilizes ATOH1 protein in CRC cells**

Tsuchiya and coworkers reported that functional, not deficient APC is capable of stabilizing ATOH1 (15). Moreover, others reported that APC re-expression in APC-mutant SW480 cells (26), resulted in a significant 2-fold increase in *MUC2* expression (**Supplementary Fig. 5**), in line with similar experiments in HT29 cells (10). We thus investigated whether the APC functional status associates with the mucinous phenotype and expression of secreted mucins in CRC samples from the TCGA cohort. Given that loss-of-function (LoF) of the APC protein is mainly caused by APC truncating mutations, and most APC missense mutations are not likely to contribute to colonic tumor development, only CRC samples carrying frameshift and nonsense mutations of APC were included in this analysis (27,28). Moreover, we and others have reported that truncated APC mutants that retain at least one of the AXIN-binding domains, retain their function to properly regulate  $\beta$ -catenin signaling (28–30). Therefore, we set a demarcation at 1581-AA for determining whether APC is deficient or not. As shown in **Fig. 4A**, MCC had a significantly lower frequency of APC deficiency than NMC. Moreover, APC deficiency was less frequent in CRC samples with high expression of the mucin-related genes *SPDEF*, *MUC2*, *MUC5AC*, and *MUC5B* (**Fig. 4A**). Consistently, the APC-deficient tumors displayed greatly reduced expression of these ATOH1 downstream genes compared with APC-proficient CRC samples (**Fig. 4B**). However, *ATOH1* mRNA expression was not significantly different in APC-deficient versus proficient tumors (**Fig. 4A, B**). These results together suggest that APC functional status correlates with the expression of ATOH1 target genes. As *ATOH1* mRNA levels are not clearly different, it implies that APC function may be linked to ATOH1 protein turnover.

Stabilized expression of ATOH1 was shown in SW480 cells while overexpressing an APC functional fragment (15). In accordance, our results showed that turnover of WT-ATOH1 in APC-mutant CRC cells (SW480 and DLD1) is much faster than in APC-wildtype lines (HCT116, LS174T, and SW48) (**Fig. 2C, D**). We further validated these results by CRISPR/Cas9-mediated generating knockout (KO) of full-length APC in

HCT116 and LS174T cells and knocking-out the truncated p.I1417fs\*2 APC variant in DLD1 cells (**Supplementary Fig. 6A, B**). WT-ATOH1 was unstably expressed in three independent APC-KO clones of all these CRC cell lines (**Fig. 4C–E**), associated with a reduced half-life (**Fig. 4F, G**).

A previous study indicated that the APC substrate  $\beta$ -catenin upregulates *ATOH1* mRNA expression (31). We sought to determine whether  $\beta$ -catenin is also involved in regulation of ATOH1 degradation. To this aim we overexpressed wild-type  $\beta$ -catenin and a degradation-resistant variant thereof (32), in APC-wildtype HCT116 and APC-mutant DLD1 cells. The stability of co-expressed ATOH1 was unaltered, demonstrating that the APC-associated ATOH1 stability is not affected by  $\beta$ -catenin expression (**Supplementary Fig. 7A, B**). This is consistent with the observation that no changes in the ATOH1 stability are observed following exposure to Wnt3a-CM or GSK3 inhibitor that induced a clear increase in endogenous  $\beta$ -catenin expression (**Fig. 3D**). Collectively, our results suggest that functional APC contributes to ATOH1 stabilization in CRC, which is independent of  $\beta$ -catenin expression.

**Fig. 4** Loss of APC reduces ATOH1 stability and leads to downregulation of ATOH1 downstream target genes. **(A)** Comparisons of APC deficiency in samples of colorectal cancer (CRC) from the TCGA cohort between mucinous CRC (MCC,  $n = 63$ ) and non-mucinous CRC (NMC,  $n = 404$ ), and between the upper ( $n = 117$ ) and lower ( $n = 117$ ) quartile of expressions of *SPDEF*, *MUC2*, *MUC5AC*, *MUC5B*, and *ATOH1*. Only samples with at least one frameshift or nonsense mutation that truncates APC before codon 1581 were considered APC-deficient, otherwise the samples were considered APC-proficient. **(B)** APC-proficient (i.e., truncation beyond 1581, denoted as “+”) CRC samples exhibited significantly higher mRNA levels of *SPDEF*, *MUC2*, *MUC5AC*, and *MUC5B* than APC-deficient (i.e., amino acids  $\leq 1580$ , denoted as “–”) tissues from the TCGA cohort, whereas similar *ATOH1* mRNA levels were observed. **(C–G)** knockout (KO) of APC reduces ATOH1 stability. **(C–E)** Parental and three independent APC-KO clones of CRC cell lines HCT116, DLD1, and LS174T, were transfected with ATOH1 vectors expressing either the wild-type (WT) or S331/342A (SA4) variant. Upper panels show protein expressions as determined using immunoblotting (GFP was used as transfection control). Lower panels show quantification of the band intensities normalized to GFP values, and then shown as relative levels to the value of the parental cell line ( $n = 2$  independent experiments). **(F, G)** Cycloheximide (CHX) chase assay showing that wild-type ATOH1 has a longer protein half-life in parental HCT116 and DLD1 cells than in the APC-KO clones. Upper panels show immunoblotting images, while the lower panels show quantifications of relative protein levels ( $n = 3$  independent experiments). All data are expressed as mean  $\pm$  SD. Significance tested using Fisher's exact test **(A)**, and Wilcoxon rank sum test **(B)**.

(see figure on next page)

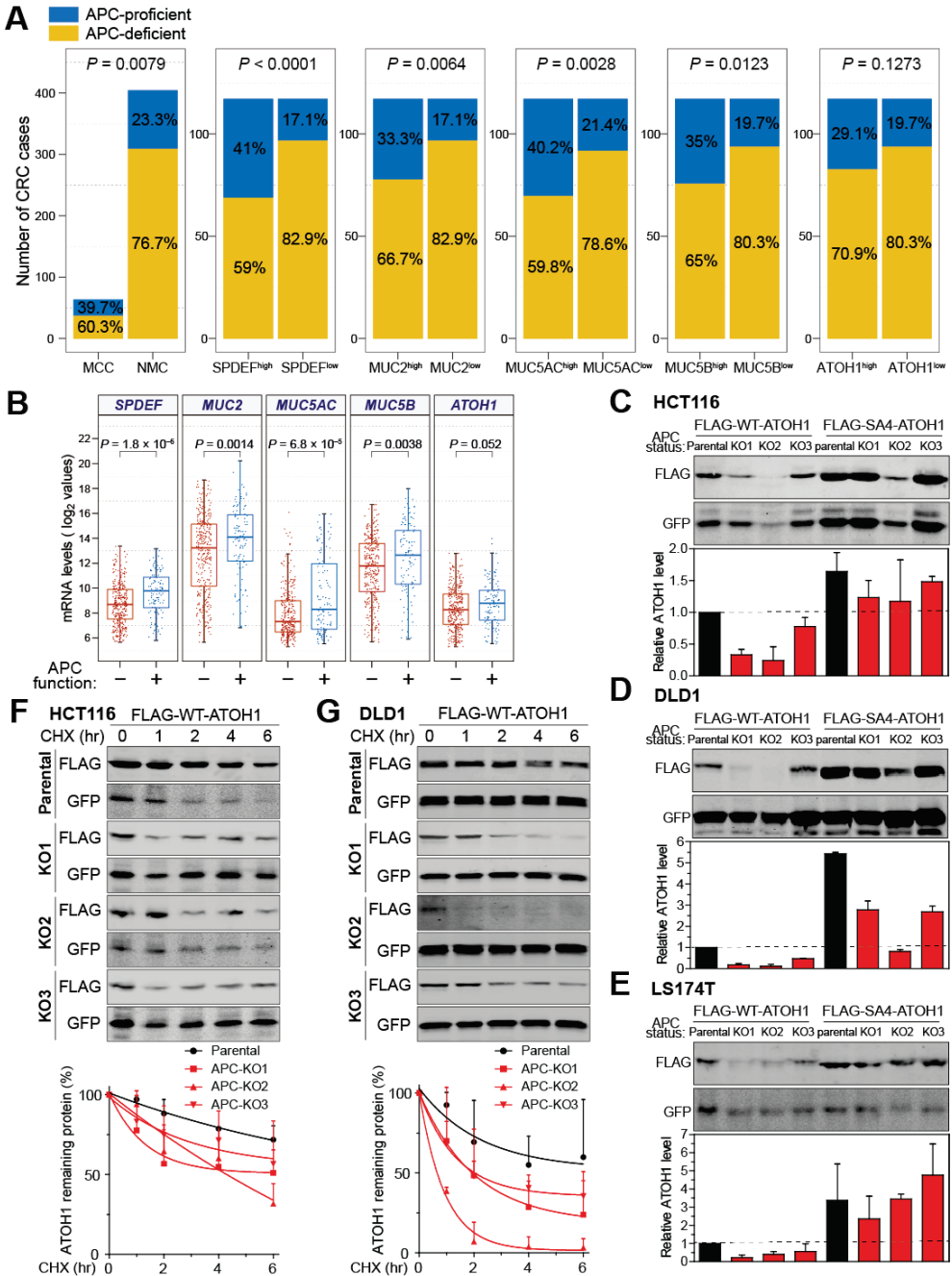


Fig. 4 (see legend on previous page)

**UBR5 interacts with and protects ATOH1 from degradation through its ubiquitin ligase activity**

To identify new ATOH1-binding proteins we employed immunoprecipitation and mass spectrometry on FLAG-tagged ATOH1 expressed in DLD1 and LS174T cells. The E-proteins TCF3 and TCF12, both well-known ATOH1 interactors (17), were identified in the anti-FLAG immunoprecipitates, indicative of the effectiveness of our mass spectrometry analysis (**Supplementary Fig. 8**). Overlaps between our dataset and a published dataset containing potential ATOH1 interactors (17) retrieved from the BioGRID (thebiogrid.org) (33) were presented as a Venn diagram (**Supplementary Fig. 8**). Only one E3 ubiquitin ligase was shared in both datasets, i.e. UBR5, which also possessed a high confidence score (**Supplementary Fig. 8**). First, we validated the interaction of UBR5 with ATOH1. GFP-tagged wild-type UBR5 and a catalytically inactive mutant ( $\Delta$ HECT) (34), co-immunoprecipitate with the FLAG-tagged WT or SA4 (S331/342A) ATOH1 variants, while the GFP-UBR5 variants were not detected in the negative control samples, suggesting that UBR5 specifically binds to ATOH1 (**Fig. 5A**).

Next, the function of UBR5 on ATOH1 breakdown was examined. Interestingly, *UBR5* knockdown leads to a reduction in WT-ATOH1 expression (**Fig. 5B**). Conversely, UBR5 overexpression increased the ATOH1 expression level in a subset of CRC cells (DLD1, GP5D, SW48, LOVO, and CW-2) (**Fig. 5C**), but not in all tested lines (**Supplementary Fig. 9**). We also observed that UBR5 mainly stabilizes the apparently phosphorylated form of ATOH1 (the upper band as shown by immunoblotting) (**Fig. 5C**; **Supplementary Fig. 10**). Importantly, comparing the effects of wild-type and the ubiquitin ligase-dead mutant ( $\Delta$ HECT) UBR5 on stabilizing ATOH1, we can conclude that UBR5 increases ATOH1 stability through its ubiquitin ligase activity since the UBR5- $\Delta$ HECT mutants failed to result in any increase in ATOH1 levels (**Fig. 5C**).

The phenotype of UBR5-mediated ATOH1 stabilization was then reproduced by three independently generated CRISPR/Cas9-mediated UBR5 knockout clones of HCT116 and DLD1 cells (**Supplementary Fig. 11**). Consistently, a reduction in the

expression level of WT-ATOH1, but not SA4 (S331/342A)-ATOH1, was observed in UBR5-KO clones, suggesting that UBR5 stabilizes ATOH1 through preventing ATOH1 from degradation (**Fig. 5D, E**). Notably, decreased WT-ATOH1 expression in the UBR5-KO cells could be rescued upon UBR5 re-expression (**Fig. 5F**). In addition, the turnover of WT-ATOH1 is faster in UBR5-knockout clones than in parental cells (**Fig. 5G, H**). Taken together, UBR5 protects ATOH1 from protein degradation through its ubiquitin ligase activity, at least in a subset of CRC cells.

We next sought to determine potential interactions of ATOH1 with APC and UBR5. In agreement with the reduction of ATOH1 expression observed in APC-KO CRC cells, the WT-ATOH1 level was attenuated by APC knockdown in DLD1 cells (**Fig. 6A**). Overexpression of UBR5, but not its catalytically inactive counterpart, led to a 2–3-fold increase in ATOH1 levels however independent of the underlying APC status. This was further confirmed in APC knockout DLD1 cells (**Fig. 6B**). These results imply that APC is not needed for UBR5 to protect ATOH1 from degradation.

---

**Fig. 5** UBR5 interacts with and stabilizes ATOH1 by its ubiquitin ligase activity. **(A)** Specific interaction of ATOH1 with UBR5 was determined by co-immunoprecipitation (co-IP) with anti-FLAG M2 affinity gel, followed by immunoblotting. HCT116 cells were co-transfected with GFP-tagged UBR5 vectors expressing wild-type (WT) or catalytically inactive ( $\Delta$ HECT) variants, and wild-type ATOH1 vector or empty vector (EV) for 48 h. IP, immunoprecipitation; IB, immunoblot. **(B)** Immunoblotting showing that siRNA-mediated knockdown of UBR5 decreased ATOH1 protein expression in HCT116 cells. **(C)** Immunoblotting showing that WT-UBR5 overexpression increased ATOH1 protein levels in the indicated colorectal cancer (CRC) cells. **(D, E)** Proteins levels of FLAG-tagged WT and S331/342A (SA4) ATOH1 vectors in UBR5 knockout (KO) clones of HCT116 and DLD1 cells were assessed using immunoblotting by the indicated antibodies. Band intensities of WT and SA4 ATOH1 were normalized to GFP values, then shown as relative levels to the value of parental cell. **(F)** Immunoblotting analysis of protein expression of WT and SA4 ATOH1 variants in parental and UBR5-KO DLD1 cells in the presence/absence of UBR5 overexpression. **(G, H)** Cycloheximide (CHX) chase assay to measure the protein half-life of wild-type ATOH1 in parental and UBR5-KO HCT116 and DLD1 cells. Upper panels, immunoblotting images. Lower panels, quantifications of relative protein levels ( $n = 3$  independent experiments). Error bars represent  $\pm$  SD from three independent experiments. Values below the immunoblotting images of **(B, C)** show quantification of relative FLAG-tagged ATOH1 protein level normalized to GFP value.

*(see figure on next page)*

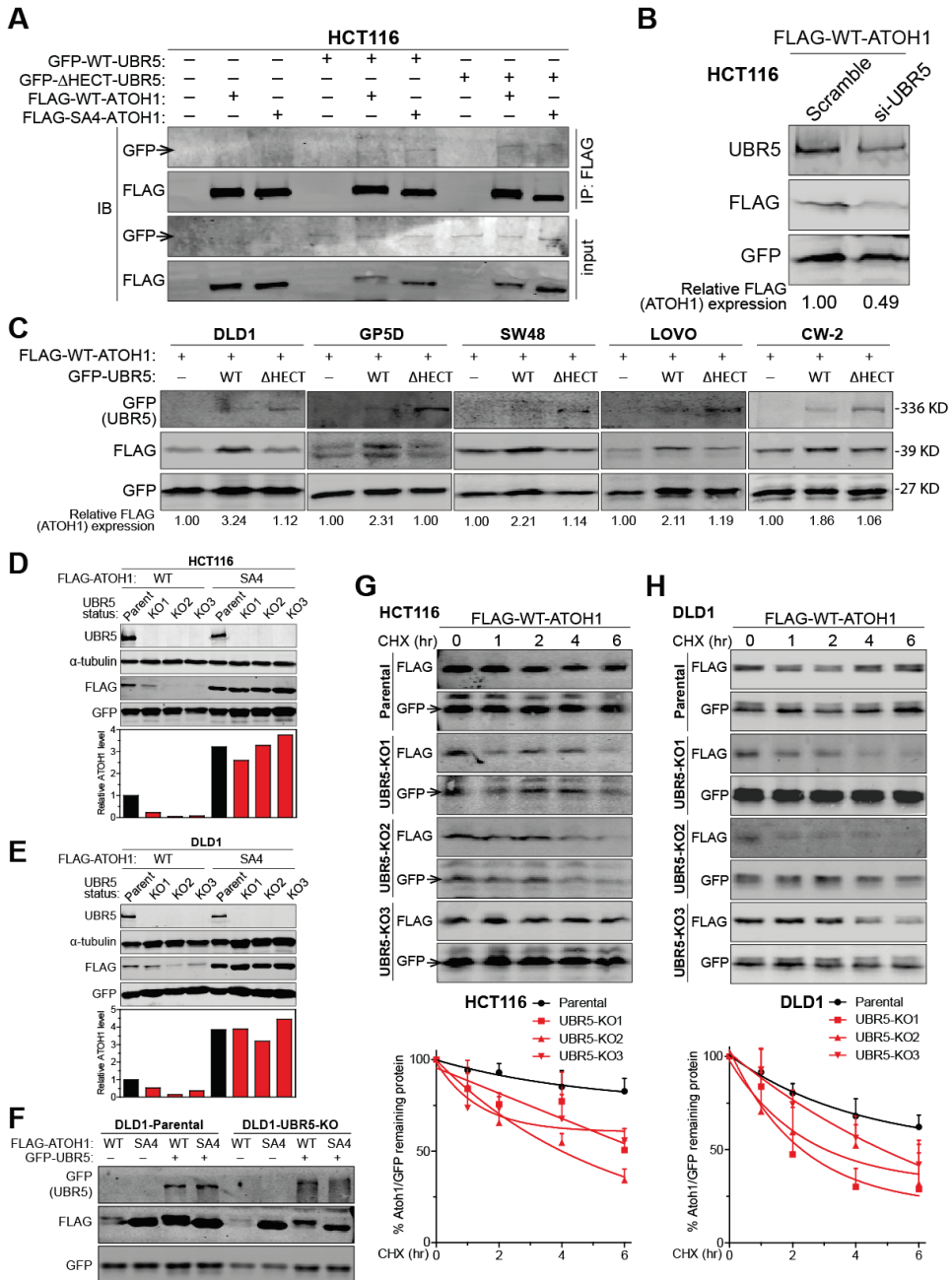
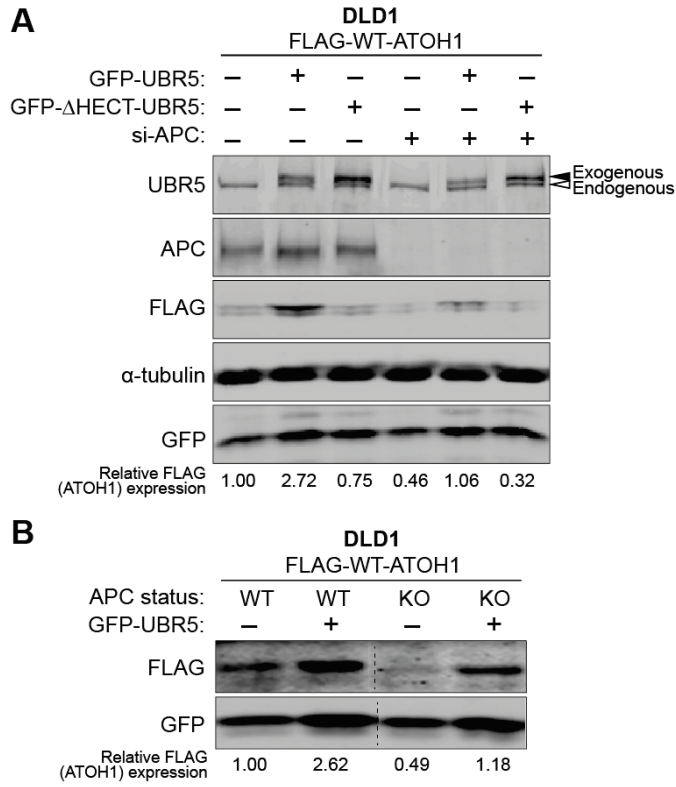


Fig. 5 (see legend on previous page)





**Fig. 6** APC is not required for UBR5-mediated ATOH1 stabilization. **(A)** Immunoblotting showing protein levels of wild-type (WT) ATOH1 upon overexpression of WT or catalytically inactive ( $\Delta$ HECT) UBR5 variants with/wo siRNA-mediated APC knockdown in DLD1 cells. **(B)** Immunoblotting showing WT and S331/342A (SA4) ATOH1 variants in parental and APC knockout (KO) DLD1 cells in the presence/absence of WT-UBR5 overexpression. The vertical dashed line indicates the removal of irrelevant lanes from the same blot.

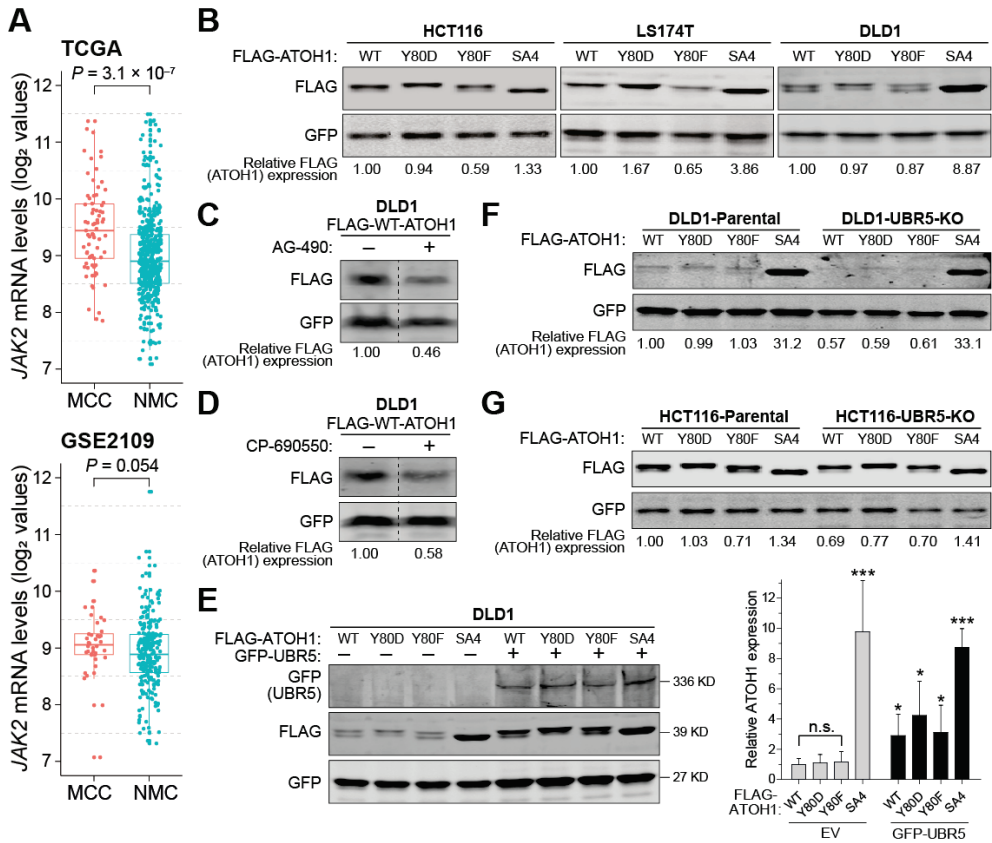
## **JAK2-mediated phosphorylation of ATOH1 at Y80 affects its stability also in CRC cells**

A previous report indicated that ATOH1 phosphorylation at tyrosine 78 (Y78, corresponding to the Y80 residue in human ATOH1) by JAK2 leads to ATOH1 stabilization in mouse brain tumors (18). Interestingly, analysis of two publicly available databases revealed that MCC exhibit higher *JAK2* mRNA levels than NMC (Fig. 7A). To investigate if Y80 phosphorylation also affects ATOH1 stability in CRC cells, we constructed Y80 variants of the human ATOH1 homolog, including Y80D (phospho-mimetic) and Y80F (phospho-dead). We tested stability of these constructs in three CRC cell lines. In HCT116 and LS174T cells the phospho-dead (Y80F) variant showed reduced stability, while this was not clearly the case in DLD1 (Fig. 7B). Applying a JAK2 inhibitor (AG-490) or pan-JAK inhibitor (CP690550) on WT-ATOH1 transfected DLD1 cells, led to a 2-fold reduction in ATOH1 levels (Fig. 7C, D). Taken together, these results indicate that also in CRC cells JAK2 may be involved in regulating ATOH1 levels.

Next, we evaluated whether Y80 phosphorylation is also relevant for the UBR5-mediated stabilization of ATOH1. Forced expression of UBR5 in DLD1 cells led to a 3–4-fold increase of all Y80 ATOH1 variants (Fig. 7E). In the converse experiment in which UBR5 was knocked-out in DLD1 and HCT116 cells, we observed reduced ATOH1 expression apparently independent of the Y80 variant being used (Fig. 7F, G). In other words, UBR5's ability to affect ATOH1 stability appears to be independent of the Y80 phosphorylation site.

## **UBR5 expression is not associated with MCC**

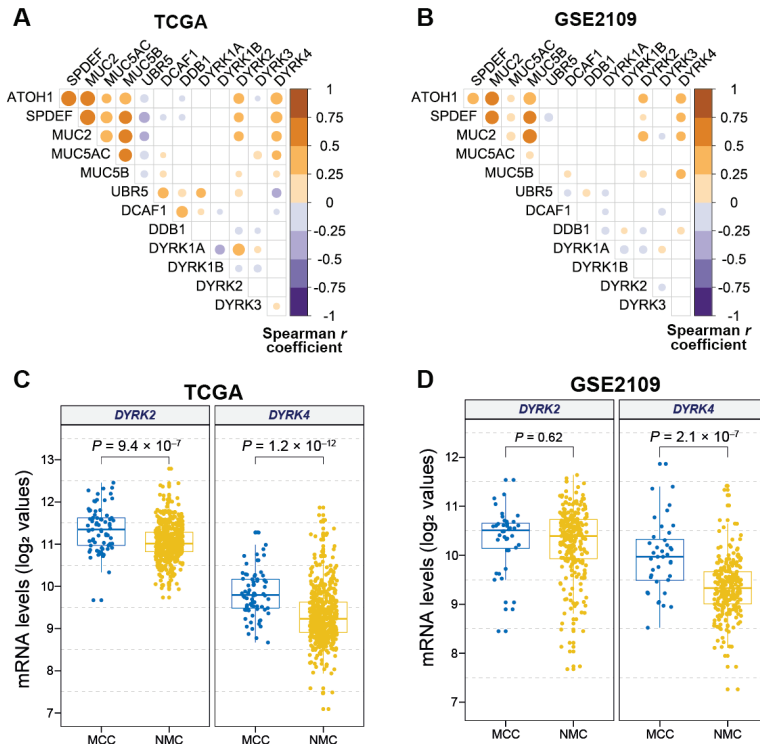
We were interested whether an association of UBR5 expression with mucinous CRC could be an explanation for the ATOH1-driven mucinous differentiation in CRC. UBR5 upregulation was however not observed in mucinous CRC cells (LS174T) compared with other cell lines exhibiting a non-mucinous phenotype (Supplementary Fig. 12A). Furthermore, the expression levels of *UBR5* were comparable in MCC and NMC present in the TCGA cohort (Supplementary Fig. 12B, C).



**Fig. 7** JAK2-mediated phosphorylation of ATOH1 at Y80 affects its stability in CRC cells. **(A)** Comparison of JAK2 mRNA levels between mucinous CRC (MCC) and non-mucinous CRC (NMC) in the TCGA (upper panel) and GSE2109 (lower panel) datasets. **(B)** Immunoblotting showing protein levels of WT, Y80D (phospho-mimetic), Y80F (phospho-dead), and SA4 ATOH1 variants in the HCT116 (left panel), LS174T (middle panel), and DLD1 (right panel) cells. **(C, D)** Immunoblotting showing protein levels of WT-ATOH1 in the DLD1 cells with/without inhibition of JAK2 using AG-490 **(C)** or inhibition of pan-JAK using CP-690550 **(D)**. The vertical dashed line indicates the removal of irrelevant lanes from the same blot. **(E)** Left panel, immunoblotting showing protein levels of WT, Y80D, Y80F, and SA4 ATOH1 variants in the absence/presence of UBR5 overexpression in DLD1 cells. Right panel, the band intensities of ATOH1 variants were normalized to GFP values, then represented as relative level to the value obtained for WT-ATOH1 without UBR5 overexpression that was arbitrarily set to 1 ( $n = 5$  independent experiments). **(F, G)** Immunoblotting showing protein levels of WT, Y80D, Y80F, and SA4 ATOH1 variants in the parental and UBR5-knockout (KO) DLD1 **(F)** and HCT116 **(G)** cells. All data are expressed as mean  $\pm$  SD. Significance tested using Wilcoxon rank sum test **(A, B)**, and unpaired Student's *t*-test **(E)**. Values below the immunoblotting images show quantification of relative FLAG-tagged ATOH1 protein level normalized to GFP value.

## Expressions of *DYRK2* and *DYRK4* positively correlate with mucinous phenotype in CRC

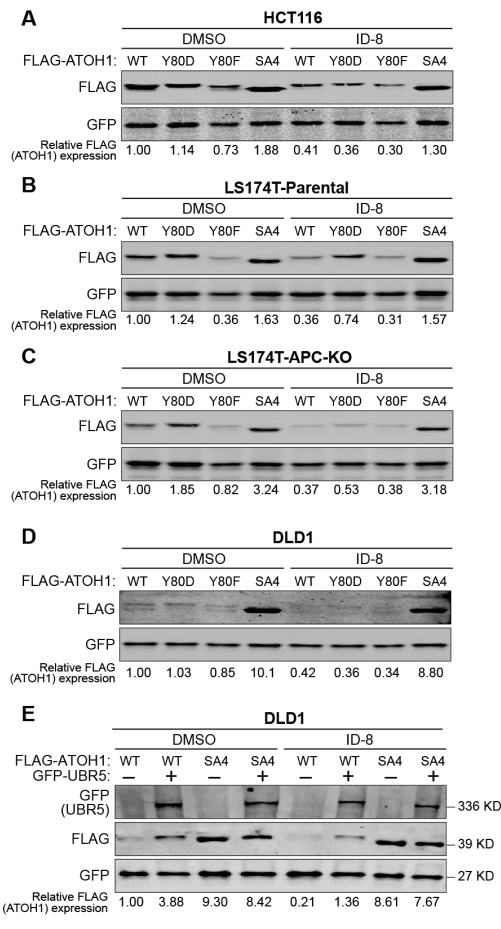
We next sought to explore whether co-effectors of UBR5 are involved in ATOH1 stabilization. Previous evidence indicated that UBR5 is able to bind with DCAF1 (aka VprBP), DDB1, and dual-specificity tyrosine phosphorylation-regulated kinase 2 (*DYRK2*) to form a complex (35). We therefore conducted analyses of Spearman's rank correlation between ATOH1 and its target genes, UBR5 and its co-effectors, and *DYRK* family members. Interestingly, only *DYRK2* and *DYRK4* correlated significantly with ATOH1 and its downstream genes (Fig. 8A, B). Furthermore, MCC samples had higher mRNA levels of *DYRK2*, and especially *DYRK4*, compared with NMC specimens (Fig. 8C, D).



**Fig. 8** Association of elevated expression of *DYRK2* and *DYRK4* with the mucinous phenotype in colorectal cancer. (A, B) Correlation matrices among the *ATOH1*-associated expression profiles of colorectal cancer (CRC) samples from the TCGA (A) and GSE2109 (B) cohorts. Spearman rank test was performed. Shading color represents the value of corresponding correlation coefficients. Non-significant ( $P > 0.05$ ) correlations are hidden. (C, D) Comparisons of mRNA levels of *DYRK2* and *DYRK4* in tumor tissues of mucinous versus non-mucinous CRC in TCGA (C) and GSE2109 (D) datasets.

DYRK inhibition impairs ATOH1 stabilization

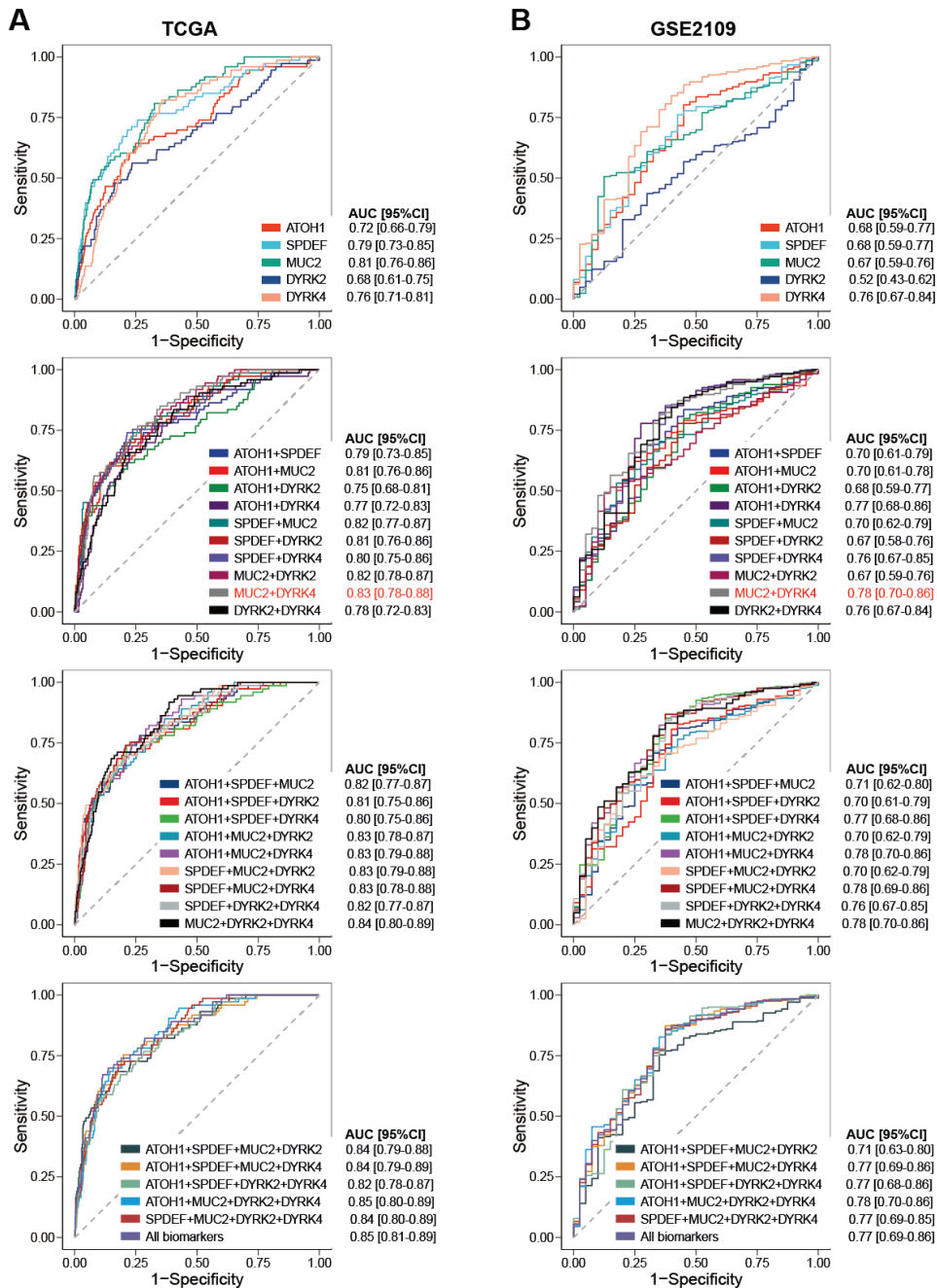
We next employed ID-8, a small molecule inhibitor targeting DYRK2 and DYRK4 (36), to uncover their functions in ATOH1 stabilization. Its application leads to a 2.4–2.8-fold reduction of wild-type ATOH1 levels in all lines tested (Fig. 9A–D, compare lanes 1 and 5), further supporting our speculation that the DYRK proteins may be functionally linked to mucinous differentiation. Next, we investigated whether DYRK’s activity may be linked to Y80 phosphorylation. However, all Y80 variants were reduced to the same extent, indicating that Y80 is not relevant for DYRK’s protective activity. Knock-out of APC investigated in one clone of LS174T also was unable to affect the reduction in ATOH1 levels induced by ID-8. Lastly, ID-8 treatment ablated the effect of UBR5 overexpression on stabilizing WT-ATOH1, suggesting that UBR5-mediated ATOH1 stabilization is compromised upon DYRK inhibition (Fig. 9E).



**Fig. 9** Inhibition of DYRKs suppresses UBR5-mediated ATOH1 stabilization. (A–D) Immunoblotting analyses showing protein levels of wild-type (WT), Y80D (phospho-mimetic), Y80F (phospho-dead), and S331/342A (SA4) ATOH1 variants in HCT116 (A), parental LS174T (B), APC knockout LS174T (C), and DLD1 (D) cells with/wo inhibition of DYRKs using ID-8. (E) Immunoblotting showing protein levels of WT and SA4 ATOH1 variants in the absence/presence of UBR5 overexpression with/wo ID-8 treatment in DLD1 cells. Values below the immunoblotting images show quantification of relative FLAG-tagged ATOH1 protein level normalized to GFP value.

**ATOH1 downstream target genes combined with *DYRK4* as a biomarker signature for detecting CRC with mucinous phenotype**

Finally, we investigated whether combination of *DYRK2* and *DYRK4* expression with ATOH1 and its downstream target genes can be applied for MCC versus NMC discrimination. To this end, we performed a receiver operating characteristic (ROC) analysis to evaluate the discriminatory ability of the individual or joint gene signatures for predicting MCC cases in the TCGA (**Fig. 10A**) and GSE2109 (**Fig. 10B**) datasets. As individual marker, *DYRK4* exhibited a good diagnostic power comparable to *ATOH1*, *SPDEF* and *MUC2*, while *DYRK2* expression is the poorest predictor of MCC (**Fig. 10A, B**). *ATOH1* alone showed comparable performance to its downstream genes (i.e., *SPDEF* and *MUC2*) in discriminating MCC from NMC across the two datasets (**Fig. 10A, B**). Importantly, the *DYRK4*-comprising multivariate panel performed improved discriminatory ability for mucinous CRC. The area under the curve (AUC) values for joint signatures in which *DYRK4* is included, were generally higher compared with the AUCs for other combinations (**Fig. 10A, B**). Especially, *DYRK4* and *MUC2* in combination yielded an AUC of 0.83 and 0.78 in the TCGA and GSE2109 cohorts, respectively (**Fig. 10A, B**). Taken together, a gene signature consisting of *DYRK4* and ATOH1 target genes could serve as a potent biomarker with good accuracy for distinguishing CRCs with mucinous phenotype from non-mucinous tumors.



**Fig. 10** Performance of individual and joint ATOH1-associated gene signatures for predicting mucinous CRC. (A, B) Receiver operating characteristic (ROC) curves showing the discriminatory ability of individual and joint biomarker signatures in predicting mucinous CRC in the TCGA cohort (A) and GSE2109 cohort (B). Area under the ROC curve (AUC) value and corresponding 95% confidence interval (CI) in brackets for each biomarker signature are provided. The AUC value and 95% CI for DYRK4 combined with MUC2 are highlighted in red.

## Discussion

To date, the molecular mechanisms modulating mucinous differentiation in the context of colonic tumor formation are still largely unclear. In the present study, we demonstrate an important role of ATOH1 in such differentiation process as shown, among others, by the induction of mucinous differentiation upon forced ATOH1 expression in the non-mucinous CRC line DLD1 (**Fig. 1F**). The phenotypic response of NMC cells to ATOH1 overexpression infers that cells in the undifferentiated state in the tumor context retain the ability to be differentiated into mucus-producing cells in the presence of sufficient ATOH1 expression. Our analysis of CRC transcriptional profiles reveal that secreted mucin genes, which are downstream targets of ATOH1, are significantly upregulated in MCC. Moreover, ATOH1-overexpressing CRC cells exhibited increases in transcription of mucin-related genes (15,23,37,38). Upregulation of ATOH1 was also observed in MCC and its precursors (hyperplastic polyp, serrated adenoma, and villous adenoma), but not in NMC and tubular adenoma (39). Combined these results highlight that the secretory differentiation in MCC is likely attributable to ATOH1 upregulation.

It is well-accepted that posttranslational modifications of ATOH1 participate in several homeostatic modulations and pathophysiological processes in brain and intestine (7,15–19). However, it remains a matter of debate which ATOH1 residues are exactly phosphorylated to control its protein stability (7,15,16,18,19,23). Previous studies indicated that phospho-dead mutants of ATOH1, in which serines 54/58 (15,38) or combined with serines 110/114/118 (23) were replaced by alanine, are resistant to proteasomal degradation in CRC cells. Mouse ATOH1 serines 328/339 and serine 334 (corresponding to human ATOH1 serines 331/342 and serine 337, respectively), located at a highly evolutionary conserved region, have also been reported as functional phosphorylation sites marking ATOH1 for degradation (16,17). In the present study, we confirmed that phosphorylation sites S331/342 are absolutely essential for ATOH1 proteasomal degradation, while the S54/58 and S110/114/118 clusters are irrelevant (**Fig. 2B**). Interestingly, we observed that ATOH1 phosphorylated at Y80 (identical to Y78 in mouse ATOH1) exerts an opposing effect



that also in colorectal cancer cells slightly increases its stability, in agreement with a recent medulloblastoma study (18) (**Fig. 6E–G**). Besides Y80, other amino residues of ATOH1, as seems likely, are phosphorylated and thereby prevent ATOH1 from degradation, as shown by the reduced levels of Y80F (phospho-dead) ATOH1 following ID-8 treatment compared with DMSO-treated controls (**Fig. 8A–D**). In addition, Tomic et al. argue that intestinal differentiation is regulated by ATOH1 simultaneously phosphorylated at multiple sites (7). Accelerated secretory maturation and elevated expression of the mucin gene signature was observed in mouse intestinal cells *in vivo* expressing a phospho-dead ATOH1 variant mutated at nine serine/threonine (S/T) residues (7). Nevertheless, whether this artificial mutant modified at nine residues is still the real ATOH1 protein, should be taken into account. Future work on this observation is needed. Taken together, there are multiple potential phosphorylation sites in ATOH1, and phosphorylations of ATOH1 at different residues may exert diverse functions that regulate its stability involved in secretory differentiation.

APC appears to assure secretory lineage differentiation in the intestine (40). APC restoration increased expression of MUC2 and lysozyme, respectively markers for mucin-expressing goblet and Paneth cells, in the intestinal crypt of APC-depleted mice (40). Similarly, restoring APC expression in the CRC line HT29 carrying a mutated APC showed elevated MUC2 expression as compared with its parental cells (10). These previous findings combined with our APC functional analysis (**Fig. 4A, B**) imply a potential correlation between APC function and expression of ATOH1 and its downstream mucin genes.

It has been proposed that  $\beta$ -catenin and ATOH1 are reciprocally targeted for degradation, i.e. when Wnt/ $\beta$ -catenin signaling is active, like in APC-mutant cells, GSK3 switches to phosphorylating ATOH1 resulting in active degradation of the latter (15). In agreement with this, ATOH1 stability in APC-wildtype CRC cells (HCT116 and LS174T) was reduced following APC knockout (**Fig. 4C, E, F**). However, knockout of the deficient APC protein (p.I1417fs\*2) in DLD1 cells still exhibited the ability to compromise ATOH1 stability (**Fig. 4D, G**). These results raise doubts about the

previous proposed scenario that ATOH1 and  $\beta$ -catenin are reciprocally phosphorylated by GSK3 $\beta$  (15). In further disagreement, we did not observe any effects of GSK3 $\beta$  on ATOH1 stability, confirmed by employing two GSK3 $\beta$  inhibitors (**Fig. 3A, B, D**). A second study argued that activation of  $\beta$ -catenin signaling may posttranscriptionally downregulate ATOH1 expression (13). However, our data showed that  $\beta$ -catenin failed to affect ATOH1 protein stability (**Fig. 3D; Supplementary Fig. 7**). Overall, our data demonstrate that APC-mediated ATOH1 stabilization does not depend on either GSK3 $\beta$  or  $\beta$ -catenin. At present, we do not understand how APC may regulate ATOH1's stability, but it does not appear to involve the JAK2-mediated phosphorylation of Y80, or to involve the UBR5/DYRK complex that we identified as a player in ATOH1 regulation.

The E3 ubiquitin ligase HUWE1 was identified to target phospho-marked ATOH1 for ubiquitination and degradation in two previous studies (16,17). Furthermore, the SHH-pathway inhibitor vismodegib (GDC-0449) was reported to induce ATOH1 phosphorylation and subsequently increase Huwe1-dependent degradation of ATOH1 (16). Theoretically, these findings corroborated by comprehensive biomedical experiments and *in vivo* mouse studies of brain and sensory cells, may be extrapolated to the human intestine. However, SHH inhibition showed an opposing effect on ATOH1 stability in human CRC cells, meaning that it increased ATOH1 levels (**Supplementary Fig. 4**). In this regard, there are probably alternative co-regulators that play more pronounced roles than the SHH signals in controlling ATOH1 proteasomal breakdown in gut, most likely owing to different driving forces for tumor formation in brain and intestine such as SHH and  $\beta$ -catenin signaling, respectively.

Importantly, we identified one other E3 ubiquitin ligase UBR5 as a binding partner of ATOH1, which seemingly serves as an antagonist or competitor of HUWE1. In this study, we report that UBR5 protects ATOH1 from breakdown through its ubiquitin ligase activity in a subset of CRC cells. However, *UBR5* mRNA and protein levels did not differ between non-mucinous and mucinous colorectal tumors (**Supplementary Fig. 12A–C**). Hence, only the presence of the UBR5-mediated ATOH1 stabilization could not be directly integrated into the paradigm for CRC with

a mucinous phenotype. On the other hand, we did not observe any noticeable increase in APC protein level upon UBR5 overexpression (**Fig. 6A**), in contrast to previous result that UBR5 interacts with and stabilizes APC (41). Interestingly, UBR5 has been demonstrated to actively participate in tumor development (42). UBR5 ubiquitylates and inactivates the TLE repressor of  $\beta$ -catenin-dependent transcription, functions as a co-effector with activated  $\beta$ -catenin, thereby sustaining  $\beta$ -catenin transactivation (43). Together, it is likely that the divergent functions of UBR5 that act in sustaining  $\beta$ -catenin transactivation and maintaining ATOH1 stability may contribute to tumor formation and simultaneously allow mucin-producing differentiation within the same tumor mass in MCC.

Our results also propose that sufficient DYRK2 and DYRK4 could stabilize ATOH1, thus favoring mucinous differentiation within colon tumors. DYRK2 has been identified in a complex with UBR5 (35). Although UBR5 RNA levels do not differ between MCC and NMC tumors, DYRK2, and especially DYRK4, are expressed higher in mucinous tumors. As such both DYRK proteins may cooperate with UBR5 to stabilize ATOH1, although it has not yet been shown that DYRK4 also associates with UBR5. In support of such a cooperation is that both UBR5 knockout and DYRK inhibition, result in reduced ATOH1 stability. The compound that we used to inhibit DYRK proteins, i.e. ID-8, has been shown to potently bind to DYRK2 and DYRK4 (36). DYRK inhibition by using ID-8 abrogated differentiation in Wnt-stimulated expansion of human stem cells (36). The underlying mechanism for such effect of the DYRK inhibitor remains poorly understood. The association of DYRK2 and DYRK4 with ATOH1 revealed in the present study provides however a possible reason explaining how DYRK inhibition may be involved in maintenance of the undifferentiated state.

MCC is arbitrarily demarcated as tumor mass comprising at least 50% mucous component, hence its diagnosis fully relies on pathologists' determination. In this study, we propose a potential biomarker signature, that is *DYRK4* and *MUC2* in combination, for molecularly distinguishing MCC. It has been shown that patients with goblet-like-subtype tumor classified by high *MUC2* expression are likely to

achieve a more favorable outcome when compared with whom had CRC clustered with other molecular signatures (44). Tumors with high expression levels of both *DYRK4* and *MUC2* might be capable of sharing major characteristics with MCCs, exhibiting a well-differentiated phenotype (3). Therefore, such ATOH1-related biomarker signature may assist in predicting clinical outcomes or decision-making concerning treatment strategy of the CRC patients with this molecular subtype, which probably have more strength than the conventional histology-based determination for MCC.

In conclusion, we have evaluated previously reported mechanisms and identified new players to stabilize ATOH1, one of the main actors to regulate mucinous differentiation in the gastrointestinal tract. We show that the APC protein contributes to ATOH1 stabilization, but in contrast to previous reports, the mechanism does not involve the GSK3 kinase or  $\beta$ -catenin signaling. SHH inhibition led to increased ATOH1 levels, opposing the effects seen in medulloblastoma. The previously reported JAK2-mediated phosphorylation of Y80 also moderately increased ATOH1 levels in CRC. Finally, we identified UBR5 to protect ATOH1 from degradation through its ubiquitin ligase activity in a subset of CRC cells. UBR5 may do this in a complex or independently from DYRK proteins, for which we observed that they are also involved in ATOH1 stabilization. Attempts to link all these mechanisms failed or were inconclusive, likely showing that ATOH1 protein regulation is a multifactorial process.

## Materials and Methods

### Cell culture

HCT116, DLD1, SW480, Caco2, GP5D, SW48, LOVO, HEK293T, and L-cells were cultured in DMEM culture medium (Lonza, Breda, The Netherlands) containing 10% FBS (Gibco, Bleiswijk, The Netherlands). LS174T cells were grown in RPMI-1640 medium (Lonza) containing 5% FBS. CW-2 cells were maintained in RPMI-1640 culture medium supplemented with 10% FBS. SNU175 cells were grown in RPMI-1640 medium containing 10% FBS and 1 mM sodium pyruvate (Gibco). All the cell lines were grown in a humidified incubator at 37 °C with 5% CO<sub>2</sub>. Short tandem repeats (STR) genotyping of all cell lines/subclones was performed by Powerplex-16 STR genotyping (Promega, Leiden, The Netherlands) in the Erasmus Molecular Diagnostics Department in October 2018. All the cells were regularly tested for mycoplasma and confirmed to be negative at Eurofins GATC-Biotech (Konstanz, Germany). APC mutations in the cell lines used in the present study are depicted in **Supplementary Table 1**.

### Treatments

Wnt3a conditioned medium (Wnt3a-CM) or control conditioned medium (Control-CM) was used at 25% (v/v) in culture medium for indicated time periods. Wnt3a-CM and Control-CM were prepared from 0.22-micron filtered supernatant of cultured mouse fibroblast L-cells stably overexpressing Wnt3a (L-Wnt3a cells) and control L-cells (L-control cells), respectively, as described previously (45).

The protein synthesis inhibitor Cycloheximide (Sigma-Aldrich, St. Louis, MO, USA), dissolved in distilled water, was used at a final concentration of 100 µg/mL for indicated time periods. Other reagents were solubilized in DMSO unless indicated otherwise. The proteasome inhibitor MG132 (Sigma-Aldrich) was used at a final concentration of 10 µM for 6 h. The GSK3 inhibitor CHIR99021 (Selleckchem, Houston, TX, USA) was used at a final concentration of 3 µM for 24 h. The GSK3 inhibitor SB216763 (Sigma-Aldrich) was used at a final concentration of 5 µM for indicated time period. The Hedgehog pathway inhibitor Vismodegib (also called

GDC-0449, Selleckchem) was used at a final concentration of 10  $\mu$ M for 8 h. The JAK2 inhibitor AG-490 (Santa Cruz Biotechnology, Santa Cruz, CA, USA) was used at a final concentration of 5  $\mu$ M for 8 h. The pan-JAK inhibitor CP-690550 (Tofacitinib, Santa Cruz Biotechnology) was used at a final concentration of 0.5  $\mu$ M for 8 h. The DYRK inhibitor ID-8 (Sigma-Aldrich) was used at a final concentration of 10  $\mu$ M for 24 h.

### Plasmids used in the present study

The ORF of the wild-type ATOH1 from the plasmid pCMV-Flag-WT-Hath1 (kindly given by Dr. Mamoru Watanabe, Tokyo Medical and Dental University, Japan) was assembled into pcDNA-5'UT-FLAG vector (kindly given by Dr. Veronique Lefebvre, Lerner Research Institute, USA) using the Gibson Assembly Master Mix (New England Biolabs, Ipswich, MA, USA) according to the manufacturers' protocol. Additional sequences (unknown and not relevant to ATOH1) located between FLAG-tag and ATOH1-ORF in the original pCMV-Flag-WT-Hath1 plasmid were removed during this molecular assembling procedure. Based on this new wild-type pcDNA-ATOH1 vector, the SA1-SA7 (see **Fig. 2A**), Y80D, and Y80F ATOH1 variants were generated using the Q5 Site-directed Mutagenesis Kit (New England Biolabs). All ATOH1 expression vectors were full-length sequence verified by Sanger sequencing. FLAG-tagged expression vectors for  $\beta$ -catenin variants (wild-type and  $\Delta$ exon3) used in this work were constructed previously (32). Plasmids GFP-UBR5 (Addgene plasmid #52050) and GFP-UBR5- $\Delta$ HECT [ $\Delta$ HECT denotes a catalytically inactivating point mutation (C2768A) presented in the HECT domain of UBR5, leading to loss of the E3 ubiquitin ligase function. Addgene plasmid #52051] were gifts from Dr. Darren Saunders (University of New South Wales, Australia) (34). Plasmid pSpCas9(BB)-2A-GFP (PX458) was a gift from Dr. Feng Zhang (Addgene plasmid #48138) (46).

### Transfection

Expression vectors were transfected with FuGENE HD reagent (Promega) as described previously (45). Transient siRNA knockdown was performed with Lipofectamine RNAiMAX (Thermo Fisher Scientific) for 48 h according to the manufacturer's

instruction. The siRNAs obtained from Dharmacon (Lafayette, CO, USA) are listed as followed:

ON-TARGETplus SMARTpool human APC siRNA (cat.# L-003869-00-0005)

ON-TARGETplus SMARTpool human UBR5 siRNA (cat.# L-007189-00-0005)

ON-TARGETplus Non-targeting Pool (scramble control, cat.# D-001810-10-05)

### **Genome editing with CRISPR/Cas9 system**

CRISPR/Cas9 genome editing was performed as described previously (45,46). Briefly, single-guide RNAs (sgRNAs) were cloned into Cas9 nuclease vector PX458. Cells transfected with PX458 vector for 48 h were FACS-sorted for GFP-overexpressing cells, and single cells were then propagated into individual colonies in 96-well plates. Expanded cells were screened by Sanger sequencing and subsequent immunoblotting analysis (if antibody is available) to determine whether protein is knocked out. The sgRNA-targeting sequence and sequence alterations of ATOH1 in LS174T knockout (KO) clones are depicted in **Supplementary Fig. 2** (a reliable commercial antibody against ATOH1 for immunoblotting analysis is unavailable so far). The sgRNA sequence, sequence alterations and immunoblotting analysis of APC in HCT116, DLD1, and LS174T KO clones are depicted in **Supplementary Fig. 6**. The sgRNA sequence and sequence alterations of UBR5 in HCT116 and DLD1 KO clones are depicted in **Supplementary Fig. 11** (immunoblotting analysis of UBR5 is shown in **Fig. 5D, E**).

### **Co-immunoprecipitation**

Co-immunoprecipitation was carried out as described previously (45). In brief, 1 µg of WT- or SA4-FLAG-ATOH1 plasmids or empty vector (EV) control, were co-transfected with 1 µg of EV or GFP-UBR5 or GFP-UBR5-ΔHECT in HCT116 cells per well in 6-well plates using FuGENE HD. Cells were lysed at 48 h post-transfection in lysis buffer (30 mM Tris-HCl, 150 mM NaCl, 5 mM NaF, 5 mM EDTA, 1% TritonX-100, protease and phosphatase Inhibitor cocktail) for 20 min on ice. Lysates were cleared by centrifugation (13,000 rpm, 10 min at 4 °C), supernatants were then incubated 2 h at 4 °C with anti-FLAG M2 Affinity Gel (cat.# A2220, Sigma-Aldrich). Following lysis

buffer washing 4 times, immunoprecipitated proteins were eluted in 2× Laemmli sample buffer (95 °C, 10 min) and examined by immunoblotting.

### **Immunoblotting analysis**

Immunoblotting analysis was performed as described previously (45). Briefly, following cell lysis, proteins were separated by SDS-PAGE and transferred onto Immobilon-P PVDF membranes (Millipore, Bedford, MA, USA), and subsequently immunoprobed with the indicated antibodies. For detection of large proteins > 300 kDa (i.e., APC and UBR5), the proteins were transferred onto membrane at 70-80 mA constant current overnight at 4 °C. The primary antibodies used in this study were: anti-FLAG (1:1000, cat.# F1804, Sigma-Aldrich), anti-GFP (1:1000, cat.# A-11122, Thermo Fisher Scientific, Waltham, MA, USA), anti-GFP (1:500, cat.# 44011, Signalway, College Park, MD, USA), anti-APC (1:500, cat.# ab58, Abcam, Cambridge, MA, USA), anti-UBR5 (1:500, cat.# 65344, Cell Signaling Technology, Danvers, MA, USA), anti-β-catenin (1:1000, cat.# 8480, Cell Signaling Technology), anti-α-tubulin (1:5000, cat.# ab4074, Abcam), anti-β-actin (1:1000, cat.# sc-47778, Santa Cruz Biotechnology). The secondary antibodies used were IRDye 680RD goat anti-mouse (1:10000, cat.# 926-68070) and IRDye 800CW goat anti-rabbit (1:5000, cat.# 926-32211) secondary antibodies (Licor-Biosciences, Lincoln, NE, USA). Following incubation with secondary antibodies, the membrane was imaged with the Odyssey Infrared Imager system (Licor-Biosciences). Densitometric quantification of immunoblots was carried out by using software Image Studio Lite version 5.2 (Licor-Biosciences). Density for each band was normalized to the GFP (transfection control) and expressed as a ratio compared to indicated control sample.

### **Cycloheximide chase assay**

Cycloheximide chase assay was performed as described previously (45). Briefly, at 40 h post-transfection, cells were treated with cycloheximide (Sigma-Aldrich) at a final concentration of 100 µg/mL in culture medium for the indicated time periods. Half-lives of ATOH1 variants were determined via nonlinear regression analysis (One phase decay) by employing software GraphPad Prism version 5.01.



### **Mass spectrometry analysis**

The Doxycyclin (Dox)-inducible LS174T and DLD1 cell lines stably expressing FLAG-tagged ATOH1 and empty vector control were collected and lysed in lysis buffer for 20 min on ice, and centrifugation-cleared lysates were subsequently subjected to immunoprecipitation with anti-FLAG Affinity Agarose Gel (cat.# A2220, Sigma-Aldrich) at 4 °C on a rotator overnight. FLAG-immunoprecipitated proteins were analyzed by liquid chromatography (LC)-tandem mass spectrometry (MS/MS), then the resulting peptide analysis were processed with the UniProt human protein sequence database, and proteins were identified by using the Mascot software search engine program (Matrix Science, London, UK, [www.matrixscience.com](http://www.matrixscience.com)). A Venn diagram showing ATOH1 potential interactors was generated by employing online software DrawVenn ([bioinformatics.psb.ugent.be/webtools/Venn/](http://bioinformatics.psb.ugent.be/webtools/Venn/)).

### **Subcutaneous xenograft model, hematoxylin and eosin (H&E), and periodic acid–Schiff (PAS) staining**

Briefly, Dox-inducible DLD1 cells stably overexpressing ATOH1 and empty vector control DLD1 cells were harvested, and 200 µL of single cell suspensions (~10<sup>6</sup> cells/100 µL PBS) were then subcutaneously implanted into opposing flanks of recipient nude mice. Following 1 week of growth, mice were provided with 2 µg/mL of Doxycyclin in their drinking water until xenografts reached a size of approximately 1 cm in diameter. Xenograft tissues were fixed overnight at 4 °C with 4% paraformaldehyde (PFA) in PBS. Fixed tissues were embedded in paraffin and divided into 5-µm-thick sections. Sections were routinely stained with hematoxylin and eosin to show cellular morphology. For PAS staining, slides were incubated with 0.05% (v/v) periodic acid solution for 10 min, followed by incubation with Schiff's reagent for 15 min. Slides were scanned using a NanoZoomer 2.0HT digital slide scanner (Hamamatsu Photonic, Shizuoka, Japan). All the experimental procedures were approved by the Animal Ethics Committee of the Erasmus MC and performed in strict accordance with Dutch and international guidelines of Institutional Animal Care and Use Committee.

## RNA isolation and RNA-seq analysis

Prior to library preparation, total RNA of parental and ATOH1-knockout LS174T clones was extracted with the NucleoSpin RNA II kit (Macherey-Nagel) according to the manufacturer's instruction. Determination of RNA integrity number (RIN) was carried out for quality control by using Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA), and the RIN values of all the RNA samples in this study were greater than 9.9. The RNA-seq libraries were constructed and then sequenced using an Illumina HiSeq 2000 sequencing platform at Eurofins GATC-Biotech (Konstanz, Germany) as described previously (47).

## Bioinformatics analysis

The raw RNA-seq read counts and clinicopathological characteristics of patients within TCGA Colon adenocarcinoma (COAD) and Rectum adenocarcinoma (READ) cohorts were downloaded from the Genomic Data Commons (GDC) data portal ([portal.gdc.cancer.gov/](http://portal.gdc.cancer.gov/), accessed on 12 September 2019). The mRNA expression data (GSE2109 and GSE76307) from Gene Expression Omnibus (GEO, [www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)) was downloaded by using GEOquery R package (48). The normalizations for mRNA microarray and RNA-seq data were carried out following the instructions detailed in the limma (49) and Deseq2 (50) R packages, respectively. The normalized RNA-seq datasets from the Genotype Tissue Expression (GTEx) "Colon" cohort and the Cancer Cell Line Encyclopedia (CCLE) "large\_intestine" cohort were accessed via UCSC Xena platform ([xenabrowser.net/datapages/](http://xenabrowser.net/datapages/)). For the GSE2109 dataset, only the samples obtained from large intestine ("Colon" and "Rectum") were included and analyzed in this study. In all cohorts, duplicated samples and samples for which no clinical information was available, were excluded. Samples containing tumor histology other than "mucinous adenocarcinoma" or "adenocarcinoma", e.g. "signet-ring cell carcinoma", were excluded.

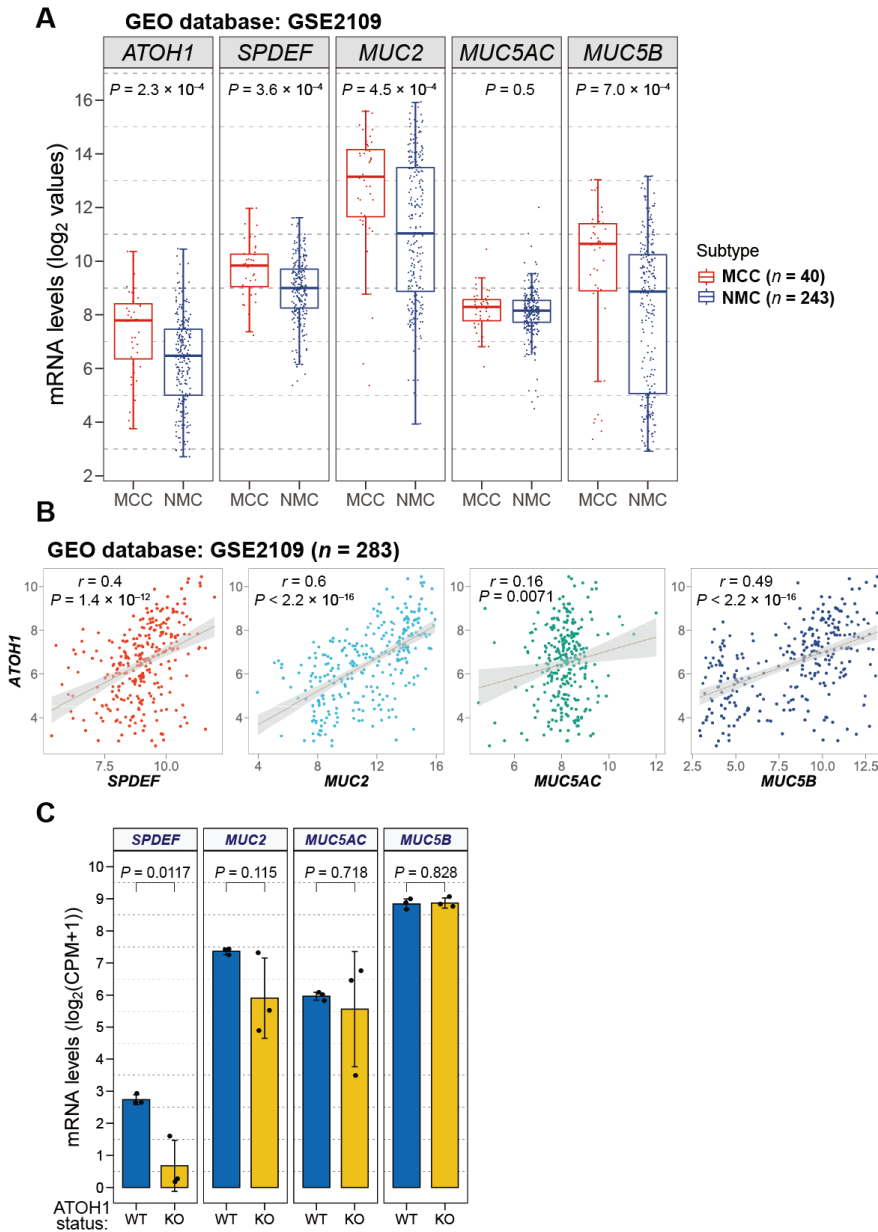
Differential gene expression analysis was performed by using the Deseq2 R package (50). Genes that are differentially abundant in mucinous versus non-mucinous TCGA-COAD&READ samples were considered as those which had an

absolute  $\log_2$ -transformed fold change  $> 1$  and false discovery rate (FDR)-adjusted  $P$  value  $< 0.05$  when comparing mRNA expression levels.

Somatic mutations presented in the COAD and READ cases in the TCGA dataset, which were identified by VarScan2 variant caller (51) and compiled in mutation annotation format (MAF) files, and corresponding clinical information were downloaded from the GDC data portal (accessed on 5 February 2020). Annotation and visualization of somatic mutations were performed using the Maftools R package (52). Samples were excluded from the mutational analysis if mutation data, expression value, and clinical information were not provided simultaneously, and duplicated and silent mutations were filtered.

### Statistical analysis

Continuous variables were expressed as mean  $\pm$  standard deviation (SD) unless otherwise noted. Two-tailed Student's  $t$ -test and non-parametric tests (Mann-Whitney  $U$  test or Wilcoxon rank sum test, as indicated in figure legends) were used for comparisons between two groups of normally and non-normally distributed variables, respectively. One-way analysis of variance (ANOVA) with Tukey's *post hoc* test or Kruskal-Wallis  $H$  test with Dunn's *post hoc* test was carried out to compare normally or non-normally distributed data, respectively, when more than two groups were analyzed. Categorical variables were compared by Fisher's exact test. Correlation between mRNA expression levels of two genes was evaluated using Pearson correlation unless indicated otherwise. Receiver operating characteristic (ROC) curves were constructed using the pROC R package (53). The predicted probability of multiple gene combinations for diagnosis of mucinous colorectal cancer was calculated based on binomial logistic regression model.  $P$  value  $< 0.05$  was considered statistically significant (symbols denote  $*P < 0.05$ ,  $**P < 0.01$ , or  $***P < 0.001$ ). n.s., not significant ( $P$  value  $> 0.05$ ). All the statistical analyses and graphic representations were performed by using software GraphPad Prism version 5.01 (GraphPad Software Inc., San Diego, California, USA) and software R (version 3.6.1) with relevant statistical packages.



**Supplementary Fig. 1** *ATOH1* and downstream target genes thereof are positively associated with the mucinous phenotype in colorectal cancer (CRC). **(A)** Boxplots showing comparisons of mRNA levels of *ATOH1* and target genes thereof in mucinous versus non-mucinous CRC in the GSE2109 dataset. All depicted genes are positively associated with MCC, except *MUC5AC*. **(B)** Association of mRNA expression of *ATOH1* with its downstream target genes in colorectal tumor samples from the GSE2109 cohort.  $r$  denotes Pearson correlation coefficient. **(C)** The mRNA levels of *SPDEF* and *MUC2* were markedly decreased in three and two, respectively, independent *ATOH1* knockout (KO) clones of LSI74T compared with the parental cells. All data are expressed as mean  $\pm$  SD. Significance tested using Wilcoxon rank sum test **(B)**, and paired Student's  $t$ -test.

**sgRNAs used for generating the ATOH1-knockout clones**

sgRNA(#1): TTGCGGGAGATGATGCGGCT PAM: **GGG** (anti-sense)

sgRNA(#2): TCTCTCGCCTGCAAAGTTGC PAM: **AGG** (anti-sense)

**LS174T ATOH1-knockout clones:****ATOH1-KO(#1): Clone 2.1**

5' ...GCCAGC**CCG**AGCCGCATCATCTCCCGCAACCGCCGCCGCCGCCGAGCCACCTGCAACTTTGCAG  
GCGAGAGAGCATCC...3' (Parental)

5' ...GCCAGCCCCAGCCG-----AACCGCCGCCGCCGCCGAGCCACCTGCAACTTTGCAG  
GCGAGAGAGCATCC...3' (-13 bp) p.H26Nfs\*75

5' ...GCCAGCCCCA--CGCATCATCTCCCGCAACCGCCGCCGCCGCCGAGCCACCTGCAACTTTGCAG  
GCGAGAGAGCATCC...3' (-2 bp) p.Q24Hfs\*26

**ATOH1-KO(#2): Clone 2.2**

5' ...GCCAGC**CCG**AGCCGCATCATCTCCCGCAACCGCCGCCGCCGCCGAGCCACCTGCAACTTTGCAG  
GCGAGAGAGCATCC...3' (Parental)

5' ...GCCAGCCCCAGCC-CATCATCTCCCGCAACCGCCGCCGCCGCCGAGCCACCTGCAACTTTGCAG  
GCGAGAGAGCATCC...3' (-1 bp) p.H26Ifs\*79

5' ...GCCAGCCCCAGC-GCATCATCTCCCGCAACCGCCGCCGCCGCCGAGCCACCTGCAACTTTGCAG  
GCGAGAGAGCATCC...3' (-1 bp) p.P25Rfs\*80

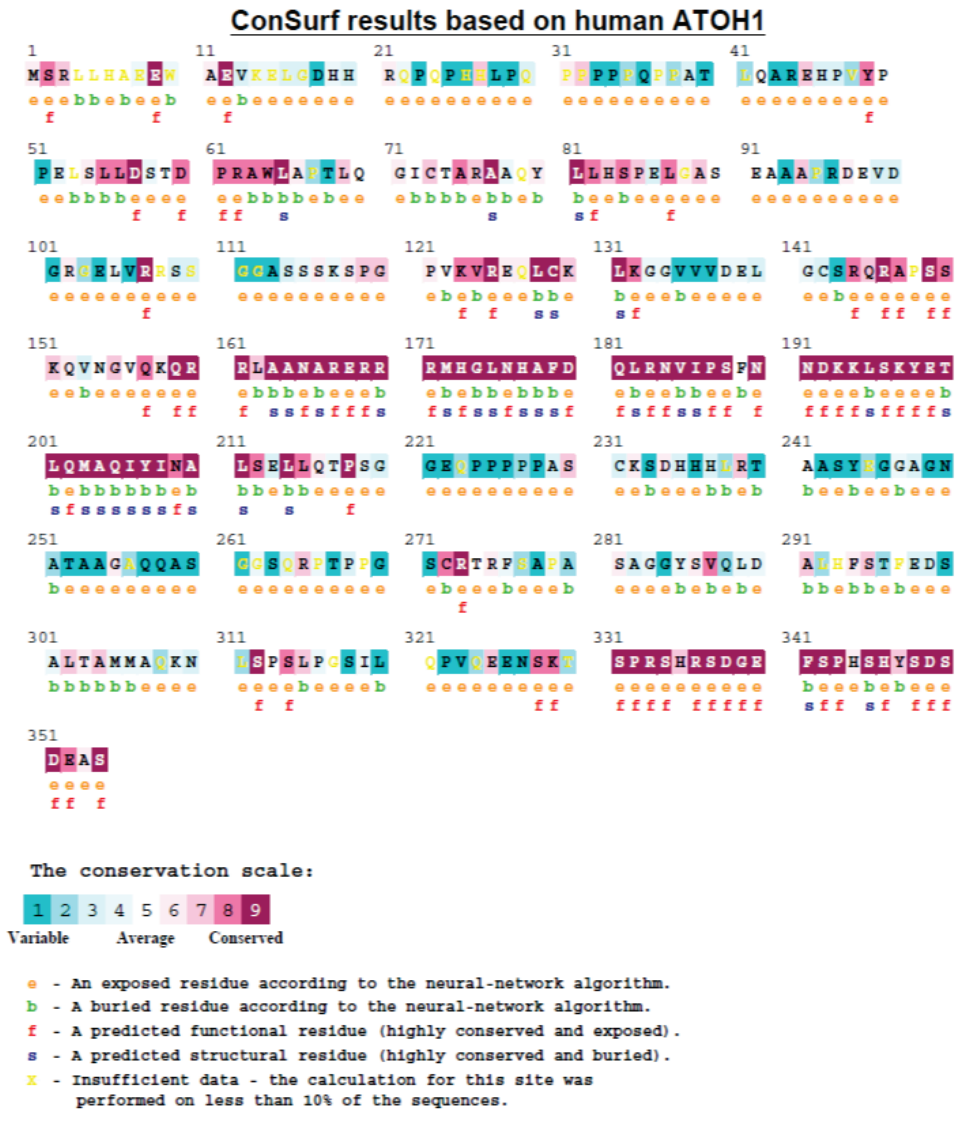
**ATOH1-KO(#3): Clone 3.1**

5' ...GCCAGCCCCAGCCGCATCATCTCCCGCAACCGCCGCCGCCGCCGAGCC**ACCT**GCAACTTTGCAG  
GCGAGAGAGCATCC...3' (Parental)

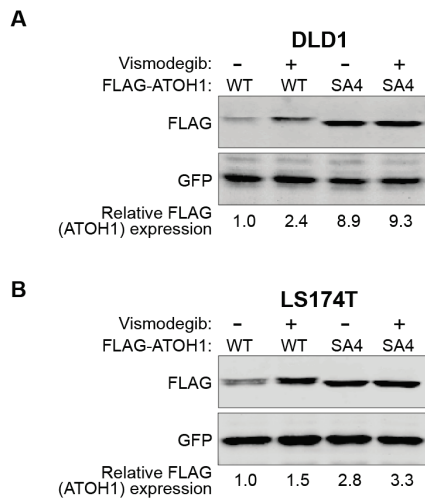
5' ...GCCAGCCCCAGCCGCATCATCTCCCGCAACCGCCGCCGCCGCCGAGCCACCTGCA-----G  
GCGAGAGAGCATCC...3' (-8 bp) p.T40Gfs\*8

5' ...GCCAGCCCCAGCCGCATCATCTCCCGCAACCGCCGCCGCCGCCGCA-----ACTTTGCAG  
GCGAGAGAGCATCC...3' (-10 bp) p.P37Lfs\*65

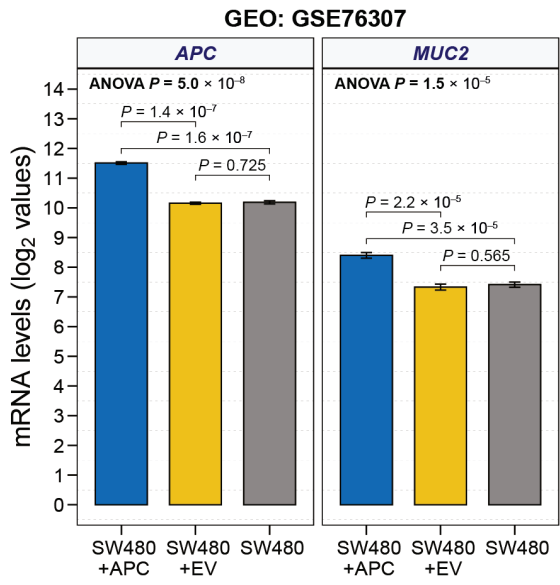
**Supplementary Fig. 2** Depiction of CRISPR/Cas9 induced mutations observed in the ATOH1 knockout clones of LS174T. The sgRNA-targeting sequences (underlined) and the protospacer-adjacent motif (PAM) sites (bold green) are shown as indicated.



**Supplementary Fig. 3** Evolutionary conservation analysis of ATOH1 using the ConSurf tool. Human ATOH1 was used as input with standard settings, and 70–99% homologous sequences of human ATOH1 were included to perform the ConSurf algorithm.



**Supplementary Fig. 4** Inhibition of the sonic hedgehog (SHH) pathway increases ATOH1 stability. (A, B) Immunoblotting showing protein levels of wild-type (WT) and S331/342A (SA4) ATOH1 variants in the absence or presence of the SHH inhibitor vismodegib in DLD1 (A) and LS174T (B) cells. Values below the immunoblotting images show quantification of relative FLAG-tagged ATOH1 protein level normalized to GFP value.



**Supplementary Fig. 5** Re-expression of APC upregulates MUC2 expression in SW480 colorectal cancer cells harboring a truncating APC mutation. GEO dataset (GSE76307) analysis showing that stable overexpression of wild-type APC increased MUC2 mRNA level in the APC-deficient colorectal cancer cell line SW480 compared with its parental cells or control cells stably expressing empty vector (EV). Significance tested using one-way ANOVA with Tukey's *post hoc* test.

## A sgRNAs used for generating the APC-knockout clones

sgRNA (targeting exon 4, mapped to the longest APC reference transcript ENST00000257430.9):  
CTGTTCTTATGGAAGCCGGGA PAM: **AGG** (sense strand)

### HCT116 APC-knockout clones:

#### APC-KO(#1): Clone 45

5' ...CTCCGTTCTTATGGAAGCCGGGA**AGG**ATCTGTATCAAGCCGTTCTGGAGA...3' (Parental)  
 5' ...CTCCGTTCTTATGGAAGC-----GGATCTGTATCAAGCCGTTCTGGAGA...3' (+1 bp, -8 bp) p.S98Tfs\*23  
 5' ...CTCCGTTCTTATGGAAGCCGGG**GA**AGGATCTGTATCAAGCCGTTCTGGAGA...3' (+1 bp) p.E100Gfs\*37

#### APC-KO(#2): Clone 48

5' ...CTCCGTTCTTATGGAAGCCGGGA**AGG**ATCTGTATCAAGCCGTTCTGGAGA...3' (Parental)  
 5' ...CTCCGTTCTTATGGAAG**A**-----ATCTGTATCAAGCCGTTCTGGAGA...3' (+1 bp, -9 bp) p.S98Rfs\*36  
 5' ...CTCCGTTCTTATGGAAG**GATC**-----  
 -----TTGTAATGGAAGCAGAGAAAGTACTGGA...3' (-4bp, -74 bp) p.S98Rfs\*2

#### APC-KO(#3): Clone 50

5' ...CTCCGTTCTTATGGAAGCCGGGA**AGG**ATCTGTATCAAGCCGTTCTGGAGA...3' (Parental)  
 5' ...CTCCGTTCTTATGGAAG-----GATCTGTATCAAGCCGTTCTGGAGA...3' (-8 bp) p.S98Rfs\*36  
 5' ...CTCCGTTCTTATGGAAGCCG-----GATCTGTATCAAGCCGTTCTGGAGA...3' (-5 bp) p.E100Ifs\*35

### DLD1 APC-knockout clones:

#### APC-KO(#1): Clone 3 (tetraploid)

5' ...CTCCGTTCTTATGGAAGCCGGGA**AGG**ATCTGTATCAAGCCGTTCTGGAGA...3' (Parental)  
 5' ...CTCCGTTCTTATGGAAGCC-----GAAGGATCTGTATCAAGCCGTTCTGGAGA...3' (-2 bp) p.E100Rfs\*36  
 5' ...CTCCGTTCTTATGGAAGC-----GATCTGTATCAAGCCGTTCTGGAGA...3' (-8 bp) p.S98Rfs\*36  
 5' ...CTCCGTTCTTATGGAAGCCG**GG**AAGGATCTGTATCAAGCCGTTCTGGAGA...3' (+1 bp) p.E100Gfs\*37  
 5' ...CTCCGTTCTTATGGAAGCCG**T**GGAAGGATCTGTATCAAGCCGTTCTGGAGA...3' (+1 bp) p.E100Gfs\*37

#### APC-KO(#2): Clone 13

5' ...CTCCGTTCTTATGGAAGCCGGGA**AGG**ATCTGTATCAAGCCGTTCTGGAGA...3' (Parental)  
 5' ...CTCCGTTCTTATGGAAG-----GAAGGATCTGTATCAAGCCGTTCTGGAGA...3' (-4 bp) p.S98Rfs\*24  
 5' ...CTCCGTTCTTATGGAAG-----GTATCAAGCCGTTCTGGAGA...3' (-19 bp) p.Y96\*

#### APC-KO(#3): Clone 18

5' ...CTCCGTTCTTATGGAAGCCGGGA**AGG**ATCTGTATCAAGCCGTTCTGGAGA...3' (Parental)  
 5' ...CTCCG-----GAAGGATCTGTATCAAGCCGTTCTGGAGA...3' (-16 bp) p.S98Kfs\*23  
 5' ...CTCCGTTCTTATGGAAG-----AGCCGGGAAGGATCTGTATCAAGCCGTTCTGGAGA...3' (-4 bp) p.Y96\*

### LS174T APC-knockout clones:

#### APC-KO(#1): Clone 3

5' ...CTCCGTTCTTATGGAAGCCGGGA**AGG**ATCTGTATCAAGCCGTTCTGGAGA...3' (Parental)  
 5' ...CTCCGTTCTTATGGAAGC-----GAAGGATCTGTATCAAGCCGTTCTGGAGA...3' (-2 bp) p.R99Gfs\*37  
 5' ...CTCCGTTCTTATGGAAGCC-----GAAGGATCTGTATCAAGCCGTTCTGGAGA...3' (-1 bp) p.E100Kfs\*23

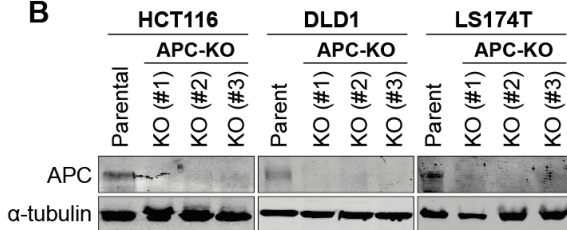
#### APC-KO(#2): Clone 8

5' ...CTCCGTTCTTATGGAAGCCGGGA**AGG**ATCTGTATCAAGCCGTTCTGGAGA...3' (Parental)  
 5' ...CTCCGTTCTTATGGAAG-----GAAGGATCTGTATCAAGCCGTTCTGGAGA...3' (-4 bp) p.S98Rfs\*24  
 5' ...CTCCGTTCTTATGGAAGCCGGGA**TATGAAATCTTGGTTGGAATTTCTTTCTTTAAGAATACTGAAATAGGCCTC**  
**CAATCTCTCTGGCTTGTAAGGTT**...3' (large insertion) p.E100Dfs\*10

#### APC-KO(#3): Clone 9

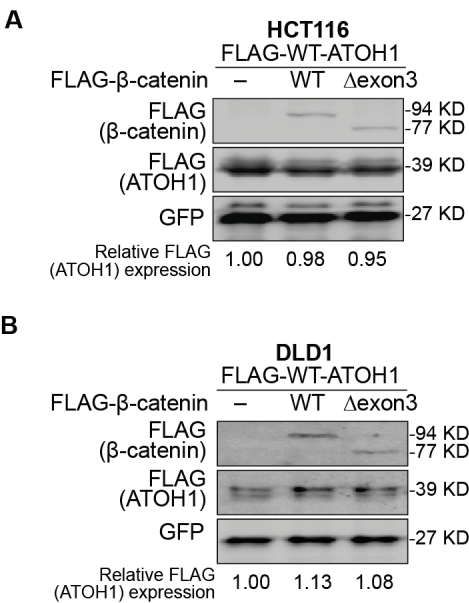
5' ...CTCCGTTCTTATGGAAGCCGGGA**AGG**ATCTGTATCAAGCCGTTCTGGAGA...3' (Parental)  
 5' ...CTCCGTTCTTATGAC**GGAGGACATTGCTACATT**GATCTGTATCAAGCCGTTCTGGAGA...3' (-14 bp, +21bp) p.Y96\*  
 5' ...CTCCGTTCTTATGGAAG-----  
 -----TATGGGTTG**CATT**  
**TCCAAGAAGAGGGTTTGTAATGGAAGCAGAGAAAGTACTGGATATTTAGAAGAACTGAGAAAGAGGATGTAACCTT**...3'  
 (-48 bp, large-segment replacement) p.R99Mfs\*26

## B

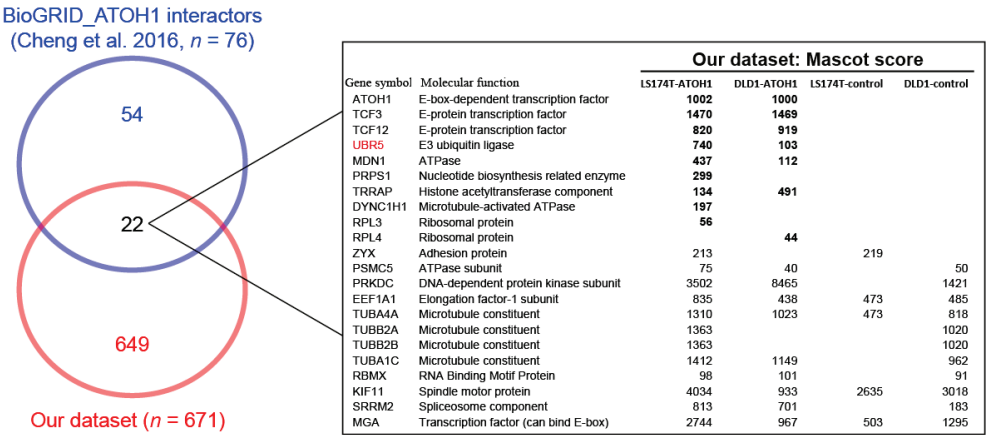


**Supplementary Fig. 6 (A)** Depiction of CRISPR/Cas9 induced mutations observed in the APC knockout clones of HCT116, DLD1, and LS174T. The sgRNA-targeting sequences (underlined) and the protospacer-adjacent motif (PAM) sites (bold green) are shown as indicated. **(B)** Immunoblotting analysis of parental and APC-knockout HCT116, DLD1, and LS174T cells.

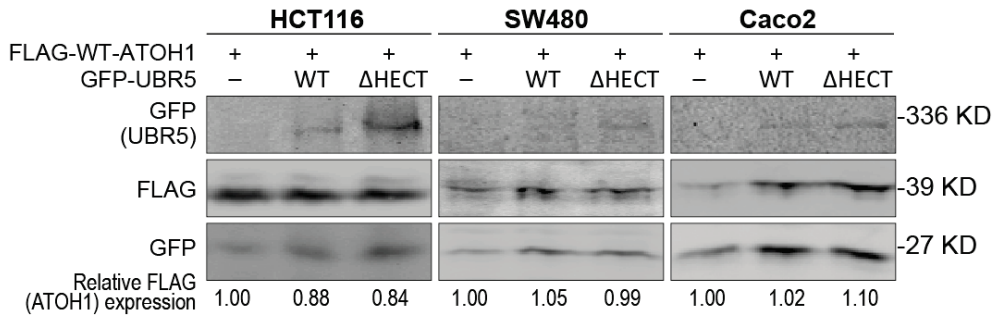




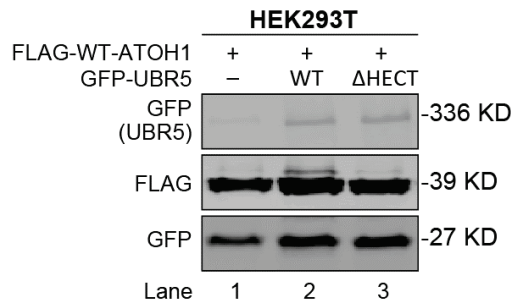
**Supplementary Fig. 7** β-catenin unlikely affects ATOH1 protein stability. (A, B) Immunoblotting showing protein levels of FLAG-tagged wild-type (WT) ATOH1 in HCT116 (A) and DLD1 (B) cells co-transfected with FLAG-tagged vectors expressing WT and degradation-resistant exon3-deletion (Δexon3) variants of β-catenin. Values below the immunoblotting images show quantification of relative FLAG-tagged ATOH1 protein level normalized to GFP value.



**Supplementary Fig. 8** Left, Venn diagram showing the number of potential ATOH1 interactors deposited in BioGRID database (blue) and ATOH1-binding candidates that we identified using mass spectrometry analysis in FLAG-immunoprecipitated samples of LS174T and DLD1 cells stably expressing FLAG-tagged ATOH1 (red). Right, table showing the overlapping ATOH1-interacting candidates and corresponding molecular functions and Mascot scores in the LS174T and DLD1 cells stably expressing ATOH1 or control vector. The gene symbol “UBR5” is highlighted with red color.



**Supplementary Fig. 9** Immunoblotting showing that WT-UBR5 overexpression increased ATOH1 protein levels in the indicated colorectal cancer (CRC) cells.



**Supplementary Fig. 10** The phosphorylated form of ATOH1 could be stabilized upon UBR5 overexpression through its ubiquitin ligase activity. Immunoblotting showing protein levels of wild-type (WT) ATOH1 in HEK293T cells co-transfected with empty vector or GFP-tagged UBR5 vectors expressing wild-type (WT) or catalytically inactive ( $\Delta$ HECT) variants. A GFP expression vector was used as transfection control.

**sgRNAs used for generating the UBR5-knockout clones**

sgRNA (targeting exon 3, mapped to the longest UBR5 reference transcript ENST00000520539.6):  
GGCTACTATTAACAGTGTG PAM: **TGG** (sense strand)

**HCT116 UBR5-knockout clones:****UBR5-KO(#1): Clone 2**

5' ... GAACAGGCTACTATTAACAGTGTG**TGG**TGGGACCAATCATGCTGCCTTTCT...3' (Parental)  
 5' ... GAACAGGCTACTATTAACAGT**T**GTGTGGTGGGACCAATCATGCTGCCTTTCT...3' (homo., +1 bp) p.C51Lfs\*11

**UBR5-KO(#2): Clone 6**

5' ... GAACAGGCTACTATTAACAGTGTG**TGG**TGGGACCAATCATGCTGCCTTTCT...3' (Parental)  
 5' ... GAACAGGCTACTATTAACAGT**T**GTGTGGTGGGACCAATCATGCTGCCTTTCT...3' (+1 bp) p.C51Lfs\*11  
 5' ... GAACAGGCTACTATTAACAGT**T**GTGTGGTGGGACCAATCATGCTGCCTTTCT...3' (-1 bp) p.C51Lfs\*40

**UBR5-KO(#3): Clone 10**

5' ... GAACAGGCTACTATTAACAGTGTG**TGG**TGGGACCAATCATGCTGCCTTTCT...3' (Parental)  
 5' ... GAACAGGCTACTATTAACAGT**T**GTGTGGTGGGACCAATCATGCTGCCTTTCT...3' (homo., +1 bp) p.C51Lfs\*11

**DLD1 UBR5-knockout clones:****UBR5-KO(#1): Clone 2 (tetraploid)**

5' ... GAACAGGCTACTATTAACAGTGTG**TGG**TGGGACCAATCATGCTGCCTTTCT...3' (Parental)  
 5' ... GAACAGGCTACTATTAACAGT**T**GTGTGGTGGGACCAATCATGCTGCCTTTCT...3' (+1 bp) p.C51Lfs\*11  
 5' ... GAACAGGCTACTATTAACAGT**-**GTGTGGGACCAATCATGCTGCCTTTCT...3' (-2 bp) p.V52Gfs\*9

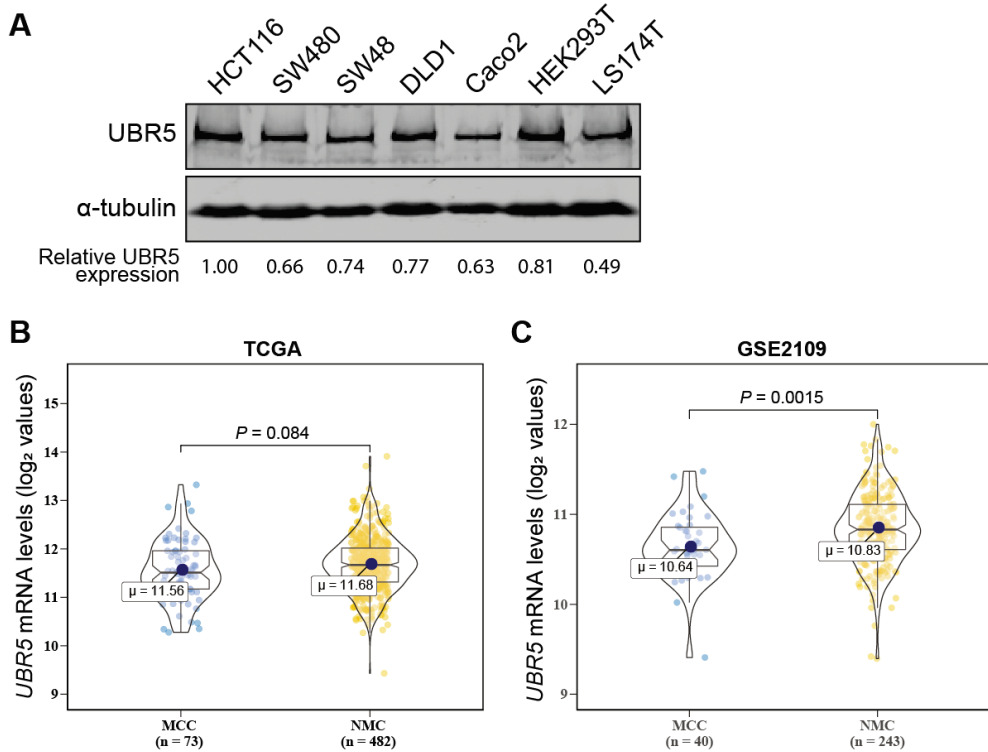
**UBR5-KO(#2): Clone 6**

5' ... GAACAGGCTACTATTAACAGTGTG**TGG**TGGGACCAATCATGCTGCCTTTCT...3' (Parental)  
 5' ... GAACAGGCTACTATTAACAGT**T**GTGTGGTGGGACCAATCATGCTGCCTTTCT...3' (+1 bp) p.C51Lfs\*11  
 5' ... GAACAGGCTACTATTAACAGT**T**GTGTGGTGGGACCAATCATGCTGCCTTTCT...3' (+2 bp) p.C51Ffs\*41

**UBR5-KO(#3): Clone 8**

5' ... GAACAGGCTACTATTAACAGTGTG**TGG**TGGGACCAATCATGCTGCCTTTCT...3' (Parental)  
 5' ... GAACAGGCTACTATTAACAGT**T**GTGTGGTGGGACCAATCATGCTGCCTTTCT...3' (homo., +1 bp) p.C51Lfs\*11

**Supplementary Fig. 11** Depiction of CRISPR/Cas9 induced mutations observed in the UBR5 knockout clones of HCT116 and DLD1. The sgRNA-targeting sequences (underlined) and the protospacer-adjacent motif (PAM) sites (bold green) are shown as indicated. homo., homozygous.



**Supplementary Fig. 12** UBR5 is comparably expressed in mucinous colorectal cancer (MCC) and non-mucinous colorectal cancer (NMC). (A) Immunoblotting showing protein levels of endogenous UBR5 in the indicated cell lines. Values below the immunoblotting images show quantification of relative UBR5 protein level normalized to the value of loading control  $\alpha$ -tubulin. (B, C) Expression levels of *UBR5* within tumor tissues of MCC and NMC from TCGA (B) and GSE2109 (C) cohorts are presented in violin plots. Significance tested in (B, C) using Wilcoxon rank sum test.

**Supplementary Table 1.** Cell lines used in this study

Cell line name	APC status
SW480	p.Q1338* homo.
Caco2	p.Q1367* het.
LOVO	p.R1114* + p.M1431fs*42
DLD1	p.I1417fs*2 het+p.R2166* het.
CW2	p.R302* het. + p.S1465fs*3 het.
GP5D	p.T1445fs*27 het.
SNU175	p.R232* het+p.N1455fs*18 het.+p.G1499* het.
LS174T	Wildtype
HCT116	Wildtype
SW48	Wildtype
HEK293T	Wildtype
L-cell	Wildtype

## References

1. F. Bray *et al.*, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**, 394-424 (2018).
2. J. Verhulst, L. Ferdinande, P. Demetter, W. Ceelen, Mucinous subtype as prognostic factor in colorectal cancer: a systematic review and meta-analysis. *J. Clin. Pathol.* **65**, 381-388 (2012).
3. N. Hugén, G. Brown, R. Glynne-Jones, J. H. de Wilt, I. D. Nagtegaal, Advances in the care of patients with mucinous colorectal cancer. *Nat. Rev. Clin. Oncol.* **13**, 361-369 (2016).
4. Q. Yang, N. A. Bermingham, M. J. Finegold, H. Y. Zoghbi, Requirement of Math1 for secretory cell lineage commitment in the mouse intestine. *Science* **294**, 2155-2158 (2001).
5. N. F. Shroyer *et al.*, Intestine-specific ablation of mouse atonal homolog 1 (Math1) reveals a role in cellular homeostasis. *Gastroenterology* **132**, 2478-2488 (2007).
6. J. H. van Es, N. de Geest, M. van de Born, H. Clevers, B. A. Hassan, Intestinal stem cells lacking the Math1 tumour suppressor are refractory to Notch inhibitors. *Nat. Commun.* **1**, 18 (2010).
7. G. Tomic *et al.*, Phospho-regulation of ATOH1 is required for plasticity of secretory progenitors and tissue regeneration. *Cell Stem Cell* **23**, 436-443 (2018).
8. Y. H. Lo *et al.*, Transcriptional regulation by ATOH1 and its target SPDEF in the intestine. *Cell. Mol. Gastroenterol. Hepatol.* **3**, 51-71 (2017).
9. W. Bossuyt *et al.*, Atonal homolog 1 is a tumor suppressor gene. *PLoS Biol.* **7**, e39 (2009).
10. C. C. Leow, M. S. Romero, S. Ross, P. Polakis, W. Q. Gao, Hath1, down-regulated in colon adenocarcinomas, inhibits proliferation and tumorigenesis of colon cancer cells. *Cancer Res* **64**, 6050-6057 (2004).
11. A. Velcich *et al.*, Colorectal cancer in mice genetically deficient in the mucin Muc2. *Science* **295**, 1726-1729 (2002).
12. K. Yang *et al.*, Interaction of Muc2 and Apc on Wnt signaling and in intestinal tumorigenesis: potential role of chronic inflammation. *Cancer Res* **68**, 7313-7322 (2008).
13. G. Peignon *et al.*, Complex interplay between  $\beta$ -catenin signalling and Notch effectors in intestinal tumorigenesis. *Gut* **60**, 166-176 (2011).
14. A. Kazanjian, T. Noah, D. Brown, J. Burkart, N. F. Shroyer, Atonal homolog 1 is required for growth and differentiation effects of notch/gamma-secretase inhibitors on normal and cancerous intestinal epithelial cells. *Gastroenterology* **139**, 918-928 (2010).
15. K. Tsuchiya, T. Nakamura, R. Okamoto, T. Kanai, M. Watanabe, Reciprocal targeting of Hath1 and  $\beta$ -catenin by Wnt glycogen synthase kinase 3b in human colon cancer. *Gastroenterology* **132**, 208-220 (2007).
16. A. Forget *et al.*, Shh signaling protects Atoh1 from degradation mediated by the E3 ubiquitin ligase Huwe1 in neural precursors. *Dev. Cell* **29**, 649-661 (2014).
17. Y. F. Cheng, M. Tong, A. S. Edge, Destabilization of Atoh1 by E3 ubiquitin ligase Huwe1 and casein kinase 1 is essential for normal sensory hair cell development. *J. Biol. Chem.* **291**, 21096-21109 (2016).
18. T. J. Klisch, A. Vainshtein, A. J. Patel, H. Y. Zoghbi, Jak2-mediated phosphorylation of Atoh1 is critical for medulloblastoma growth. *eLife* **6**, e31181 (2017).
19. W. R. Xie *et al.*, An Atoh1-S193A phospho-mutant allele causes hearing deficits and motor impairment. *J. Neurosci.* **37**, 8583-8594 (2017).
20. J. M. Carethers, B. H. Jung, Genetics and genetic biomarkers in sporadic colorectal cancer. *Gastroenterology* **149**, 1177-1190 (2015).
21. N. The Cancer Genome Atlas, Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330 (2012).
22. V. S. Li *et al.*, Wnt signaling through inhibition of  $\beta$ -catenin degradation in an intact Axin1 complex. *Cell* **149**, 1245-1256 (2012).

23. Y. Kano *et al.*, The acquisition of malignant potential in colon cancer is regulated by the stabilization of Atonal homolog 1 protein. *Biochem. Biophys. Res. Commun.* **432**, 175-181 (2013).
24. H. Ashkenazy *et al.*, ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344-350 (2016).
25. J. Mulvaney, A. Dabdoub, Atoh1, an essential transcription factor in neurogenesis and intestinal and inner ear development: function, regulation, and context dependency. *Journal of the Association for Research in Otolaryngology* **13**, 281-293 (2012).
26. L. E. King, C. G. Love, O. M. Sieber, M. C. Faux, A. W. Burgess, Differential RNA-seq analysis comparing APC-defective and APC-restored SW480 colorectal cancer cells. *Genomics Data* **7**, 293-296 (2016).
27. L. Zhang, J. W. Shay, Multiple roles of APC and its therapeutic implications in colorectal cancer *J. Natl. Cancer Inst.* **109**, (2017).
28. M. Christie *et al.*, Different APC genotypes in proximal and distal sporadic colorectal cancers suggest distinct WNT/ $\beta$ -catenin signalling thresholds for tumourigenesis. *Oncogene* **32**, 4675-4682 (2013).
29. R. Smits *et al.*, Apc1638T: a mouse model delineating critical domains of the adenomatous polyposis coli protein involved in tumorigenesis and development. *Genes Dev.* **13**, 1309-1321 (1999).
30. J. P. von Kries *et al.*, Hot spots in  $\beta$ -catenin for interactions with LEF-1, conductin and APC. *Nat. Struct. Biol.* **7**, 800-807 (2000).
31. F. Shi, Y. F. Cheng, X. L. Wang, A. S. Edge,  $\beta$ -catenin up-regulates Atoh1 expression in neural progenitor cells by interaction with an Atoh1 3' enhancer. *J. Biol. Chem.* **285**, 392-400 (2010).
32. P. Liu *et al.*, Oncogenic mutations in Armadillo repeats 5 and 6 of  $\beta$ -Catenin reduce binding to APC, increasing signaling and transcription of target genes. *Gastroenterology* **158**, 1029-1043 (2020).
33. R. Oughtred *et al.*, The BioGRID interaction database: 2019 update. *Nucleic Acids Res.* **47**, D529-D541 (2019).
34. T. Gudjonsson *et al.*, TRIP12 and UBR5 suppress spreading of chromatin ubiquitylation at damaged chromosomes. *Cell* **150**, 697-709 (2012).
35. S. Maddika, J. Chen, Protein kinase DYRK2 is a scaffold that facilitates assembly of an E3 ligase. *Nat. Cell Biol.* **11**, 409-419 (2009).
36. K. Hasegawa *et al.*, Wnt signaling orchestration with a small molecule DYRK inhibitor provides long-term xeno-free human pluripotent cell expansion. *Stem Cells Translational Medicine* **1**, 18-28 (2012).
37. K. Fukushima *et al.*, Atonal homolog 1 protein stabilized by tumor necrosis factor alpha induces high malignant potential in colon cancer cell line. *Cancer Sci* **106**, 1000-1007 (2015).
38. M. Aragaki *et al.*, Proteasomal degradation of Atoh1 by aberrant Wnt signaling maintains the undifferentiated state of colon cancer. *Biochem. Biophys. Res. Commun.* **368**, 923-929 (2008).
39. E. T. Park *et al.*, HATH1 expression in mucinous cancers of the colorectum and related lesions. *Clin. Cancer Res.* **12**, 5403-5410 (2006).
40. L. E. Dow *et al.*, Apc restoration promotes cellular differentiation and reestablishes crypt homeostasis in colorectal cancer. *Cell* **161**, 1539-1552 (2015).
41. R. Ohshima *et al.*, Putative tumor suppressor EDD interacts with and up-regulates APC. *Genes Cells* **12**, 1339-1345 (2007).
42. R. F. Shearer, M. Ionomou, C. K. Watts, D. N. Saunders, Functional roles of the E3 ubiquitin ligase UBR5 in cancer. *Mol. Cancer Res.* **13**, 1523-1532 (2015).
43. J. E. Flack, J. Mieszczanek, N. Novcic, M. Bienz, Wnt-dependent inactivation of the Groucho/TLE co-repressor by the HECT E3 ubiquitin ligase Hyd/UBR5. *Mol. Cell* **67**, 181-193 (2017).
44. A. Sadanandam *et al.*, A colorectal cancer classification system that associates cellular phenotype and responses to therapy. *Nat. Med.* **19**, 619-625 (2013).
45. S. Li *et al.*, Commonly observed *RNF43* mutations retain functionality in attenuating Wnt/ $\beta$ -catenin signaling and unlikely confer Wnt-dependency onto colorectal cancers. *Oncogene* **39**, 3458-3472 (2020).

46. F. A. Ran *et al.*, Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281-2308 (2013).
47. W. H. Wang *et al.*, Oncogenic STRAP supports hepatocellular carcinoma growth by enhancing Wnt/ $\beta$ -catenin signaling. *Mol. Cancer Res.* **17**, 521-531 (2019).
48. S. Davis, P. S. Meltzer, GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics* **23**, 1846-1847 (2007).
49. M. E. Ritchie *et al.*, limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
50. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol. Evol.* **15**, 550 (2014).
51. D. C. Koboldt *et al.*, VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568-576 (2012).
52. A. Mayakonda, D. C. Lin, Y. Assenov, C. Plass, H. P. Koeffler, Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* **28**, 1747-1756 (2018).
53. X. Robin *et al.*, pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**, 77 (2011).

# Chapter 6

## **Commonly Observed *RNF43* Mutations Retain Functionality in Attenuating Wnt/ $\beta$ -Catenin Signaling and Unlikely Confer Wnt-Dependency Onto Colorectal Cancers**

Shan Li<sup>1,2</sup>, Marla Lavrijsen<sup>1</sup>, Aron Bakker<sup>1</sup>, Marcin Magierowski<sup>1,3</sup>, Katarzyna Magierowska<sup>1,3</sup>, Pengyu Liu<sup>1</sup>, Wenhui Wang<sup>1,4</sup>, Maikel P. Peppelenbosch<sup>1</sup>, Ron Smits<sup>1</sup>

<sup>1</sup>Department of Gastroenterology and Hepatology, Erasmus MC-University Medical Center, Rotterdam, The Netherlands

<sup>2</sup>Department of Hepatobiliary Surgery, Daping Hospital, Army Medical University, Chongqing, China

<sup>3</sup>Department of Physiology, Jagiellonian University Medical College, Cracow, Poland

<sup>4</sup>State Key Laboratory of Natural Medicines, Department of Pharmacology, China Pharmaceutical University, Nanjing, China

***Oncogene.*** 2020, 39:3458–3472.





## Abstract

Cancer-associated *RNF43* mutations lead to activation of  $\beta$ -catenin signaling through aberrantly increasing Wnt-receptor levels at the membrane. Importantly, inactivating *RNF43* mutations have been suggested to render cancer cells sensitive to Wnt-based therapeutics. However, the extent to which *RNF43* mutations lead to impaired regulation of Wnt/ $\beta$ -catenin signaling has been poorly investigated. Here, we observed that tumors with a functional mismatch repair system show a predominant 5'-location of truncating *RNF43* mutations, suggesting C-terminal truncations such as the most commonly reported p.G659fs mutation, do not affect  $\beta$ -catenin signaling. In accordance, expressing C-terminal truncation mutants and wild-type RNF43, showed equal effects on  $\beta$ -catenin signaling, Wnt-receptor turnover and DVL-binding. We confirmed these observations at endogenous levels by CRISPR-Cas9-mediated knockout of G659fs RNF43 expression in KM12 cells and generating comparable mutations in HEK293T cells. We could not confirm previous reports linking RNF43 to p53 and E-cadherin breakdown. Our data also suggest that only colorectal cancer cells harboring N-terminal mutations of *RNF43* convey Wnt-dependency onto the tumor cells. Results of this study have potentially important clinical implications indicating that Wnt-based therapeutics should be applied cautiously in cancer patients harboring *RNF43* mutations.

**Keywords:** RNF43; Wnt-targeting agents; Wnt/ $\beta$ -catenin signaling; Gastrointestinal cancers.

## Introduction

The  $\beta$ -catenin signaling pathway is frequently deregulated in cancer. In colorectal cancer (CRC) this is mainly through loss of function mutations in the *APC* tumor suppressor gene, while several other tumor types acquire activating *CTNNB1* or inactivating *AXIN1/AXIN2* mutations. All these mutations lead to aberrant  $\beta$ -catenin stabilization, which constitutively activates Wnt/ $\beta$ -catenin target genes and triggers a genetic program resulting in tumor formation.

During the last decade, mutations in *ZNRF3* and *RNF43* are also recognized to drive tumorigenesis by aberrantly activating  $\beta$ -catenin signaling (1–3). They encode for transmembrane E3-ubiquitin ligases that target Wnt-receptors, composed of Frizzled (FZD) and LRP proteins, for endocytosis and degradation (1,2). As such, they represent a negative-feedback mechanism to keep Wnt/ $\beta$ -catenin signaling in check. In addition, RNF43 has been shown to also inhibit  $\beta$ -catenin signaling by tethering TCF7L2 (also known as TCF4) to the nuclear membrane (4).

*RNF43* mutations observed in cancers are mostly inactivating through protein truncation. They have been observed in a variety of cancers including those of the colon, stomach, pancreas, endometrium, and ovarium, at frequencies ranging from 4–21% (3,5–14). The functional consequence of these mutations is considered to be a reduced capability to downregulate Wnt-receptors, making these cancers potentially hypersensitive and dependent on Wnt-ligand activation. These *RNF43*-mutant cancers have, therefore, gained substantial therapeutic interest, as they may identify tumors that respond to extracellular Wnt-inhibitors that have been developed, such as FZD antibodies and Wnt-secretion inhibitors (15–17).

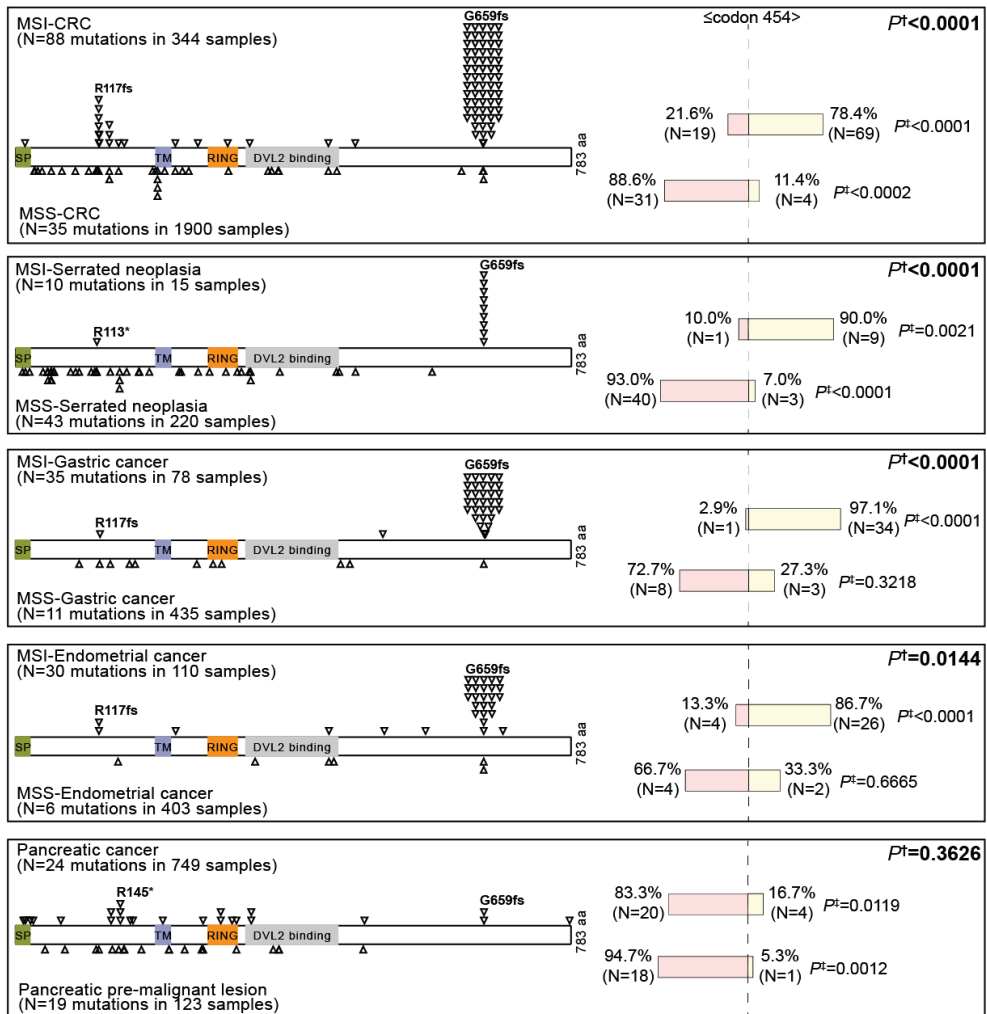
Importantly, the acquisition of several recurrent *RNF43* mutations is highly associated with mismatch repair (MMR)-deficiency, characterized by the accumulation of multiple mutations at the nucleotide level, especially within sequences containing stretches of mono- or di-nucleotide repeats. These tumors acquire thousands of mutations within genes, many of which are likely ‘bystander’ mutations, meaning that they will not strongly contribute to the tumorigenic process. *RNF43* mutation frequency in MMR-deficient colonic tumors is about 80% and more than 50% in those of the endometrium and stomach (5,9,18). More importantly, 50–

60% of these *RNF43* mutations are enriched in two hotspots at codons R117 and especially G659, respectively frameshift mutations occurring at 6×C or 7×G mononucleotide repeats (5). Both mutations arise more frequently than expected by chance. However, a thorough functional analysis to determine whether these mutations indeed confer a Wnt-dependency onto the tumor cells has not been performed. Here, we show that the most commonly reported p.G659fs *RNF43* mutation and several other C-terminal truncating mutations may in fact be 'bystander' mutations that have no or little impact on aberrantly regulating Wnt/ $\beta$ -catenin signaling.

## Results

### **Tumors with a functional mismatch repair system show a predominant 5'-location of truncating *RNF43* mutations**

Truncating mutations have been reported along the entire *RNF43* coding-sequence and are especially common in MMR-defective tumors (5). For various tumor types we plotted *RNF43* mutations separately for lesions with a defective (MSI) and functional (MSS) MMR-system (**Fig. 1; Supplementary Table 1**) (5,9,14,18–28). We stratified mutations to the first and second half of the coding region using codon 454 as cut-off, which demarcates the Dishevelled-2 (DVL2) binding domain, the most C-terminal binding protein currently identified in more detail (29). The location of *RNF43* mutations is significantly different between MSI and MSS cancers (**Fig. 1**). In parallel, comparison of observed and expected frequencies of truncating mutations showed that 3'-*RNF43* truncating mutations are strongly enriched in MSI-tumors, while 5'-*RNF43* truncating mutations are strongly associated with MSS-tumors. For patients with pancreatic cancers and premalignant lesions, which are typically MSS (30), *RNF43* truncating mutations are likewise almost exclusively identified in the 5'-half (**Fig. 1**) (31–35). Taken together, these results show that short *RNF43* mutants (shorter than 454-aa) are predominantly selected in tumors with a functional MMR-system and thus more likely represent the true 'driver' mutations in carcinogenesis, while the long *RNF43* mutants observed in MMR-deficient tumors may in fact represent 'bystander' mutations.



**Fig. 1** Distributions of *RNF43* truncating mutations in colorectal cancers, colorectal serrated neoplasia, gastric cancers, endometrial cancers, and pancreatic tumors. Only samples with one truncating *RNF43* mutation, as for most missense mutations the functional significance is unknown, and known MSI-status were included. MSI, microsatellite instability-high; MSS, microsatellite-stable; †, significant difference in location of *RNF43* truncating mutations before and after codon 454 between MSI and MSS groups or between pancreatic cancer and pre-malignant lesions was tested using a Fisher's exact test; ‡, *P* values correspond to Fisher's exact test comparing observed versus expected frequency. The vertical dashed line indicates the demarcation of *RNF43* mutation locations at codon 454.

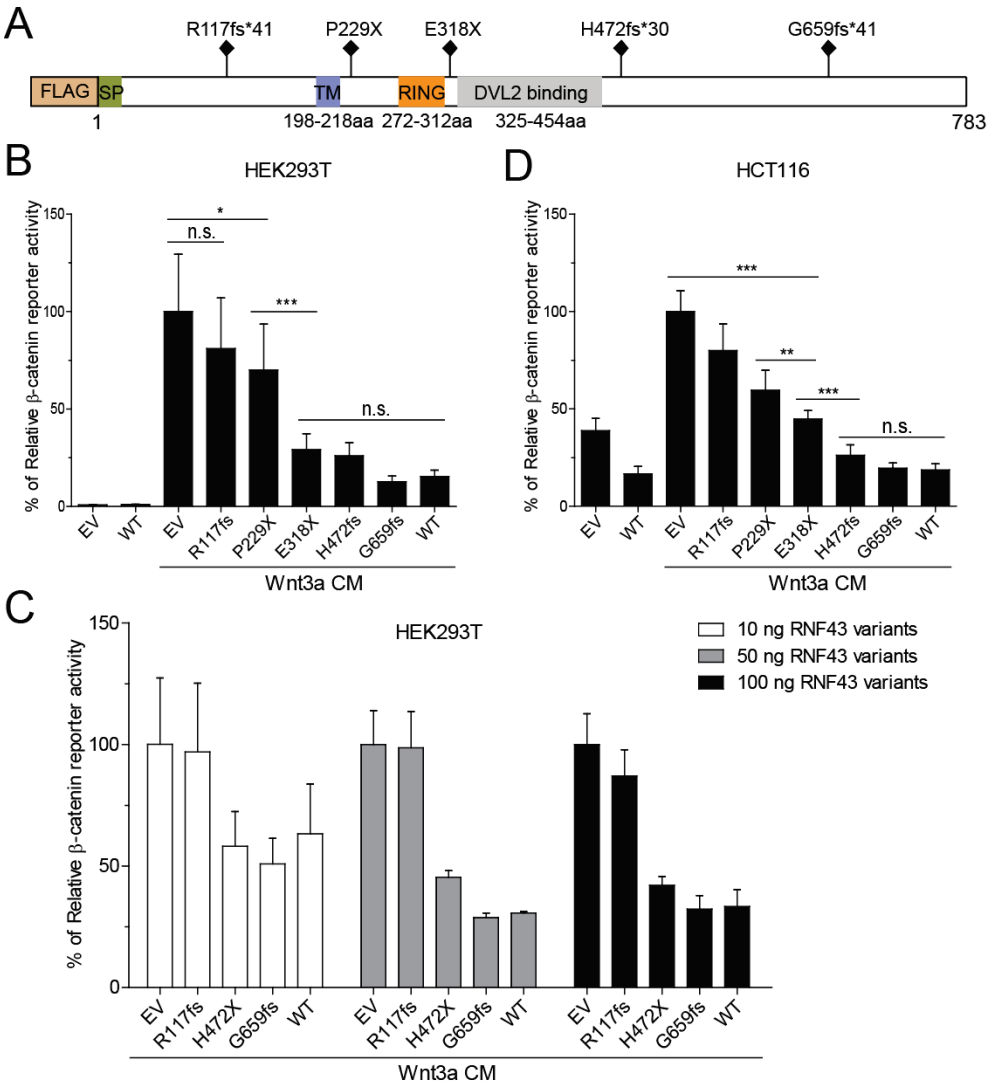
## C-terminal truncating RNF43 mutants retain negative regulation of $\beta$ -catenin signaling

To investigate to what extent RNF43 mutants retain functionality in regulating  $\beta$ -catenin signaling, we generated vectors expressing wild-type RNF43 and various truncated mutants thereof, which exactly recapitulate the most common *RNF43* truncating mutations observed in human cancers or delineate known functional domains (**Fig. 2a**) (14,29). As expected, wild-type RNF43 strongly inhibited Wnt-induced  $\beta$ -catenin signaling in HEK293T cells. Surprisingly, the G659fs variant demonstrated the same inhibitory effect, while the shorter H472fs and E318X variants also showed a significant reduction (**Fig. 2b**). A comparable effect on Wnt-induced  $\beta$ -catenin signaling was observed when only low amounts of RNF43 vectors were expressed, and in the Wnt-responsive CRC cell line HCT116 expressing only short RNF43 mutants (homozygous p.R117fs) (**Fig. 2c, d**). Overall, these results show that C-terminal truncating RNF43 mutants retain negative regulation of  $\beta$ -catenin signaling.

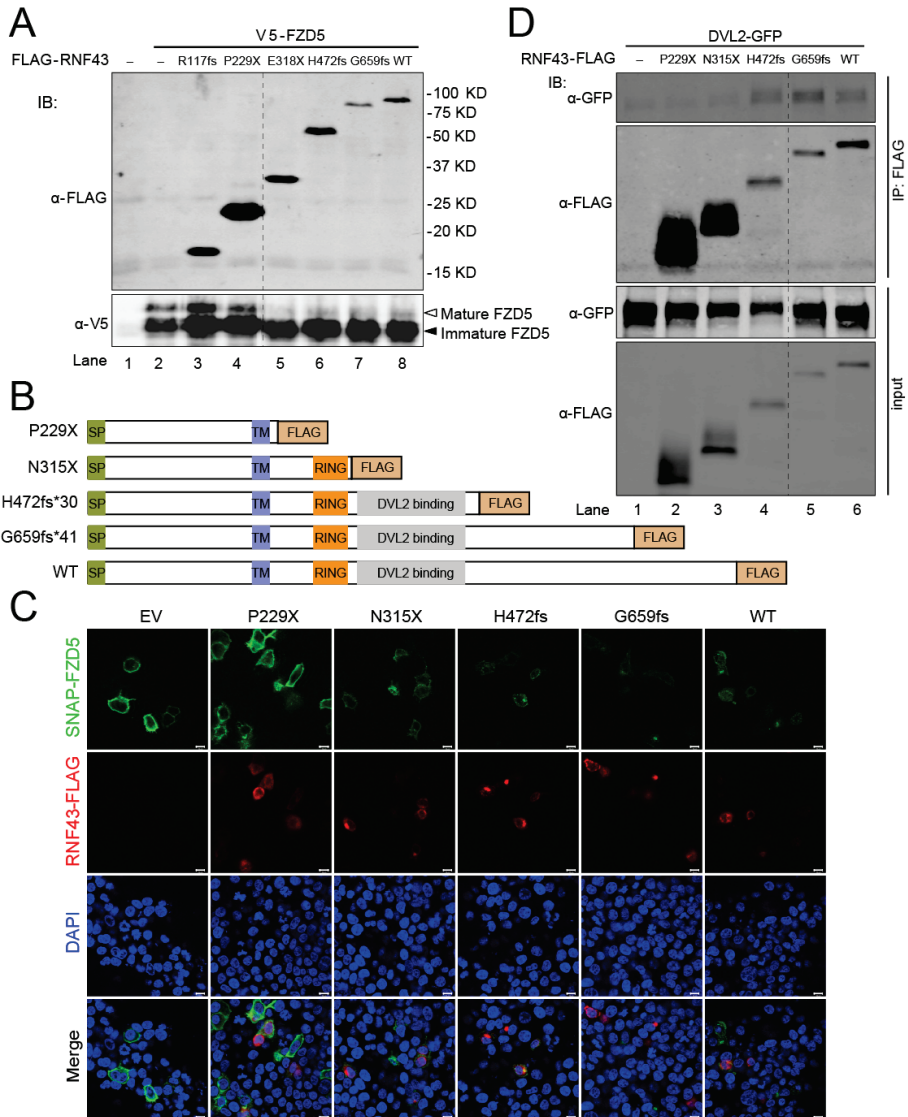
## Long RNF43 truncation mutants retain turnover capability of the Wnt-receptor

Next we examined the function of RNF43 truncation mutants on promoting turnover of the Wnt-receptor FZD using two independent methods. In both assays wild-type RNF43 and also the G659fs, H472fs and E318X/N315X variants effectively reduced mature-FZD5 levels (**Fig. 3a–c**; **Supplementary Fig. 1** shows the subcellular location of overexpressed wildtype-RNF43). In contrast, variants lacking the RING domain (R117fs and P229X) failed to downregulate mature-FZD5. Collectively, C-terminal truncation mutants retaining at least the RING domain still are effective in Wnt-receptor turnover when overexpressed.

Previously, it was shown that RNF43 requires binding to DVL to efficiently promote Wnt-receptor turnover (29). In line with expectation the H472fs variant and longer ones can still efficiently interact with DVL2 (**Fig. 3d**).



**Fig. 2** Characterization of *RNF43* truncating mutations in regulating  $\beta$ -catenin signaling. **a** Schematic image depicting the constructed N-terminal FLAG-tagged *RNF43* mutant expression vectors, mapped to the domain structure of *RNF43* protein. SP, signal peptide; TM, transmembrane. **b** HEK293T cells were co-transfected with WRE/MRE luciferase  $\beta$ -catenin reporter plasmids and vectors expressing wild-type (WT), mutant *RNF43* or empty vector (EV), in the absence or presence of Wnt3a-CM.  $\beta$ -catenin reporter activity was measured (in triplicate,  $n=3$  independent experiments). WRE/MRE ratio obtained for EV with Wnt3a-CM treatment was arbitrarily set to 100%. **c** Wnt3a-induced  $\beta$ -catenin reporter activity was measured in HEK293T cells transiently transfected with indicated amounts of vector (in triplicate,  $n=2$ ). Values are normalized to EV. **d**  $\beta$ -catenin reporter activity was measured in HCT116 cells (in triplicate,  $n=3$ ). Values are normalized to EV with Wnt3a-CM. All data shown in **b–d** are expressed as mean  $\pm$  SD. Significance tested in **b** and **d** using one-way ANOVA followed by Tukey's *post hoc* test.



**Fig. 3** Long RNF43 truncation mutants retain turnover capability of the Wnt receptor. **a** Immunoblot showing that levels of mature (glycosylated) V5-tagged Frizzled5 (FZD5) are substantially reduced in HEK293T cells upon transient co-transfection with long (p.E318X, p.H472fs, G659fs) truncation mutants and wild-type (WT) RNF43, but not with short (p.R117fs and p.P229X) RNF43 truncation mutants and empty vector (EV). **b** Schematic view showing the expression vectors encoding C-terminal FLAG-tagged RNF43 mutants. **c** Long (p.N315X, p.H472fs, and G659fs) mutants and wild-type RNF43 induce rapid internalization of membrane-located SNAP-FZD5. HCT116 cells were co-transfected with SNAP-tagged FZD5 and RNF43 variants. Cell-surface SNAP-FZD5 was labelled with SNAP-surface488 for 15min and then chased for 5 min. Scale bars, 10  $\mu$ m. **d** Interaction between ectopically expressed FLAG-tagged RNF43 mutants and GFP-tagged DVL2 in HCT116 cells, assessed by a co-immunoprecipitation assay. The vertical dashed line indicates removed irrelevant lanes. IB, immunoblotting; IP, immunoprecipitation. Original immunoblot images are shown in Supplementary Fig. 13.



### Evaluation of other reported interactors of RNF43

RNF43 was shown to potentially also inhibit  $\beta$ -catenin signaling by tethering TCF7L2 to the nuclear membrane (4). In their experiments RNF43 truncated after amino-acid 447 failed to bind to TCF7L2, suggesting it interacts with the C-terminal half. Accordingly, we observed a weak interaction with wild-type and G659fs RNF43, while the H472fs variant reproducibly binds substantially more TCF7L2 for unknown reasons (**Supplementary Fig. 2a**). Shorter variants did not bind to TCF7L2.

Next, we induced  $\beta$ -catenin signaling by exposing RNF43-transfected cells to extracellular Wnt-ligand and/or by blocking GSK3 activity. The latter will induce signaling intracellularly by blocking the  $\beta$ -catenin breakdown complex. In case the nuclear function was more dominant, both treatments would have shown a reduced signaling activity in the presence of RNF43. However, we observed that RNF43 and longer variants thereof, were only capable of inhibiting  $\beta$ -catenin signaling in cells exposed to Wnt-ligand (**Supplementary Fig. 2b, c**), suggesting the Wnt-receptor turnover function is the most relevant.

Some reports have linked RNF43 also to p53 breakdown (36–40). To this aim we transfected p53-negative H1299 cells with low amounts of a p53 expression vector and all RNF43 variants. As shown in **Supplementary Fig. 3**, none of the variants including wild-type RNF43 were able to inhibit p53-signaling, in contrast to the well-known p53 ubiquitin ligase MDM2 (41).

More recently, RNF43 was also suggested to degrade E-cadherin (42). In that report, A549 cells accumulated high levels of E-cadherin following *RNF43* knockdown, which we are however not able to reproduce (**Supplementary Fig. 4**). Likewise, we do not see any increase in E-cadherin levels in HEK293T and KM12 cells with RNF43 knockout.

### Endogenous levels of the G659fs variant reduce $\beta$ -catenin signaling

Thus far we have shown that overexpressed RNF43 and long variants thereof can effectively inhibit  $\beta$ -catenin signaling. To test this at endogenous levels, we employed CRC cell lines with various reported *RNF43* mutations (**Supplementary Table 2**). Evaluation of *RNF43* and *ZNRF3* expression (**Supplementary Figs. 5 and 6**) shows a clear expression in most cell lines. KM12 cells homozygously express G659fs *RNF43* at

substantial levels (in RKO cells this is basically undetectable). Therefore, we generated CRISPR/Cas9-mediated RNF43-knockout (KO) KM12 cells (**Supplementary Fig. 7**). *AXIN2* RNA levels, a well-established  $\beta$ -catenin target gene, were increased 1.8-fold in RNF43-KO cells (**Fig. 4a**). Wnt3a addition resulted in a slight increase in both parental and KM12-RNF43-KO cells, but especially when combined with R-Spondin, which mediates membrane-clearance of RNF43 and ZNRF3 (1,29,43–46), a significant increase in *AXIN2* expression was observed. KM12-RNF43-KO cells lack RNF43 protein that can be blocked by R-Spondin, so the increase is likely attributed to inhibition of ZNRF3. Therefore, we knocked-down *ZNRF3* expression. Combined addition of Wnt3a and R-spondin results in the highest  $\beta$ -catenin reporter levels in both KM12-parental and RNF43-KO cells (**Fig. 4b**). Importantly, similar levels are observed in Wnt3a-exposed RNF43-KO cells combined with *ZNRF3*-knockdown, while the presence of the G659fs variant in the KM12-parental cells prevents maximum activation (compare lanes 6).

Analysis of V5-FZD5 levels showed that mature-FZD5 is clearly higher in the KM12-RNF43-KO cells indicating a reduced Wnt-receptor turnover (**Fig. 4c**). Phosphorylated DVL2 (pDVL2) levels are higher in KM12-RNF43-KO cells under all tested conditions (**Fig. 4d**), which is directly triggered by FZD and thus generally is regarded as a straightforward indicator for Wnt-receptor activation (47,48).

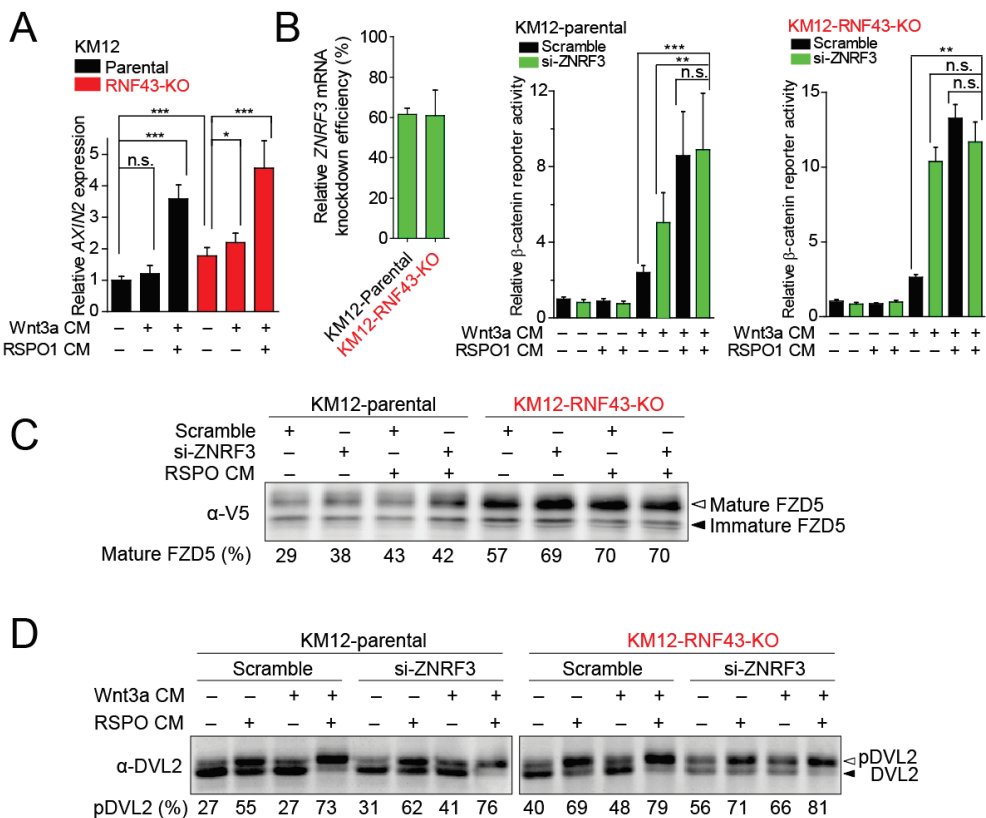
Taken together, at endogenous levels the long truncated G659fs RNF43 variant also retains functionality in regulating  $\beta$ -catenin signaling.

**Fig. 4** The G659fs truncation of RNF43 expressed at endogenous levels in KM12 cells, preserves  $\beta$ -catenin-suppressive properties. **a** *AXIN2* expression was measured in parental and RNF43-knockout (KO) KM12 cells, in the presence of Wnt3a-CM with/wo RSPO1-CM (in triplicate, n=3, normalized to *GAPDH* and to DMSO-treated parental KM12 cells). **b** Left panel, determination of knockdown efficiency of *ZNRF3* siRNA in KM12-parental and RNF43-knockout cells (in duplicate, n=3). Right panels, parental and RNF43-KO KM12 cells were co-transfected with WRE/MRE plasmids and *ZNRF3* siRNA in the absence or presence of Wnt3a- and RSPO1-CM as indicated.  $\beta$ -catenin reporter activity was then measured (in triplicate, n=3). Data are presented as WRE/MRE ratios, which are arbitrarily set to 1 for the untreated samples (lanes 1). **c** Immunoblot showing levels of mature V5-FZD5 in parental and RNF43-KO KM12 cells co-transfected with *ZNRF3* siRNA with/wo RSPO1-CM as indicated. **d** Immunoblot showing phosphorylated DVL2 (pDVL2) levels following different treatments, in parental and RNF43-KO KM12 cells. All data are expressed as mean  $\pm$  SD. Significance tested using one-way ANOVA followed by Tukey's *post hoc* test [**a** and **b** (parental KM12 cells)], and Kruskal-Wallis *H* test with a Dunn's *post hoc* test [**c** (RNF43-KO KM12 cells)]. Original immunoblot images are shown in Supplementary Fig. 13.

(see figure on next page)

**Confirmation in HEK293T cells not carrying additional Wnt-signaling related mutations**

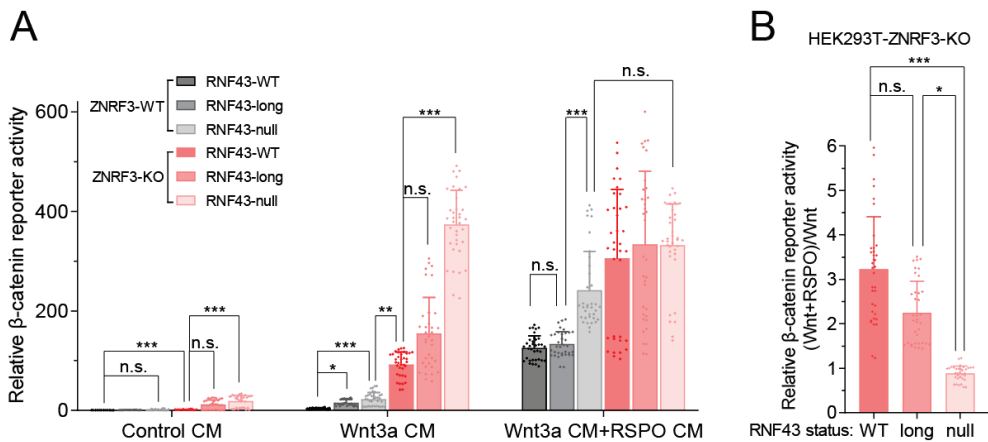
The KM12 cell line, besides carrying a G659fs *RNF43* mutation, is also mutant for AXIN1 (**Supplementary Table 2**). As such, the resulting level of  $\beta$ -catenin signaling is not exclusively dictated by mutant RNF43. In fact, this is the case for most cell lines in our panel. To determine in a more controlled manner the level of signaling imposed by the G659fs mutation at endogenous levels, we generated HEK293T clones expressing no or long truncated versions of RNF43. ZNRF3 was knocked-out simultaneously to prevent functional compensation by this homologous protein. We obtained three independent clones with all possible combinations of *ZNRF3* and *RNF43* mutations (**Supplementary Fig. 8**). Double-mutant cells show a very strong induction of  $\beta$ -catenin reporter activity when exposed to Wnt3a (**Fig. 5a** and



**Fig. 4** (see legend on previous page)

**Supplementary Fig. 9).** Importantly, ZNRF3-knockout clones expressing either wild-type or long truncated RNF43 variants show an intermediate effect on signaling, with no significant difference between both variants.

Next, we tested the R-spondin responsiveness of all Wnt3a-exposed ZNRF3-knockout clones. As expected,  $\beta$ -catenin signaling was not further increased in Wnt3a/R-spondin exposed double-knockout clones (**Fig. 5b**), while clones expressing wild-type or long RNF43 variants showed a 2–4-fold increase. These observations further confirm that long RNF43 mutants expressed at endogenous levels, retain  $\beta$ -catenin regulation that can be blocked by R-spondin.



**Fig. 5** Long G659fs-like RNF43 truncation mutants expressed at endogenous levels retain regulation of  $\beta$ -catenin signaling, confirmed by combinations of RNF43/ZNRF3 knockout in HEK293T cells. **a**  $\beta$ -catenin reporter activity was measured in ZNRF3 wild-type (WT) and knockout (KO) HEK293T cells simultaneously expressing CRISPR-Cas9 induced null, long mutant (R654fs/K655fs/R656fs/R657fs) or wild-type RNF43. Three independent clones were analyzed for each genetic combination. Wnt3a-CM (1:9) and/or RSPO1-CM (1:9) were added as indicated. **b** Relative R-spondin responsiveness of ZNRF3 KO clones showing that G659fs-like RNF43 truncations can still be inhibited by R-spondin to levels approaching those of wild-type RNF43. Data from **a** are shown as relative levels of Wnt3a and RSPO1 co-stimulated  $\beta$ -catenin reporter activity to Wnt3a-stimulated values. All data are represented as mean  $\pm$  SD (two technical replicates from three biological replicates,  $n=2$  independent experiments, achieved with combination of three independently derived clones for each genotype). Significance tested using Kruskal-Wallis  $H$  test with a Dunn's *post hoc* test (**a** and **b**).

**The G659fs *RNF43* mutation is not affected by nonsense-mediated decay or reduced protein stability**

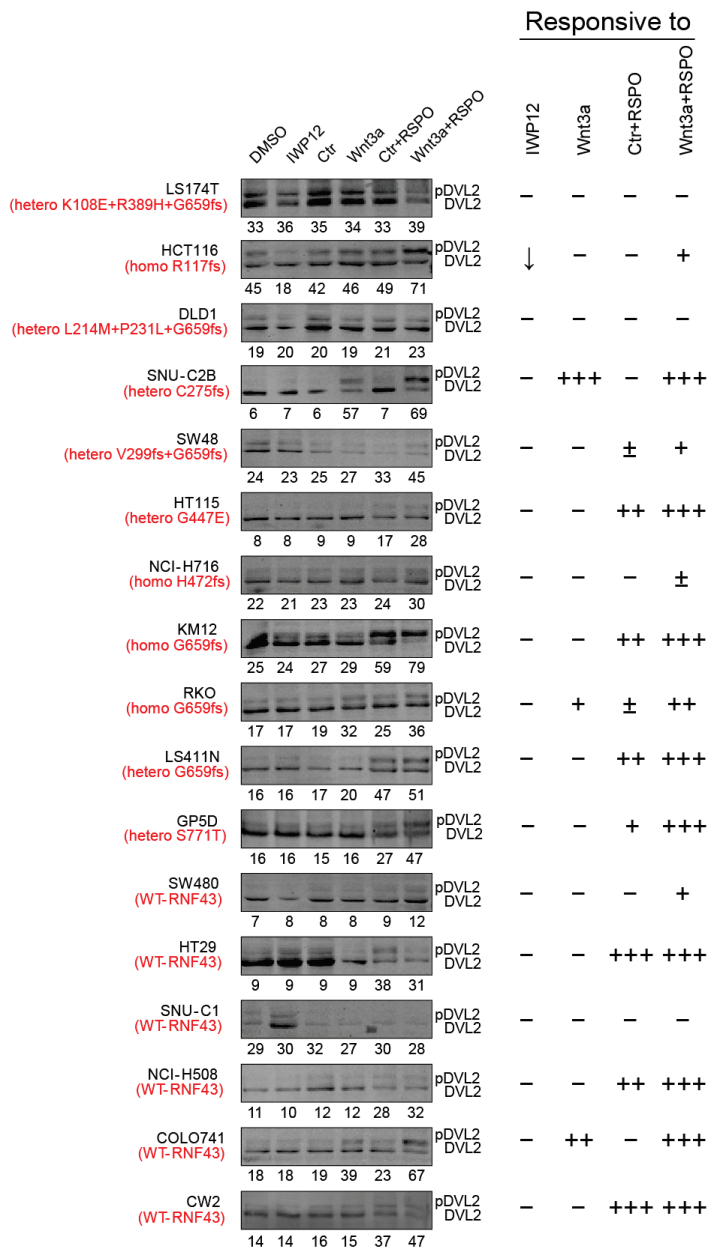
Truncating *RNF43* mutations may be affected by nonsense-mediated mRNA decay (NMD) (49). However, sequencing of genomic DNA and cDNA from three heterozygous G659fs-mutant CRC cell lines revealed no reduction in levels of mutant *RNF43* (**Supplementary Fig. 10**). In addition, both the wild-type and G659fs protein show a similar half-life (**Supplementary Fig. 11**). Collectively, NMD and reduced protein stability are unlikely to contribute to *RNF43*-mediated tumorigenesis for the G659fs mutations.

**Blocking Wnt-secretion does not affect pDVL2 levels in the majority of *RNF43*-mutant CRC cell lines**

Therapeutics targeting extracellular Wnt-ligand exposure can inhibit the growth of a subset of CRCs, including *RNF43*-mutant cancers (17,50,51). DVL2 phosphorylation occurs upstream of the  $\beta$ -catenin breakdown complex, hence it can be used to investigate the overall functionality of the Wnt receptor complex (including *RNF43/ZNRF3*), independent from downstream-acting mutations in APC or  $\beta$ -catenin (47,48). Therefore, we evaluated pDVL2 levels following treatment with the Porcupine inhibitor IWP12. In addition, we exposed the cells to Wnt3a, R-Spondin or a combination thereof.

Among the 17 CRC cell lines tested, only HCT116 cells (homozygous R117fs) showed a substantial pDVL2 reduction using IWP12 (**Fig. 6**). None of the other lines including the ten carrying *RNF43* mutations were affected in pDVL2 levels. Only three cell lines (SNU-C2B;RKO;COLO741) showed increased pDVL2 levels following Wnt3a treatment alone, which may be attributable to low *RNF43/ZNRF3* RNA levels or the presence of *ZNRF3* mutations (**Supplementary Figs. 5 and 6**). Finally, most cell lines showed a clear pDVL2 increase when exposed to R-Spondin, either alone or in combination with Wnt3a.

Collectively, these analyses show that (i) all tested *RNF43*-mutant lines except HCT116 do not require extracellular Wnt-ligand exposure to maintain pDVL2 at baseline levels; (ii) most cell lines express functional levels of *ZNRF3* and/or *RNF43* that can be blocked by R-Spondin to further increase pDVL2 levels. Keeping our preceding analyses in mind, these functional proteins can also be the long truncated *RNF43* variants.



**Fig. 6** Phosphorylation status of endogenous DVL2 in CRCs harboring different *RNF43* mutations, and its responsiveness to several treatments. (Left panels) pDVL2 levels following different treatments assessed by immunoblot. Values below the images represent percentage of total DVL2 that is in the phosphorylated form (upper band). *RNF43* mutations in the CRC cells are shown in between parentheses. (Right panels) The responsiveness of each cell line to respective treatments is summarized, according to changes in pDVL2 levels observed in DMSO control (↓, <50%; ±, >30%; +, >50%; ++, >100%; +++, >200%). Homo, homozygous; Hetero, heterozygous; Ctrl, Control-conditioned medium; Wnt3a, Wnt3a-conditioned medium; RSPO, RSPO1-conditioned medium; pDVL2, phosphorylated DVL2. Original immunoblot images are shown in Supplementary Fig. 13.

## Cell viability and anchorage-independent growth assays following IWP12 treatment

Investigating the growth-inhibitory effects of IWP12 on RNF43-mutant cell lines showed that HCT116 viability was significantly impaired (**Fig. 7a**), while lines harboring wild-type (SW480;LOVO) and biallelic C-terminal *RNF43* truncating mutations (NCI-H716;KM12) were not affected.

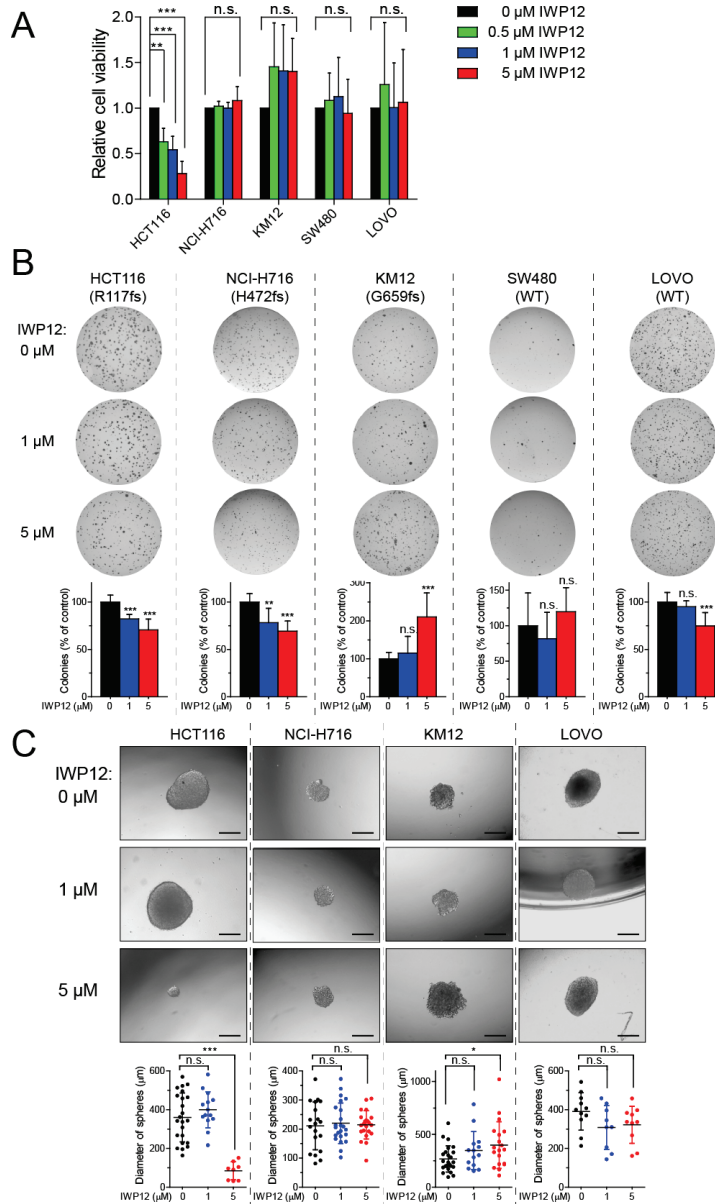
IWP12 also significantly reduced the number and size of HCT116 colonies in anchorage-independent assays (**Fig. 7b, c**). NCI-H716 and LOVO showed a significant reduction in colony numbers, while size was unaffected. For KM12 the counterintuitive result was obtained that IWP12 led to an increase in colony size and numbers. The non-RNF43-mutant line SW480 was not affected in colony numbers (**Fig. 7b**), while it failed to grow in the sphere forming assay. Collectively, these results show that only HCT116 cells respond consistently to Porcupine inhibition.

## C-terminal truncating *RNF43* mutations often co-occur with other $\beta$ -catenin enhancing mutations

Most of the RNF43-mutant lines in our panel (10/11) harbor mutations in other genes that are expected to enhance  $\beta$ -catenin signaling. This seems in apparent contrast with previous statements about mutual exclusivity with *APC* mutations (5). To further explore the co-occurrence of *RNF43* truncating mutations with other  $\beta$ -catenin enhancing mutations, we screened cancer genomic datasets (**Supplementary Table 3**) (19–21). Codon 454 was used as demarcation between 5'- and 3'-RNF43-mutant cancers. The majority of patients carrying C-terminal truncating RNF43 mutations concurrently harbor additional  $\beta$ -catenin enhancing mutations (34/58, 58.6%), which is significantly higher than patients carrying N-terminal mutations (15/48, 31.3%) (Fisher's exact test,  $P = 0.0062$ ).

**Fig. 7** Cell viability and anchorage independent growth of CRC cell lines harboring homozygous RNF43 truncating mutations following IWP12 treatment. **a** CRC cell lines harboring different homozygous *RNF43* mutations were treated with different concentrations of Porcupine inhibitor IWP12 for 72 hours. SW480 and LOVO were taken along as RNF43 wild-type controls. Only HCT116 cells expressing R117fs RNF43 show reduced viability. Cell viability was assessed by the AlamarBlue assay and presented as values relative to the DMSO control. Mean  $\pm$  SD of five independent experiments (8–12 replicates).

(legend continued on next page)



**Fig. 7** (legend continued from previous page) **b** A soft-agar colony formation assay to test anchorage-independent growth was performed on CRC cell lines, treated with Porcupine inhibitor IWP12 at different concentrations for 14 days. Medium was refreshed every third day. Representative images are shown in the upper panels. Lower panel shows the number of colonies relative to the DMSO control (in triplicate,  $n=3$ ). **c** Representative images showing that IWP12 treatment only strongly reduces sphere formation of HCT116 cells. Single cells of respective CRC cell lines obtained from flow cytometry were subjected to a sphere formation assay and treated with IWP12 (0, 1, and 5 μM; DMSO as solvent control) for 14 days. The bottom dot plots show the diameter of the spheres that were observed (performed in 20 replicates,  $n=3$ ). SW480 failed to form spheres. Scale bars, 200 μm. Data shown as mean  $\pm$  SD. Significance tested using one-way ANOVA followed by Tukey's *post hoc* test (**a-c**).



## Discussion

*RNF43* mutations have gained substantial interest during the last decade, as they represent a commonly observed alternative route to aberrantly enhance Wnt/ $\beta$ -catenin signaling in cancers. Moreover, they may denote tumors that respond to Wnt-inhibitors. *RNF43* mutations are mostly truncating ones and especially common in MMR-deficient tumors, with a characteristic G659fs mutation observed in the majority of *RNF43*-mutant tumors. Here we show that this and other C-terminal truncating mutations do not confer a clear Wnt-dependency onto CRC cells.

Our finding is based on the following observations: (i) Separately analyzing MMR-proficient tumors shows that truncating mutations predominantly occur before codon 454, strongly suggesting that long truncating mutations do not provide a selective advantage. (ii) Overexpression of long truncated *RNF43* variants suppresses  $\beta$ -catenin signaling similarly to wild-type *RNF43*, in accordance with previous work (52). (iii) Also when expressed at endogenous levels in HEK293T cells, long truncated *RNF43* variants retain a significant level of  $\beta$ -catenin regulation that can still be inhibited by R-spondin. Furthermore, knockout of endogenous G659fs expression in KM12 cells, resulted in increased  $\beta$ -catenin signaling. (iv) The majority of tumors carrying C-terminal truncations show co-occurrence with other gene mutations expected to enhance  $\beta$ -catenin signaling. (v) Among our panel of *RNF43*-mutant CRC cell lines, only HCT116, carrying a homozygous N-terminal truncating *RNF43* mutation, shows reduced pDVL2 levels and a consistent growth inhibition following Wnt-secretion blockade.

The ZNRF3/*RNF43* module harbors a strong level of functional redundancy. For example, knockout of each gene separately has no discernable effect on mouse intestinal homeostasis (2). Only when both genes are inactivated simultaneously, a clear increase in the proliferative compartment was observed. Moreover, mouse intestinal organoids can only grow R-spondin independently when both genes are inactivated (2,13). Likewise, we observed that  $\beta$ -catenin signaling is maximally enhanced in HEK293T and KM12 cells when both genes are inactivated. For KM12 this required the knockout of the G659fs allele. In mouse intestinal organoids, G659fs-like

mutations can result in R-spondin independent growth, but importantly only in the context of simultaneous ZNRF3-knockout (14). Unfortunately, no  $\beta$ -catenin signaling analyses were performed, as our results and those of Tsukiyama et al. predict that the ones carrying the G659fs-like mutations will have lower  $\beta$ -catenin signaling activity compared with full knockout clones (52).

The minor loss in functionality observed for the G659fs mutation, raises doubts whether the mere identification of this mutation within a tumor marks it as a prime candidate for treatment with Wnt-inhibitors, especially in the context of retained ZNRF3 function. This may explain why the majority of G659fs tumors carry additional mutations expected to enhance  $\beta$ -catenin signaling, such as truncating *APC* and oncogenic  $\beta$ -catenin mutations, for which the tumorigenic potential is well-recognized. We also observe truncating *AXIN1/2* mutations in about 20% of the G659fs lesions. The functional relevance of these *AXIN1/2* mutations for CRC is still not fully understood (53). However, in support of a tumor-promoting role through enhanced  $\beta$ -catenin signaling, acquired *AXIN1* loss in CRC cells carrying an *RSPO3*-fusion, was observed to lead to resistance for Wnt-inhibitors (51).

Although our analyses are based on cell lines, the results are largely in line with reported primary CRC organoid-cultures. Van de Wetering et al. showed that a RNF43-A355fs mutant organoid was impeded by Porcupine inhibition, while a G659fs-mutant organoid was unresponsive (54). Fujii et al. established four organoids carrying RNF43 mutations (55). Two organoids with N-terminal truncating or missense mutations at conserved residues were Wnt/R-spondin dependent, while among two G659fs-mutant organoids, only the one carrying an additional *ZNRF3* mutation required Wnt/R-spondin for its growth.

The underlying mechanism for the retained functionality most likely resides in the ability to bind to DVL, which is essential for RNF43/FZD-receptor interaction and FZD-turnover (29). DVL-binding mapped to residues 325–454, which fits our observation that RNF43-H472fs can still interact. An independent investigation mapped DVL-binding to residues 478–596 (52), which appears in disagreement with our results. Looking at evolutionary conserved RNF43 residues calculated by the

ConSurf tool (<http://consurf.tau.ac.il/2016/>) (56) (**Supplementary Fig. 12**), the region between aa 438–512 seems highly conserved, possibly suggesting that this extended domain may be required for optimal DVL association. The conserved residue analysis also shows that C-terminal RNF43 is the least conserved. No specific function has been attributed to this domain, but our results show that it appears not to be important for regulating  $\beta$ -catenin signaling. Currently regulating Wnt-signaling, both canonical as well as non-canonical (52), is the only known function for RNF43. Other functions may be uncovered in the future. In support Neumeyer et al. observed gastric hyperproliferation in RNF43-mutant mice without a concomitant increase of  $\beta$ -catenin target genes, although an effect on non-canonical Wnt-signaling was not investigated (57).

Our observations have important implications for several upstream Wnt-inhibitors in development by pharmaceutical companies and in clinical trials (ClinicalTrials.gov Identifier: NCT01351103; NCT02278133; NCT02521844; NCT01973309; NCT01345201; NCT02092363). All these agents are directed at blocking the extracellular activation of Wnt-receptors, either by blocking Wnt-secretion or antibodies directed towards the FZD/LRP receptors. Prime candidates to show responsiveness are tumors expressing R-Spondin fusions and RNF43 mutant cancers (3). Our results show that among the RNF43-mutant cancers only the ones carrying N-terminal truncations are likely to respond. As tumors carrying the G659fs mutation are the most common, clinical trials should evaluate these tumors cautiously as this mutation may have no or little potential for inducing Wnt/ $\beta$ -catenin driven tumorigenesis.

## Materials and methods

Information about plasmids used in the present study, immunoblotting, immunocytofluorescence staining, nonsense-mediated decay determination, cycloheximide chase assay, quantitative real-time PCR (qRT-PCR), transfection protocols, measurement of cell viability, soft agar colony formation assay, and sphere formation assay is described in the Supplementary Methods. Original immunoblot images and primers used for qRT-PCR are provided, respectively, in **Supplementary Fig. 13** and **Supplementary Table 4**.

### Cell lines and culture

HEK293T, HCT116, SW480, SW48, GP5D, LOVO, DLD1, HT29, Caco2, RKO, HT115, CAR1, and A549 cells were maintained in Dulbecco's Modified Eagle's medium (DMEM) (Lonza, Breda, The Netherlands) supplemented with 10% fetal bovine serum (FBS; Gibco, Bleiswijk, The Netherlands). CW-2, SNU-C2B, SNU-C1, LS411N, NCI-H508, COLO741, and H1299 cells were cultured in Roswell Park Memorial Institute (RPMI)-1640 culture medium (Lonza) containing 10% FBS. NCI-H716 and KM12 cells were cultured in RPMI-1640 culture medium containing 10% FBS and 1 mM sodium pyruvate (Gibco). LS174T cells were cultured in RPMI-1640 culture medium supplemented with 5% FBS. Culture media were changed every 2–3 days. All the cell lines were cultured in a humidified incubator maintained at 37 °C with 5% CO<sub>2</sub>. Identity of all cell lines and clones thereof, was confirmed by the Erasmus Molecular Diagnostics Department, using Powerplex-16 STR genotyping (Promega, Leiden, The Netherlands) in October 2018. All cell lines tested negative for mycoplasma based on the real-time PCR method at Eurofins GATC-Biotech (Konstanz, Germany). *RNF43* mutation status depicted in **Supplementary Table 2** was confirmed in all cell lines by Sanger sequencing and was consistent with those reported at COSMIC, the Catalogue Of Somatic Mutations In Cancer (<http://cancer.sanger.ac.uk>) (58).

For the Wnt3a treatment, growth medium containing 25% L-Wnt3a conditioned medium (Wnt3a-CM) or L-control conditioned medium (Control-CM) was used as described previously (59). To obtain the L-Wnt3a or L-control conditioned media,

genetically engineered mouse fibroblast L-cell line stably expressing Wnt3a (L-Wnt3a cells) or control L-cells (L-control cells) were seeded onto 100-mm culture dishes at 10% cell confluency in DMEM plus 10% FBS for 4 days. After collecting the first batch of conditioned medium, dishes were replenished with fresh growth medium, and the cells were incubated for an additional 3 days. The harvested medium was then filtered and stored at 4 °C. Roof plate-specific Spondin-1 (R-spondin 1, RSPO1)-conditioned medium was prepared using the same procedures as described for Wnt3a-CM. For the RSPO1-CM treatment, growth medium containing 10% RSPO1-CM was used.

### **Construction of RNF43 variant expression vectors**

Expression vectors for N-terminal FLAG-tagged RNF43 variants were generated using the pcDNA-5'UT-FLAG vector as basis (generously given by Dr. Veronique Lefebvre, Lerner Research Institute, Cleveland, Ohio, USA). The full-length cDNA of the p.R117fs variant of human RNF43 was acquired from the human colorectal cancer HCT116 cell line by RNA purification using NucleoSpin RNA kit (Macherey-Nagel, Düren, Germany) and reverse transcription (RT)-PCR using specific primers and High capacity cDNA reverse transcription kit (Applied Biosystems, Foster City, CA, USA), and then the ORF of the p.R117fs RNF43 variant was assembled into the N-terminal FLAG-tagged vector using the Gibson Assembly Master Mix (New England Biolabs, Ipswich, MA, USA) according to the manufacturers' protocol. The wild-type, p.P229X, p.E318X, p.H472fs, and p.G659fs N-terminal FLAG-tagged RNF43 variants were generated by the Q5 Site-directed Mutagenesis Kit (New England Biolabs) using the p.R117fs RNF43 clone as basis. The corresponding C-terminal FLAG-tagged RNF43 variants were constructed following the same strategy with the Q5 site-directed Mutagenesis Kit (New England Biolabs) by first removing the N-terminal FLAG epitope sequence and then adding it to the C-terminal end of wild-type RNF43. All C-terminal variants with frameshift mutations were generated such that also the frameshift encoded amino acids are included in the final protein. Primer sequences are available upon request. All RNF43 expression plasmids were full-length sequence verified.

## CRISPR-Cas9 genome editing

RNF43-knockout (KO) KM12 cells were generated via CRISPR-Cas9 genome editing. A single guide RNA (sgRNA) targeting *RNF43* exon 2 was designed using the following CRISPR design tool (<http://crispr.mit.edu/>), and cloned into pSpCas9(BB)-2A-GFP (PX458), a gift from Feng Zhang (Addgene plasmid # 48138), using standard procedures (60). The sgRNA sequence is listed in **Supplementary Fig. 7**. Cells were seeded into 6-well plates and transfected with 1500ng of PX458 using FuGENE HD transfection reagent (Promega) following the manufacturer's instruction at a 3:1 ratio of FuGENE HD. After transfection for 48 h, single GFP-positive cells were sorted out and plated into 96-well culture plates by a fluorescence activated cell sorter (FACS) FACSaria II cell sorter (BD Biosciences, San Jose, CA, USA). Three weeks later, Genomic DNA was extracted from expanded single cell clones by QuickExtract DNA Extraction Solution (Epicentre, Madison, WI, USA). A 550 bp fragment encompassing the sgRNA site was PCR-amplified followed by Sanger sequencing (Macrogen, Amsterdam, Netherlands) to verify whether *RNF43* is knocked out (disruption of RNF43 ORF). The sequence alterations in *RNF43* are depicted in **Supplementary Fig. 7**. Knockout clones (combination of *RNF43* and/or *ZNRF3* mutations) of HEK293T were generated following the same strategy mentioned above. The sequence alterations in *RNF43* and *ZNRF3* of the HEK293T knockout clones are depicted in **Supplementary Fig. 8**. The primers used for identifying sequence alterations are available upon request.

## Luciferase reporter assays

The luciferase reporter assays were carried out as previously reported (59). Briefly, after transient transfection for 24 h, cells were stimulated 24 h with Wnt3a-CM, Control-CM and/or GSK-3 inhibitor SB216763 (Sigma-Aldrich, Zwijndrecht, the Netherlands) before harvest, and then Firefly and Renilla luciferase activity was measured with the Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer's instruction on LumiStar Optima luminescence counter (BMG LabTech, Ortenberg, Germany). All luciferase reporter assays were carried out in

triplicate. Firefly luciferase value was normalized to its internal control Renilla luciferase readout to calculate relative luciferase unit (RLU), and WRE/MRE ratio was reported unless otherwise specified.

### **Co-immunoprecipitation**

HEK293T and HCT116 cells in 6-well plates were transiently transfected with 1 $\mu$ g of C-terminal FLAG tagged-RNF43 plasmids or empty vector control, or DVL2-GFP (kindly given by Dr. Mariann Bienz) per well using FuGENE HD for 48 h. Cells were washed twice with cold PBS and lysed with 500  $\mu$ l of IP buffer (30 mM Tris-HCl, 150 mM NaCl, 5 mM EDTA, 5 mM NaF, 1% TritonX-100) containing Halt Protease and Phosphatase Inhibitor Single-use Cocktail (Thermo Fisher Scientific) and incubated 20 min on ice. After pelleting the cell debris by centrifugation at 13,000 rpm for 10 min at 4 °C, supernatants (cell extracts) were incubated with 20  $\mu$ l of pre-cleared anti-FLAG M2 Affinity Agarose Gel (cat. # A2220, Sigma-Aldrich) for 2 h at 4 °C with rotating. After 4 times extensive washing with IP buffer, the immunoprecipitated bead conjugates were heated in 2 $\times$  Laemmli sample buffer (95 °C, 10 min) and subjected to SDS-PAGE gel, followed by immunoblotting.

### **Surface SNAP-labeling and immunofluorescence analysis**

HCT116 cells were grown on 8-well Lab-Tek II chamber sides (Thermo Scientific). After 48 h cells were co-transfected with 1 $\mu$ g of SNAP-FZD5 (kindly given by Dr. Mariann Bienz) (2) and 1  $\mu$ g of C-terminal FLAG-RNF43 plasmids. Cells were labelled with 1  $\mu$ M SNAP-surface488 (New England Biolabs) substrate at room temperature for 15 min. After labeling medium was removed, cells were washed with growth medium once, then new growth media was added. Labelled surface FZD5 proteins were chased for 5 or 30 min at 37 °C, then cells were fixed with 4% PBS-buffered formaldehyde solution at room temperature for 10 min. After PBS washing 3 times, cells were permeabilized with 0.2% Triton X-100 in PBS for 10 min, washed with PBS and incubated with blocking solution (BSA 3% in PBS) for 30 min at room temperature. Cells were then stained with anti-FLAG antibody (1:500, cat. # F1804,

Sigma-Aldrich) overnight at 4 °C. After 3 times washing with PBS, cells were incubated for 1 h at room temperature with donkey anti-mouse Alexa594 secondary antibody (1:300, cat.# A21203, Invitrogen) protected from light. Immunofluorescence was performed by mounting cells in Vectashield antifade mounting medium with DAPI (cat. # H-1200, Vector Laboratories, Burlingame, CA, USA). Images were captured by a Zeiss LSM510 Meta confocal laser scanning microscope using ZEN 2009 software with constant parameter setting.

## Statistical analysis

The Fisher's exact test was used to compare proportions. All continuous variables are presented as the mean  $\pm$  standard deviation (SD). For comparisons of the normally distributed data between two groups, a two-tailed Student's *t*-test was used, whereas non-normally distributed data were compared using non-parametric Mann-Whitney *U* test. Significance of difference between three or more groups was determined using one-way analysis of variance (ANOVA) followed by Tukey's *post hoc* test for normally distributed variables, while Kruskal-Wallis *H* test with a Dunn's *post hoc* test was performed in the case of non-normally distributed data. The Kolmogorov-Smirnov normality tests were performed for normal distribution. Statistical differences were considered significant when the *P* value was  $<0.05$  on a two-sided basis (\**P*  $< 0.05$ , \*\**P*  $< 0.01$ , and \*\*\**P*  $< 0.001$ ). n.s. denotes not significant. All the statistical analyses were carried out using software GraphPad Prism version 5.01 (GraphPad Software Inc., San Diego, California, USA).

## Supplementary Methods

### Plasmids used in the present study

RNF43 variant expression vectors constructed as described in the Methods. PG13-luc (a luciferase reporter plasmid containing 13 copies of p53-binding consensus sequence) was a gift from Dr. Bert Vogelstein (Johns Hopkins University, USA; Addgene plasmid #16442). Plasmids HA-MDM2 and HA-p53 used in this work were a kind gift from Dr. Aart G. Jochemsen (LUMC, Leiden, The Netherlands). HA-tagged RNF43 plasmid was



a kind gift from Dr. Markus Gerhard (Technische Universität München, Munich, Germany).

### **Immunoblotting analysis**

Immunoblotting was carried out using standard methods. Cells were lysed in 2× Laemmli sample buffer with 0.1 M dithiothreitol (DTT) and heated for 10 min at 95°C. Proteins were separated on 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The proteins were then transferred to Immobilon-P PVDF membranes (Millipore, Bedford, MA, USA). Membranes were blocked 1 h with Odyssey blocking buffer (Licor-Biosciences, Lincoln, NE, USA) at room temperature and incubated overnight with primary antibodies at 4°C. After washing with PBS/0.05% Tween20 (PBST) buffer 10 min three times, the membranes were incubated 1 h with IRDye 680 goat anti-mouse (1:10.000) or IRDye 800 goat anti-rabbit (1:5.000) secondary antibodies (Licor-Biosciences), and then washed with PBST 10 min three times. The membranes were then scanned on the Odyssey Infrared Imaging System (Licor-Biosciences). The primary antibodies used above were anti-V5 (1:1000, cat. # 13202, Cell Signaling Technology), anti-DVL2 (1:1000, cat. # 3216, Cell Signaling Technology, Danvers, MA, USA), anti-FLAG (1:1000, cat. # F1804, Sigma-Aldrich, St. Louis, MO, USA), anti-GFP (1:1000, cat. # A-11122, Thermo Fisher Scientific, Waltham, MA, USA), anti-TCF7L2 (1:1000, cat. # ab76151, Abcam, Cambridge, MA, USA), anti-RNF43 (1:500, cat. # HPA008079, Sigma-Aldrich), anti-E-cadherin (1:500, cat. # 1702-1, Epitomics, Burlingame, CA, USA), anti-β-actin (1:1000, cat. # sc-47778, Santa Cruz, CA, USA). For electrochemiluminescence (ECL)-based detection methods, Immobilon® Block - FL (Fluorescent Blocker) blocking buffer was used (cat. # WBAVDFL01, Millipore). The primary antibodies were diluted in this blocking buffer. Membranes were washed with TBS containing 0.05% Tween 20 (TBST). The following secondary antibody used was goat anti-rabbit Immunoglobulins/HRP(1:10000, cat. # P0448, DAKO, Santa Clara, USA). Membranes were then incubated with Immobilon ECL Ultra Western HRP Substrate (Millipore) and visualized by using Amersham Imager 600 (GE Healthcare). Band intensity was

quantified by densitometry using software Image Studio Lite version 5.2 (Licor-Biosciences). Phosphorylation levels of DVL2 are presented as the percentage of total DVL2 that is in its phosphorylated form.

### **Immunocytofluorescence staining**

HCT116 cells were seeded on cover slips and reverse transfected with 1 µg of RNF43 expression plasmid, either with C-terminal HA- or FLAG-tag. The following day, cells were washed with PBS, fixed in ice-cold methanol/acetone (1:1) for 15 min, followed by 3 washes in PBS. Next, cells were incubated at RT for 15 min with blocking solution (PBS, 3% BSA, 1% saponin, 0.5% Triton X-100). Samples were incubated with primary antibodies at 4°C overnight in blocking solution, followed by 3 washes with PBS containing 3% BSA and 1% saponin. Primary antibodies were diluted as follows: anti-FLAG (1:300, cat. # F1804, Sigma-Aldrich); HA-tag (1:150, cat. # 2367, Cell Signaling Technology); anti-Calnexin (1:150, cat. # 2679, Cell Signaling Technology). The following secondary reagents were used: Donkey anti-Rabbit-Alexa 488 (cat. # A-21206), Donkey anti-Mouse-Alexa 594 (cat. # A-21203); all from Thermo Fisher Scientific, at 1:400 dilution at RT for 1 h. Following 3 washes with PBS containing 3% BSA and 1% saponin, slides were mounted with Vectashield antifade mounting medium with DAPI (H-1200, Vector Laboratories). Images were generated using a Zeiss LSM700 confocal laser scanning microscope.

### **Genomic DNA and cDNA sequencing for checking nonsense-mediated mRNA decay**

The genomic DNA and cDNA from SW48, DLD1, and LS411N cells were prepared according to the procedures mentioned before. Sanger sequencing was performed to check if nonsense-mediated mRNA decay is present. The primers are available upon request.

### **Cycloheximide chase assay**

To determine the turnover rates of the RNF43 variants in HEK293T cells, a cycloheximide chase assay was performed. Briefly, cycloheximide (Sigma-Aldrich) was added at a final concentration of 100 µg/mL in normal growth medium after 40 h transfection, with cells harvested at each time point (0, 1, 3, 5, 7, and 9 h). Half-lives of RNF43 variants in HEK293T cells were shown by a nonlinear regression analysis (one phase exponential decay) using software GraphPad Prism version 5.01 (GraphPad Software Inc., San Diego, California, USA).

### **Quantitative real-time PCR (qRT-PCR)**

Briefly, total RNA was isolated by using the NucleoSpin RNA II kit (Macherey-Nagel), then the RNA was reverse transcribed with Primescript RT reagent kit (TaKaRa) according to the manufacturer's instruction. Quantitative PCR was performed in the StepOne Real-Time PCR System (Applied Biosystem). Analyses were performed by using the StepOne version 2.0 software (Applied Biosystem) with the comparative  $\Delta\Delta CT$  method and normalized with the human housekeeping gene *GAPDH*. All experiments were performed in triplicate. Primer sequences are provided in **Supplementary Table 4**.

### **Transient transfection**

Transient transfections of RNF43 variant expression vectors in HEK293T and HCT116 cells were performed with FuGENE HD transfection reagent (Promega) following the manufacturer's instruction at 3:1 ratio of FuGENE HD. For luciferase reporter assays, briefly, cells were seeded onto 24-well plates in their growth medium and left for 24 h, allowing cell confluency to reach 80–90%. Each well was co-transfected with 200ng Wnt responsive element (WRE) or Mutant responsive element (MRE) Firefly luciferase reporter constructs, 10ng pCMV-Renilla control vector, and 10-100ng (amount was specified in the text) of RNF43 plasmid for 48 h. For co-transfection of plasmids and siRNA in the HCT116, SW480, NCI-H716, and KM12 cells each well was co-transfected with 200-400ng WRE or MRE reporter constructs, 10-20ng pCMV-

Renilla control vector, and 0.5-1µl of 20µM ON-TARGETplus SMARTpool siRNA (RNF43-cat.# L-007004-00-0005, ZNRF3-cat# L-010747-00-0005, Dharmacon, Lafayette, CO, USA) or ON-TARGETplus Non-targeting Pool (scramble control, cat.# D-001810-10-05) using Lipofectamine 2000 transfection reagent (Invitrogen) for 48 h according to the manufacturer's instruction. For siRNA interference only, cells were transfected with ON-TARGETplus SMARTpool siRNAs using Lipofectamine RNAiMAX (Thermo Fisher Scientific) for 48 h following the manufacturer's protocol.

The individual catalog numbers and target sequences of the siRNAs in ON-TARGETplus SMARTpool are listed as followed:

Scramble-#1 (D-001810-01): 5'-UGGUUUACAUGUCGACUAA-3';  
 Scramble-#2 (D-001810-02): 5'-UGGUUUACAUGUUGUGUGA-3';  
 Scramble-#3 (D-001810-03): 5'-UGGUUUACAUGUUUUCUGA-3';  
 Scramble-#4 (D-001810-04): 5'-UGGUUUACAUGUUUUCCUA-3'.  
 siRNF43-#1 (J-007004-09): 5'-GCAGAACAGAAAGCUAUUA-3';  
 siRNF43-#2 (J-007004-10): 5'-UAUGAUGUGUGGAUCCUAA-3';  
 siRNF43-#3 (J-007004-11): 5'-GGAGAAAGCUAUUGCACAG-3';  
 siRNF43-#4 (J-007004-12): 5'-GGUGGAGUCUGAAAGAUCA-3'.  
 siZNRF3-#1 (J-010747-05): 5'-UCAAGAGGCCGGUGGUGUA-3';  
 siZNRF3-#2 (J-010747-06): 5'-UCACUUCUGUCAUCCUAUU-3';  
 siZNRF3-#3 (J-010747-07): 5'-GAAAUUGGGAACUGUAUGG-3';  
 siZNRF3-#4 (J-010747-08): 5'-GCAGAUUUAGGUUAGGUA-3'.

### Measurement of cell viability

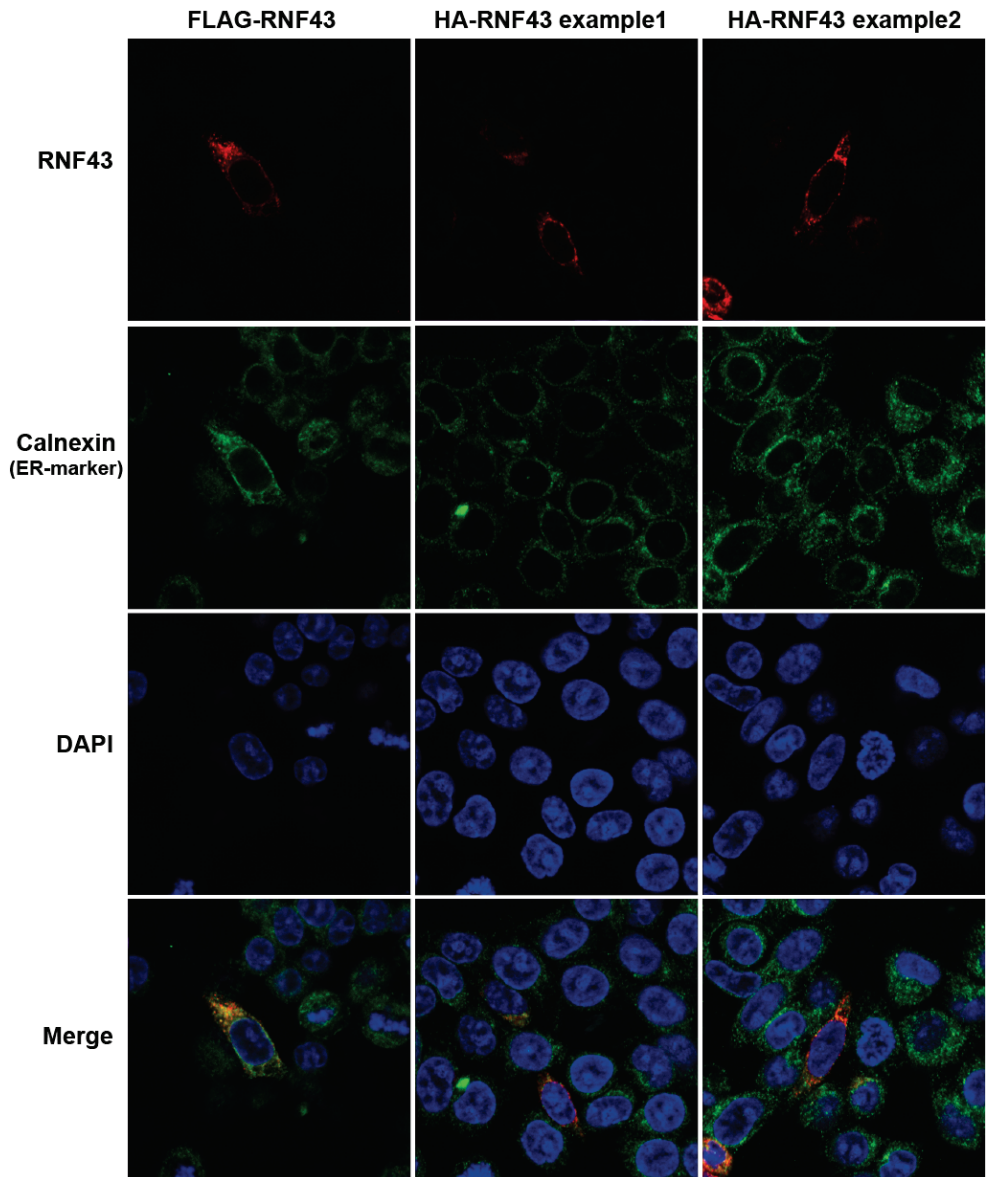
The effect of IWP12 (Sigma-Aldrich) on cell proliferation was evaluated using the AlamarBlue Cell Viability Assay (Thermo Fisher Scientific) according to manufacturer's instructions. Briefly, cells were incubated in culture medium and treated with IWP12 (0.5µM, 1µM, and 5µM) or DMSO control for 72 h, and cells were then exposed to AlamarBlue reagent for 3 h. The fluorescence signal of the AlamarBlue assay was then measured by using the CytoFluor Series 4000 Fluorescence Multi-Well Plate Reader (PerSeptive Biosystems).

### **Soft agar colony formation assay**

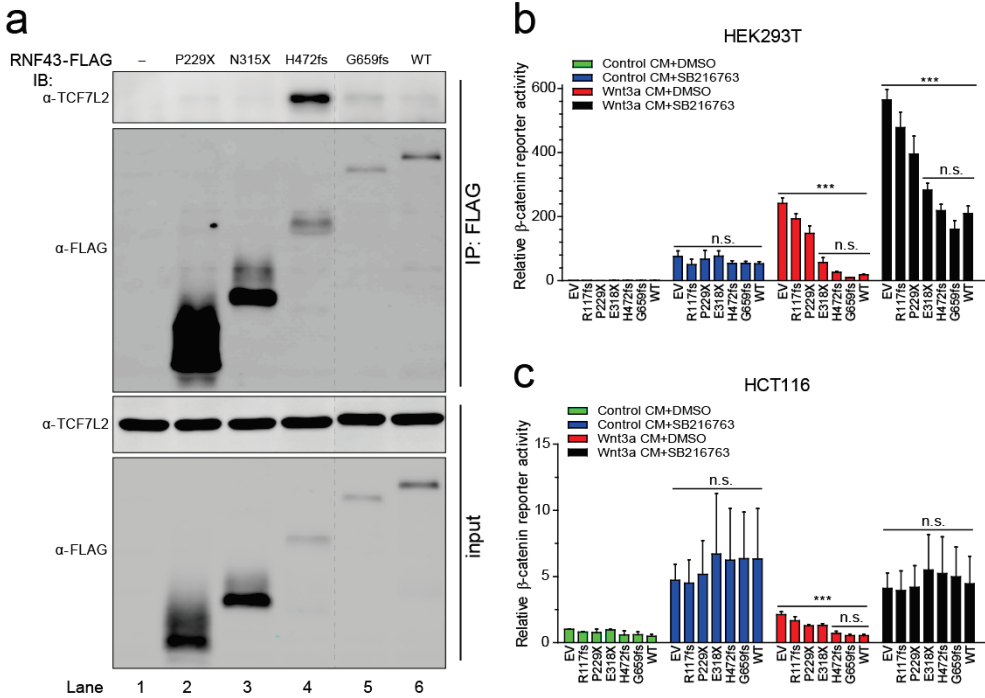
HCT116 (1000 cells), SW480 (2000 cells), LOVO (2000 cells), KM12 (2000 cells), and NCI-H716 (2000 cells) were suspended in 0.3% low-melting-point agar (Sigma) with the sphere forming medium (SFM) consisting of serum-free DMEM/F12 media supplemented with B-27 supplement (1:50, Gibco), N-2 supplement (1:100, Gibco), 20 ng/mL epidermal growth factor (EGF, Gibco), 20 ng/mL fibroblast growth factor 10 (FGF-10, Peprotech), penicillin and streptomycin. The cell-containing agar suspension was gently added onto the bottom layer of 0.6% agar in 24-well plate. 0.5 mL of SFM containing the indicated concentration of IWP12 was added on top of the upper cell-containing agar layer, which was replaced every 3 days. On day 14 after seeding, colonies were fixed and stained with 0.005% Crystal violet in 10% formalin diluted in PBS buffer. Colonies were automatically counted with ImageJ software (US National Institutes of Health). All soft agar colony formation assays were repeated three times, and three wells were replicated each time for each IWP12 concentration. The mean and standard deviation were calculated.

### **Sphere formation assay**

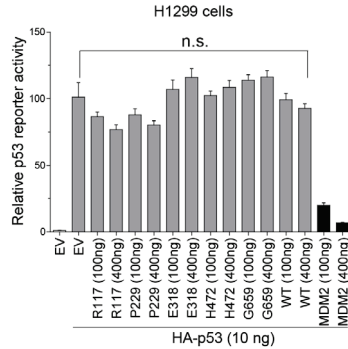
FACS-sorted single cells were grown in SFM medium in 96-wells Ultra-Low Attachment plates (Corning, Amsterdam, The Netherlands) at a density of one viable cell for 14 days. Fresh aliquots of medium was added every 3 days. The assay was performed in 20 replicates for each IWP12 concentration in three independent experiments, and only spheres observed under a microscope were captured through a Zeiss AxioCam ICc3 camera with AxioVision SE64 software. Diameter of spheres was measured with ImageJ software (US National Institutes of Health).



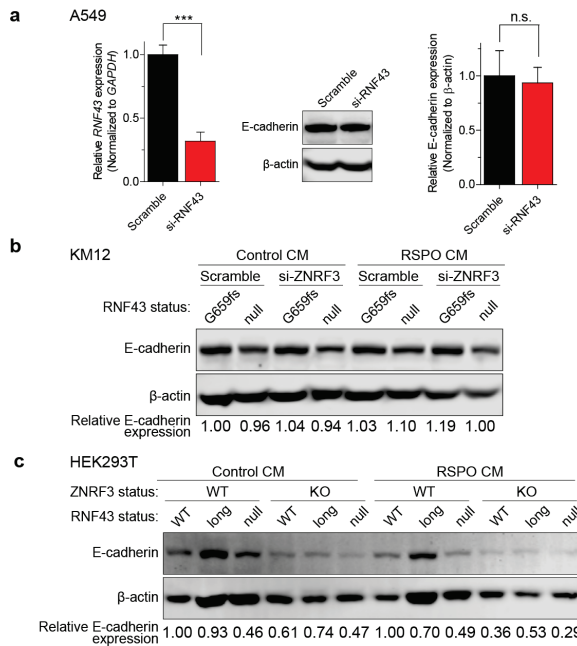
**Supplementary Fig. 1.** Immunofluorescence of transfected RNF43 shows predominant cytoplasmic and endoplasmic reticulum (ER) location. HCT116 cells were transfected with RNF43 expression plasmids containing either a C-terminal FLAG- or HA-tag. Following fixation, cells were stained for the ER-marker Calnexin, and either the FLAG- or HA-tag. Images were generated using a Zeiss LSM700 confocal electroscope (40× objective and 2× digital zoom).



**Supplementary Fig. 2.** Evaluation of TCF7L2 (aka TCF4) interaction and relative contribution to signaling. **A**, Interaction between ectopically expressed C-terminal FLAG-tagged RNF43 mutants and endogenous TCF7L2, assessed by co-immunoprecipitation assay in HCT116 cells. The vertical dashed line indicates removed irrelevant lanes. **B and C**, β-catenin reporter activity was measured in HEK293T or HCT116 cells transiently expressing wild-type (WT), mutant RNF43 or empty vector (EV) as indicated. Cells were treated with Wnt3a-conditioned medium (CM), 5μM of GSK-3 inhibitor SB216763, or a combination thereof (in triplicate, n=3). The relative β-catenin reporter activity was normalized to the value of EV. All data shown in **B** and **C** are represented as mean ± SD. Significance tested using Kruskal-Wallis *H* test with a Dunn's *post hoc* test. Original immunoblot images are shown in Supplementary Fig. 13.

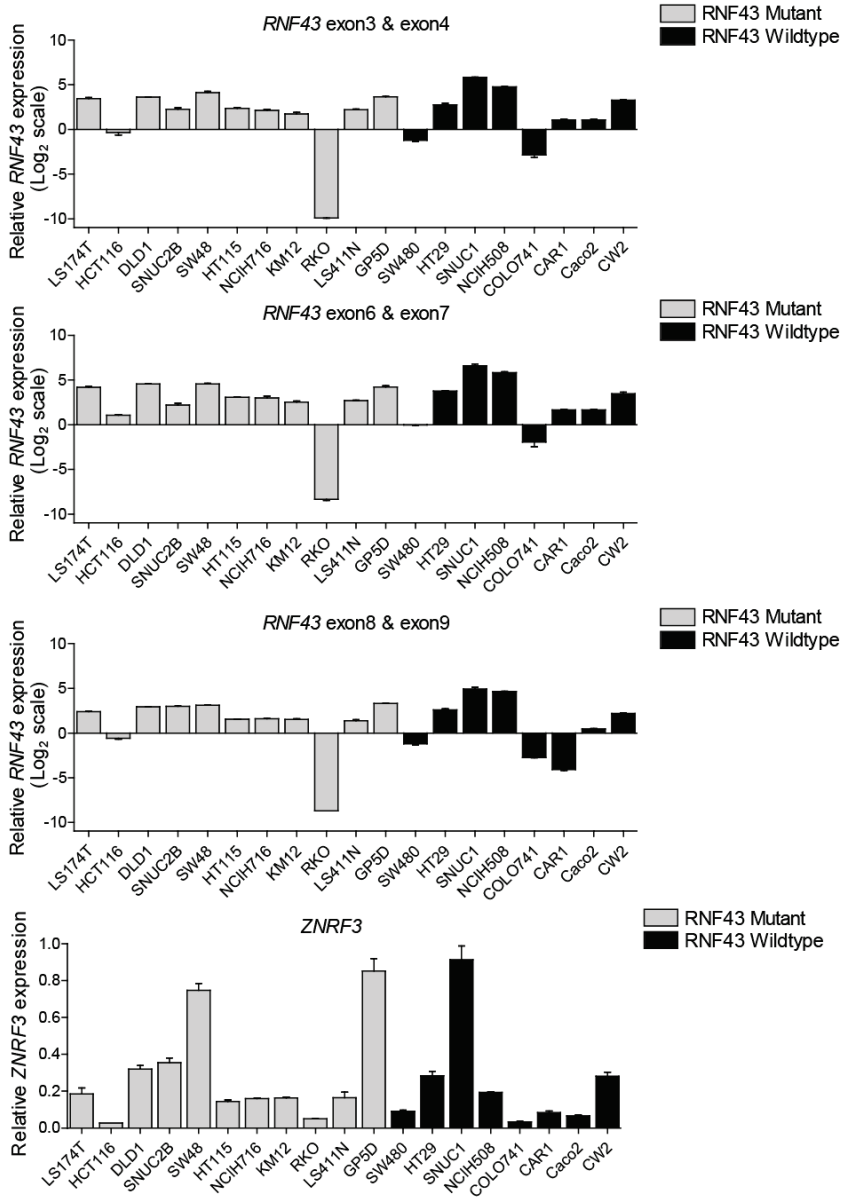


**Supplementary Fig. 3.** Overexpression of RNF43 variants, even at high dosage, may have little or no function on down-regulating p53 signaling. H1299 cells were transiently co-transfected with PG13-luc luciferase reporter plasmid (containing p53-binding consensus sequence) and vectors expressing wild-type (WT), mutant RNF43 or empty vector (EV) or MDM2 expression plasmid (served as a positive functional control), with or without HA-tagged p53 overexpression (in triplicate,  $n=2$ ). The reporter activities shown are relative to EV without p53 overexpression arbitrarily set to 1. The amounts of plasmids used for transfection are indicated (ng DNA). Data are expressed as mean  $\pm$  SD. Significance tested using Kruskal-Wallis  $H$  test with a Dunn's *post hoc* test.

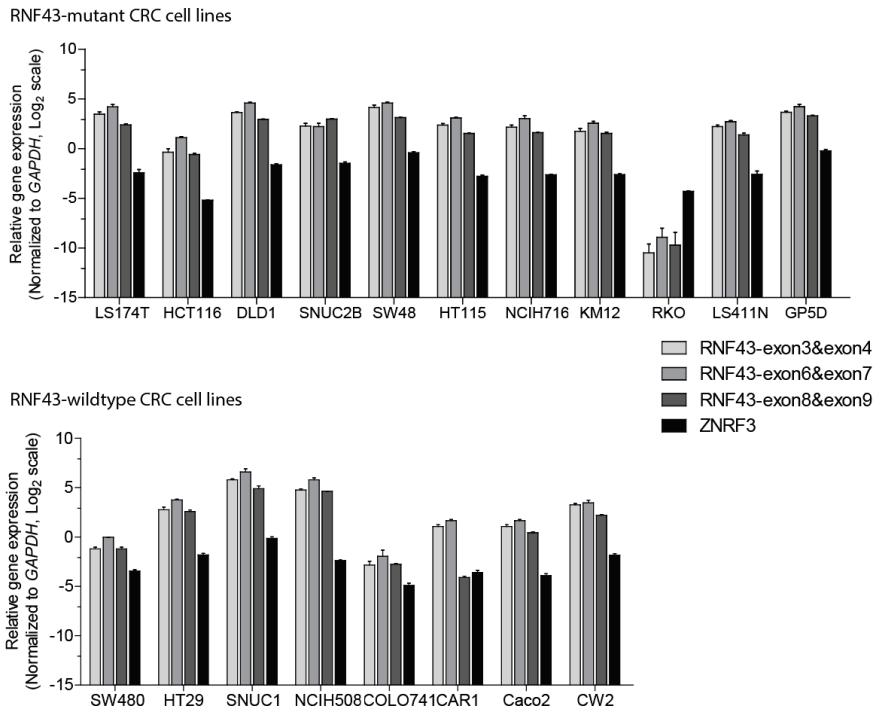


**Supplementary Fig. 4.** RNF43 probably does not exhibit a negative regulatory effect on E-cadherin protein. **a** Left panel, qPCR validation of *RNF43* knockdown efficiency in A549 cells; middle panel, representative immunoblot showing expression levels of E-cadherin protein in A549 cells transfected with/without siRNA targeting *RNF43* for three days; right panel, quantification of the immunoblot. Data shown are presented as mean  $\pm$  SD for three biological replicates, and significance was tested using unpaired Student's  $t$ -test. **b**, **c** representative immunoblot showing expression levels of E-cadherin protein in KM12 cells with/without knockout of RNF43 and HEK293T cells expressing different RNF43/*ZNRF3* mutants as indicated, with/without *ZNRF3*-knockdown or RSPO1-CM treatment. The band densities of E-cadherin were normalized to  $\beta$ -actin (used as a loading control).





**Supplementary Fig. 5.** Bar graphs showing relative mRNA levels of *RNF43* and *ZNRF3* in CRC cell lines analyzed by qRT-PCR using three exon-exon primer sets for *RNF43* (upper three panels) and one primer set for *ZNRF3* (bottom panel). The mRNA expression levels were calculated by normalization to *GAPDH* values. Data shown as mean  $\pm$  SD of three technical replicates from one representative experiment.



**Supplementary Fig. 6.** *RNF43* is more strongly expressed in most colorectal cancer (CRC) cell lines as compared with *ZNRF3*. Bar graphs showing relative mRNA levels of *RNF43* and *ZNRF3* in CRC cell lines analyzed by qRT-PCR using three exon-exon primer sets for *RNF43* and one primer set for *ZNRF3*. The mRNA expression levels were calculated by normalization to GAPDH values (in log<sub>2</sub> scale). Data shown as mean ± SD of three technical replicates from one representative experiment. Data acquired from the same experiment as in Supplementary Fig. 5.

CRISPR/Cas9 genome editing site in Wildtype-RNF43: (Highlighted with red color is the sg-RNA, and TGG is the PAM site)

.....TAGCATGAGTGGTGGCCACCAGCTGCAGCTGGCTGCCCTCTGGCCCTGGCTGCTGATGGCTA  
CCCTGCAGGCAGGCTTTGGACGCACAGGACTGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCA  
GAACAGAAAGCTATTATCAGAGTGATCCCTTGAAAATGGACCCACAGGAAACTGAATCTCACTTT  
GGAAGGTGTGTTTGTGCTGTTGCTGAAATACTCCAGCAGAAGGAAAATTAATGCAGGCAAGTATAA  
CTTTATTATATTTCACTTTCCATCTGTTTGAAATAT.....

exon 2 as the on-target locus.						
Oligo name	Oligo sequence	Score	guide sequence	on-target locus	number of offtarget sites	on-target locus at Reference sequence of the RNF43 CDs (transcript 2 in NCBI)
hRNF43-crF2	CACCgAGGC TTTGGACGC ACAGGAC	80	AGGCTTTGGA CGCACAGGAC TGG	chr17:+ 5841543 3	138 (0 are in genes)	bp 69-88
hRNF43-crR2	AAACGTCT GTGCGTCCA AAGCCTc					

KM12\_RNF43-KO

(Parental) ..CCCTGCAGGCAGGCTTTGGACGCACAG-**ACTGGTACTGGCAGCAGCGG**

.....CCCTGCAGGCAGGCTTTGGACGCACAG-**ACTGGTACTGGCAGCAGCGG** (-1bp)

.....CCCTGCAGGCAGGCTTTGGACGCACAG-**ACTGGTACTGGCAGCAGCGG** (-1bp)

**Supplementary Fig. 7.** Depiction of CRISPR/Cas9 induced mutations observed in the RNF43-knockout clone of KM12. The sg-RNA used and the PAM site are shown as indicated.

**sgRNAs used for generating the HEK293T clones**

RNF43-null (targeting RNF43-exon2) **sgRNA:** AGGCTTTGGACGCACAGGAC **PAM:** TGG  
 RNF43-long (targeting RNF43-exon9) **sgRNA:** ACCCACAGAGGAAAAGGCGG **PAM:** GGG  
 ZNRF3-KO (targeting ZNRF3-exon2) **sgRNA:** CATGGTTTCGGGTCCAATTC **PAM:** TGG

**HEK293T knockout clones:****RNF43-WT/ZNRF3-WT (#1): Clone B12**

RNF43 and ZNRF3 sequences: wild-type

**RNF43-WT/ZNRF3-WT (#2): Clone B18**

RNF43 and ZNRF3 sequences: wild-type

**RNF43-WT/ZNRF3-WT (#3): Clone G6**

RNF43 and ZNRF3 sequences: wild-type

**RNF43-long/ZNRF3-WT (#1): Clone A13**

RNF43 sequence:

CCACACAGAGGAAAAGCGGGGGTCCCTCCGAGCCACCCCTGGCTCTCGGCCCCAGGAT (parental)  
 CCCACACAGAG.....AGCCACCCCTGGCTCTCGGCCCCAGGAT (-23bp) p.K655Afs\*82  
 CCCACACAGAG.....CCACCCCTGGCTCTCGGCCCCAGGAT (-25bp) p.R654Sfs\*36

ZNRF3 sequence: Wild-type

**RNF43-long/ZNRF3-WT (#2): Clone E3**

RNF43 sequence:

CCACACAGAGGAAAAGCGGGGGTCCCTCCGAGCCACCCCTGGCTCTCGGCCCCAGGAT (parental)  
 CCCACACAGAGAAAAG.....AGCCACCCCTGGCTCTCGGCCCCAGGAT (-17bp) p.R657Afs\*82  
 CCCACACAGAGGAAA.....CCCCCCCCCTGGCTCTCGGCCCCAGGAT (-19bp) p.R656Tfs\*36

ZNRF3 sequence: Wild-type

**RNF43-long/ZNRF3-WT (#3): Clone E8**

RNF43 sequence:

CCACACAGAGGAAAAGCGGGGGTCCCTCCGAGCCACCCCTGGCTCTCGGCCCCAGGAT (parental)  
 CCCACACAGAGGAAA.....GCCACCCCTGGCTCTCGGCCCCAGGAT (-19bp) p.R656Sfs\*36  
 CCCACACAGAGGAAA.....CGGGGGTCCCTCCGAGCCACCCCTGGCTCTCGGCCCCAGGAT (-2bp) p.R656Tfs\*88

ZNRF3 sequence: Wild-type

**RNF43-null/ZNRF3-WT (#1): Clone C4**

RNF43 sequence:

TTTGGACGCACAGGACTGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACAGAAA (parental)  
 TTTGGACGCAC.....TGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACAGAAA (-5bp)  
 TTTGG.....TACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACAGAAA (-14bp)

ZNRF3 sequence: wild-type

**RNF43-null/ZNRF3-WT (#2): Clone C5**

RNF43 sequence:

TTTGGACGCACAGGACTGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACAGAAA (parental)  
 TTTGG.....GACTGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACAGAAA (-8bp)  
 TTTGGACGCACAG..ACTGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACAGAAA (-1bp)  
 TTTGGACGCACAGNNCTNTGACTGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAG (+8bp)

ZNRF3 sequence: wild-type

**RNF43-null/ZNRF3-WT (#3): Clone C16**

RNF43 sequence:

TTTGGACGCACAGGACTGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACAGAAA (parental)  
 TTTGGACGCAC..GACTGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACAGAAA (-2bp)  
 TTTGGACGCACAGTACCAGACTGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAAC (+5bp)

ZNRF3 sequence: wild-type

**RNF43-WT/ZNRF3-KO (#1): Clone D25**

RNF43 sequence: wild-type

ZNRF3 sequence:

TGGTGAAGCTGGAACAGCCAGAAATTGGACCCGAAACCATGCCTCACTGTCCTAGGC (parental)  
 TGGTGA.....CCCGAAACCATGCCTCACTGTCCTAGGC (-22bp)  
 TGGTGAAGCTGGAACAGCCAGAACTTGGACCCGAAACCATGCCTCACTGTCCTAGG (+1bp)

**RNF43-WT/ZNRF3-KO (#2): Clone D47**

RNF43 sequence: wild-type

ZNRF3 sequence:

TGGTGAAGCTGGAACAGCCAGAAATTGGACCCGAAACCATGCCTCACTGTCCTAGGC (parental)  
 TGGTGAAGCTGGAACAGCCAGAA.....ACCATGCCTCACTGTCCTAGGC (-11bp)  
 TGGTGAAGCTGGAACAGCCAGAAATTGGACCCGAAACCATGCCTCACTGTCCTAGG (+1bp)

**RNF43-WT/ZNRF3-KO (#3): Clone H17**

RNF43 sequence: wild-type

ZNRF3 sequence:

TGGTGAAGCTGGAACAGCCAGAAATTGGACCCGAAACCATGCCTCACTGTCCTAGGC (parental)  
 TGGTGA.....CCCGAAACCATGCCTCACTGTCCTAGGC (-22bp)  
 TGGTGAAGCTGGAACAGCCAGAACTTGGACCCGAAACCATGCCTCACTGTCCTAGG (+1bp)

**Supplementary Fig. 8.** Depiction of CRISPR/Cas9 induced mutations observed in RNF43-knockout clones of HEK293T. The sg-RNA used and the PAM site are shown as indicated.

(figure continued on next page)

**RNF43-long/ZNRF3-KO (#1): Clone A8**

RNF43 sequence:  
 CCCACAGAGGAAAAGCGGGGGTCCCTCCGAGCCACCCCTGGCTCTCGGCCCCAGGAT (parental)  
 CCCACAGAGGA.....TCCCTCCGAGCCACCCCTGGCTCTCGGCCCCAGGAT (-13bp) p.K655Ifs\*39  
 CCCACAGAGG.....GCGGGGGGTCCCTCCGAGCCACCCCTGGCTCTCGGCCCCAGGAT (-5bp) p.K655Afs\*88  
 ZNRF3 sequence:  
 AAGCTGGAAACGCCAGAATTGGACCCGAAACCATGCCTCACTGTCCTAGGCA (parental)  
 AAGCTGGAAACGCCAGAAATTGGACCCGAAACCATGCCTCACTGTCCTAGGC (+1bp)  
 AAGCTGGAAACGCCAGAA.....ACCATGCCTCACTGTCCTAGGCA (-11bp)

**RNF43-long/ZNRF3-KO (#2): Clone F16**

RNF43 sequence:  
 CCCACAGAGGAAAAGCGGGGGTCCCTCCGAGCCACCCCTGGCTCTCGGCCCCAGGAT (parental)  
 CCCACAGAGGAA.....CGGGGGGTCCCTCCGAGCCACCCCTGGCTCTCGGCCCCAGGAT (-4bp) p.K655Nfs\*42  
 CCCACAGAGGAAAAGGC.....CACCCCTGGCTCTCGGCCCCAGGAT (-19bp) p.K657Pfs\*35  
 ZNRF3 sequence:  
 TGCTGAAGCTGGAACAGCCAGAATTGGACCCGAAACCATGCCTCACTGTCCTAGGC (parental)  
 TGCTGA.....CCCGAAACCATGCCTCACTGTCCTAGGC (-22bp)  
 TGCTGAAGCTGGAACAGCCAGAACTGGACCCGAAACCATGCCTCACTGTCCTAGG (+1bp)

**RNF43-long/ZNRF3-KO (#3): Clone F23**

RNF43 sequence:  
 CCCACAGAGGAAAAGCGGGGGTCCCTCCGAGCCACCCCTGGCTCTCGGCCCCAGGAT (parental)  
 CCCACAGAGGAA.....CGGGGGGTCCCTCCGAGCCACCCCTGGCTCTCGGCCCCAGGAT (-4bp) p.K655Nfs\*42  
 CCCACAGAGGA.....GCCACCCCTGGCTCTCGGCCCCAGGAT (-22bp) p.K655Sfs\*36  
 ZNRF3 sequence:  
 TGCTGAAGCTGGAACAGCCAGAATTGGACCCGAAACCATGCCTCACTGTCCTAGGC (parental)  
 TGCTGA.....CCCGAAACCATGCCTCACTGTCCTAGGC (-22bp)  
 TGCTGAAGCTGGAACAGCCAGAACTGGACCCGAAACCATGCCTCACTGTCCTAGG (+1bp)

**RNF43-null/ZNRF3-KO (#1): Clone B28**

RNF43 sequence:  
 GGCCTTTGGACGCACAGGACTGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACAG (parental)  
 GGCCTTTGGACGCACAG..ACTGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACAG (-1bp)  
 GGCCTTTGGACGCACAG.....GTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACAG (-5bp)  
 GGCCTTTGGACGCACAGAGACTGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACA (+1bp)  
 ZNRF3 sequence:  
 GGAGTGGTGAAGCTGGAACAGCCAGAATTGGACCCGAAACCATGCCTCACTGTCCTAGG (parental)  
 GGAGTGGTGAAGCTGGAACAGCC.....CGAAACCATGCCTCACTGTCCTAGG (-11 bp)  
 GGAGTGGTGAAGCTGGAACAGCC.....CGAAACCATGCCTCACTGTCCTAGG (-11 bp)

**RNF43-null/ZNRF3-KO (#2): Clone C54**

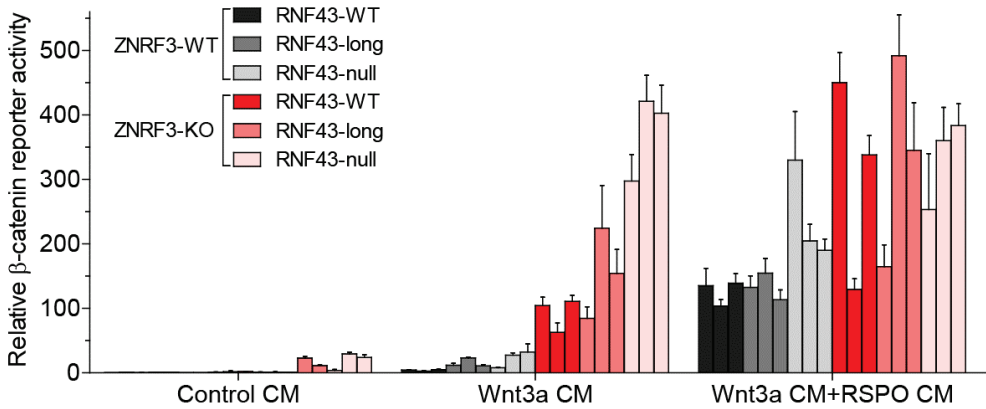
RNF43 sequence:  
 GGCCTTTGGACGCACAGGACTGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACAG (parental)  
 GGCCTTTGGAC.....AGSACTGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACAG (-4bp)  
 GGCT.....GGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACAG (-16bp)  
 ZNRF3 sequence:  
 TGCTGAAGCTGGAACAGCCAGAATTGGACCCGAAACCATGCCTCACTGTCCTAGG (parental)  
 TGCTGAAGCTGGAACAGCC.....GACCCGAAACCATGCCTCACTGTCCTAGG (-7bp)  
 TGCTGAAGCTGGAACAGCC.....GACCCGAAACCATGCCTCACTGTCCTAGG (-7bp)

**RNF43-null/ZNRF3-KO (#3): Clone I49**

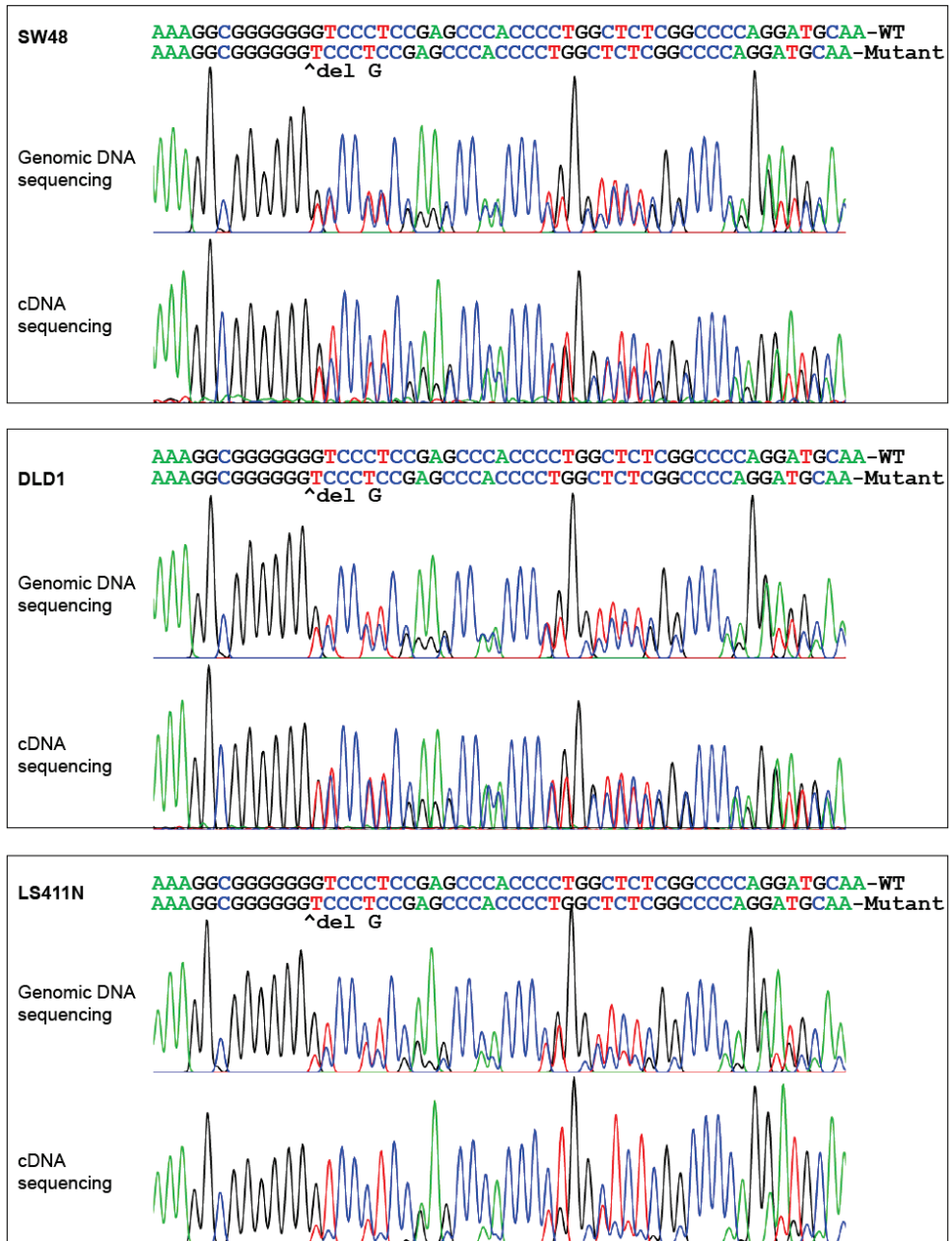
RNF43 sequence:  
 GGCCTTTGGACGCACAGGACTGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACAG (parental)  
 GGCCTTTGGACGCACAG..ACTGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACAG (-1bp)  
 GGCCTTTGGACGCACAG..ACTGGTACTGGCAGCAGCGGTGGAGTCTGAAAGATCAGCAGAACAG (-1bp)  
 ZNRF3 sequence:  
 TGCTGAAGCTGGAACAGCCAGAATTGGACCCGAAACCATGCCTCACTGTCCTAGG (parental)  
 TGCTGA.....CCCGAAACCATGCCTCACTGTCCTAGG (-22bp)  
 TGCTGAAGCTGGAACAGCCAGAACTGGACCCGAAACCATGCCTCACTGTCCTAGG (+1bp)

(figure continued from previous page)

**Supplementary Fig. 8.** Depiction of CRISPR/Cas9 induced mutations observed in RNF43-knockout clones of HEK293T. The sg-RNA used and the PAM site are shown as indicated.

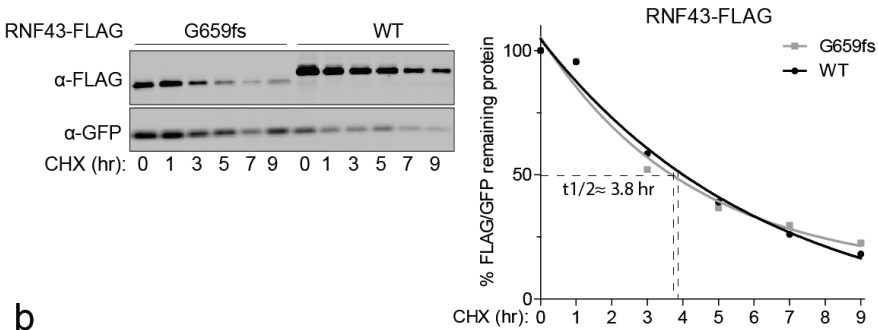


**Supplementary Fig. 9.**  $\beta$ -catenin reporter activity was measured in ZNRF3 wild-type (WT) and knockout (KO) HEK293T cells simultaneously expressing CRISPR-Cas9 induced null, long mutant (R654fs/K655fs/R656fs/R657fs) or wild-type RNF43, here shown for each clone separately. Data acquired from the same experiment as in Fig. 5a, b.

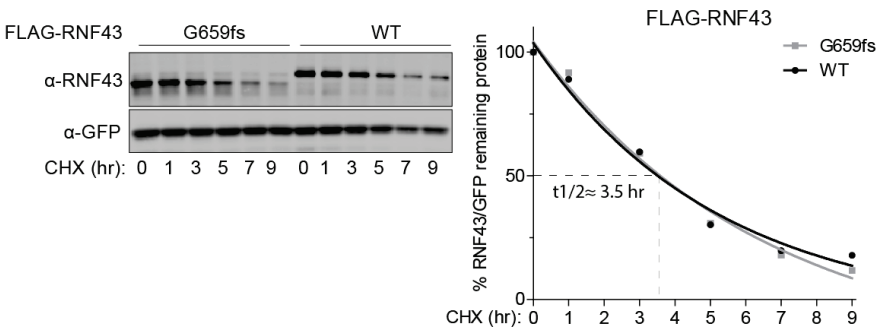


**Supplementary Fig. 10.** The *RNF43* mutant mRNA encoding the p.G659fs variant is not targeted and suppressed by nonsense-mediated decay. Representative chromatograms showing Sanger sequencing of *RNF43* amplicons spanning the p.G659fs (c.1976delG) mutation obtained from genomic DNA and cDNA of colorectal cancer cells (SW48, DLD1, and LS411N) harboring a heterozygous p.G659fs *RNF43* mutation. No reduction in presence of the *RNF43* c.1976delG indel was observed in cDNA sequencing results compared to genomic DNA sequencing data (the position of the deleted nucleotide is shown above the sequencing data). Based on peak height, LS411N appears to carry two wild-type and one mutant chromosome.

a



b



**Supplementary Fig. 11.** Cycloheximide chase assay showing that wild-type and G659fs RNF43 variants have a similar protein half-life. HEK293T cells were co-transfected with the RNF43 expression vectors and GFP plasmid (as a transfection control), followed by cycloheximide (CHX) treatment for the indicated times. **A**, left panel, immunoblot showing the protein expression of C-terminal FLAG-tagged RNF43 (RNF43-FLAG) vectors using anti-FLAG antibody. **B**, left panel, the protein level of N-terminal FLAG-tagged RNF43 (FLAG-RNF43) vectors was assessed by immunoblotting using anti-RNF43 antibody. **A and B**, Right panels, for half-life ( $t_{1/2}$ ) determination of the RNF43 variants we used a nonlinear regression analysis (one phase exponential decay). Band intensities of RNF43 expression after cycloheximide addition for the indicated time points were normalized to GFP values, and then represented as a percentage to the value at time zero ( $t=0$  was taken as 100%). Original immunoblot images are provided in Source Data File.

### ConSurf results based on human RNF43



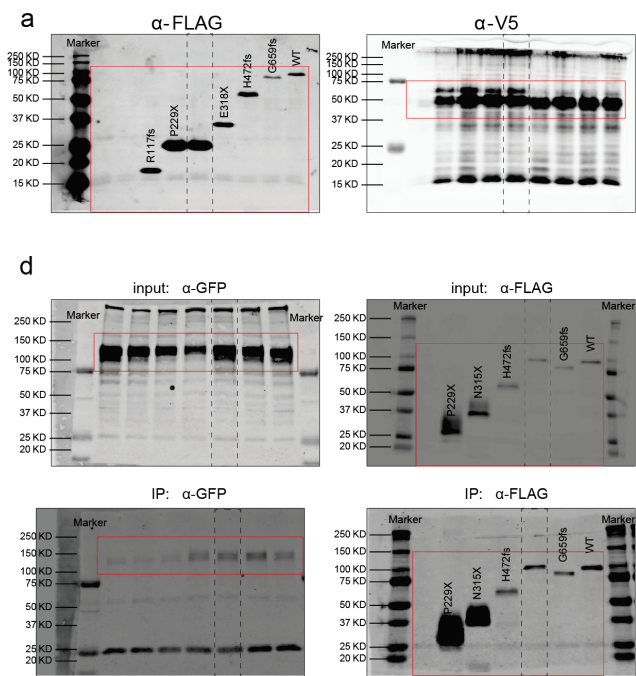
The conservation scale:



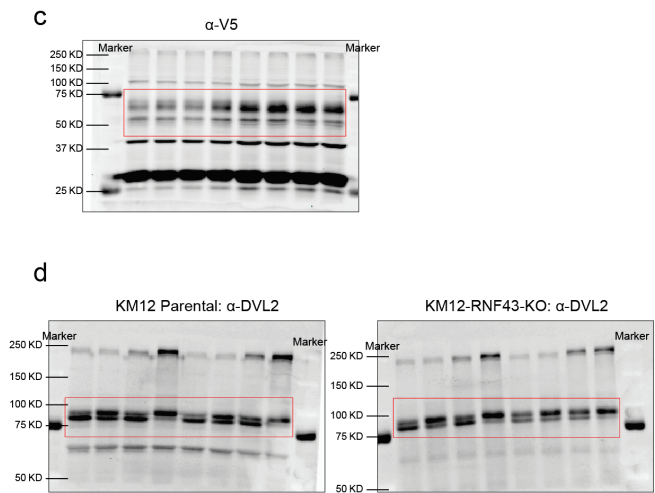
**Supplementary Fig. 12.** Analysis of Evolutionary Conserved RNF43 Residues using the ConSurf tool. Human RNF43 was used as input with standard settings and homology between 70–99%.



Original IB images for Fig. 3



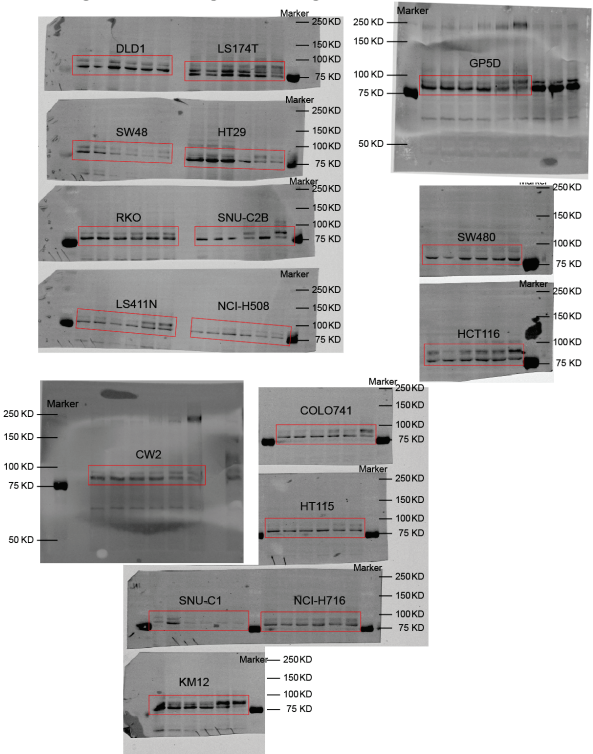
Original IB images for Fig. 4



**Supplementary Fig. 13.** Original immunoblot images presented in this study.

*(figure continued on next page)*

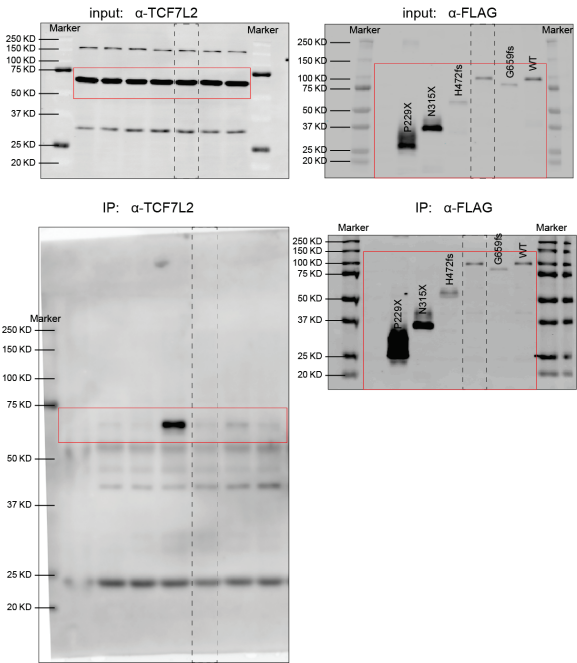
Original IB images for Fig. 6



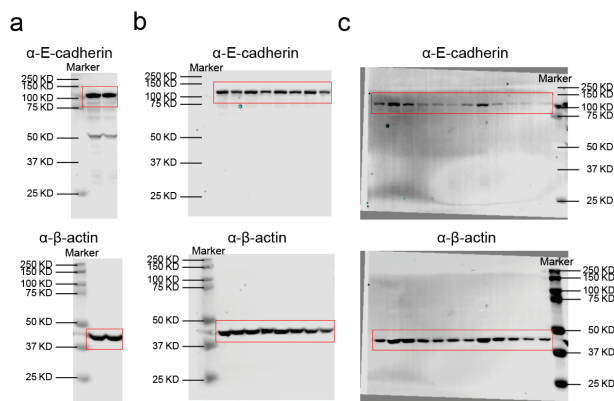
Supplementary Fig. 13. Original immunoblot images presented in this study.

(figure continued on next page)

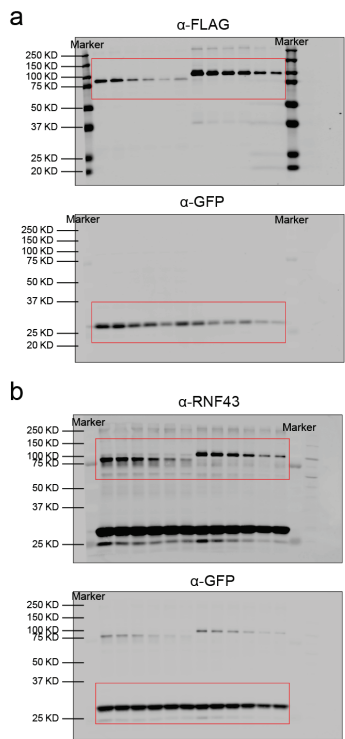
Original IB images for Supplementary Fig. 2a



Original IB image for Supplementary Fig. 4



Original IB image for Supplementary Fig. 11



**Supplementary Fig. 13.** (figure continued from previous page) Original immunoblot images presented in this study.

### Supplementary Table 1.

*RNF43* mutations and MSI status in various tumor types. Only tumors with one *RNF43* truncating mutation were included. Data is available at:

[https://static-content.springer.com/esm/art%3A10.1038%2Fs41388-020-1232-5/MediaObjects/41388\\_2020\\_1232\\_MOESM3\\_ESM.pdf](https://static-content.springer.com/esm/art%3A10.1038%2Fs41388-020-1232-5/MediaObjects/41388_2020_1232_MOESM3_ESM.pdf)

**Supplementary Table 1a.** *RNF43* mutations and MSI status reported in MSI-colorectal cancers. Total 344 MSI colorectal cancer samples. Total 88 truncating *RNF43* mutations, including 81 frameshift mutations and 7 nonsense mutations.

**Supplementary Table 1b.** *RNF43* mutations and MSI status reported in MSS-colorectal cancers. Total 1900 MSS colorectal cancer samples. Total 35 truncating *RNF43* mutations, including 19 frameshift mutations and 16 nonsense mutations.

**Supplementary Table 1c.** *RNF43* mutations and MSI status reported in MSI-colorectal serrated neoplasia. Total 15 MSI colorectal serrated neoplasia samples. Total 10 truncating *RNF43* mutations, including 9 frameshift mutations and 1 nonsense mutations.

**Supplementary Table 1d.** *RNF43* mutations and MSI status reported in MSS-colorectal serrated neoplasia. Total 220 MSS colorectal serrated neoplasia samples. Total 43 truncating *RNF43* mutations, including 29 frameshift mutations and 14 nonsense mutations.

**Supplementary Table 1e.** *RNF43* mutations and MSI status reported in MSI-gastric cancer. Total 78 MSI-gastric cancer samples. Total 35 truncating *RNF43* mutations, including 35 frameshift mutations and 0 nonsense mutations.

**Supplementary Table 1f.** *RNF43* mutations and MSI status reported in MSS-gastric cancer. Total 435 MSS/MSI-L gastric cancer samples. Total 11 truncating *RNF43* mutations, including 6 frameshift mutations and 5 nonsense mutations.

**Supplementary Table 1g.** *RNF43* mutations and MSI status reported in MSI-endometrial cancer. Total 110 MSI-H endometrial cancer samples. Total 30 truncating *RNF43* mutations, including 29 frameshift mutations and 1 nonsense mutations.

**Supplementary Table 1h.** *RNF43* mutations and MSI status reported in MSS-endometrial cancer. Total 403 MSS/MSI-L endometrial cancer samples. Total 6 truncating *RNF43* mutations, including 4 frameshift mutations and 2 nonsense mutations.

**Supplementary Table 1i.** *RNF43* mutations reported in pancreatic cancer. Total 749 pancreatic cancer samples. Total 24 truncating *RNF43* mutations, including 14 frameshift mutations and 10 nonsense mutations.

**Supplementary Table 2.** Cell lines used in this study. Important mutations and microsatellite instability (MSI) status are depicted. MSI-H, microsatellite instability-high; MSI-L, microsatellite instability-low; MSS, microsatellite stable; homo., homozygous; het., heterozygous.

Cell line name	RNF43	ZNRF3	APC	CTNNB1	AXIN1/2	BRAF	KRAS	PIK3CA	TP53	MSI status
LS174T	p.K108E het.+p.R389H het.+p.G659fs*41 het.			p.S45F homo.	AXIN2 p.R834W het.	p.D211G het.	p.G12D het.	p.H1047R het.		MSI-H
HCT116	p.R117fs*41 homo.			p.S45del het.	AXIN2 p.G665s*24 het.		p.G13D het.	p.H1047R het.		MSI-H
DLD1	p.G659fs*41 p.L214M het.	het.+	p.I11417fs*2 het.+p.R2166* het.		AXIN1 p.L396M het.		p.G13D het.	p.D649N+p.E545K het.	p.S241F het.	MSI-H
SNU-C2B	p.C275fs*143 het.	p.V138fs*65 het.+p.R207W het.					p.G12D het.	p.D725G het.	p.R273H+p.R273C+p.S185S het.	MSI-H
SW48	p.V299fs*143 p.G659fs*41 het.	het.+		p.S33Y het.				p.G914R het.		MSI-H
HT-115	p.G447E het.		p.S1196* het.+ p.T1556fs*3 het.					p.R88Q+p.R770Q+p.R770Q het.	p.R213* het.	MSI-L
NCI-H716	p.H472fs*30 homo.						p.R97I het.		p.E224D homo.	MSS
KM12	p.G659fs*41 homo.		p.G471E het.+p.N1818fs*2 het.		AXIN1 p.G265fs*149 het.+ p.E640fs*65 het.	p.A404fs*9+p.A712 T het.			p.V73fs*50+p.H179R het.	MSI-H
RKO	p.G659fs*41 homo.	p.V249fs*117 het.				p.V600E het.		p.H1047R het.		MSI-H
LS-411N	p.G659fs*41 het.	p.A280T het.	p.Q789* het.+ p.T1556fs*3 het.			p.V600E het.			p.Y126* homo.	MSI-H
GP5D	p.S771T het.		p.T1445fs*27 het.		AXIN2 p.G665s*24 het.	p.T529A het.	p.G12D het.	p.H1047L het.		MSI-H
SW480			p.Q1338* homo.				p.G12V homo.		p.R273H+p.P309S homo.	MSS
HT-29			p.E853* het.+ p.T1556fs*3 het.			p.V600E het.+p.T119S het.		p.P449T het.	p.R273H homo.	MSS
SNU-C1					AXIN2 p.A69T het.				p.S166* homo.	MSS
NCI-H608						p.G596R het.		p.E545K het.	p.R273H homo.	MSS
COL0741						p.V600E het.			p.P322fs*24 het.	MSS
CAR1									p.V272M homo.	MSS
Caco2			p.Q1367* het.			p.V600E het.			p.E204* het.	MSS
CW2			p.R302* het.+ p.S1466fs*3 het.	p.R682Q het.	AXIN1 p.P37fs*47 het.+ AXIN2 p.N660fs*24 het.		p.P140H het.	p.P283S het.		MSI-H
LOVO			p.R1114* + p.M1431fs*42				p.G13D het.			MSI-H

**Supplementary Table 3.** Cancer genomic datasets for showing  $\beta$ -catenin enhancing mutations in colorectal cancer samples (data retrieved from cBioPortal). Data is available at:

[https://static-content.springer.com/esm/art%3A10.1038%2Fs41388-020-1232-5/MediaObjects/41388\\_2020\\_1232\\_MOESM3\\_ESM.pdf](https://static-content.springer.com/esm/art%3A10.1038%2Fs41388-020-1232-5/MediaObjects/41388_2020_1232_MOESM3_ESM.pdf)

**Supplementary Table 4.** Primers used for qRT-PCR.

Gene name	Forward primer (5'-3')	Reverse primer (5'-3')
<i>GAPDH</i>	GTCTCCTCTGACTTCAACAGCG	ACCACCCTGTTGCTGTAGCCAA
<i>RNF43</i> (exon3 & exon4)	ACATCACTGAGGATCGAGCT	TCCTTCAGCTCAATCCTCACA
<i>RNF43</i> (exon6 & exon7)	CTGTGTGTGCCATCTGTCTG	GTCCGATGCTGATGTAACCA
<i>RNF43</i> (exon8 & exon9)	CCTTCTGAATGGAGTTCTGAC	GCTAGGCCTGAACATCTCACA
<i>ZNRF3</i>	TGCTGTCAGGGCCAACTT	CAGTTCCCAATTTCCAGGTAAG
<i>AXIN2</i>	TATCCAGTGATGCGCTGACG	TTACTGCCCACACGATAAGG

## Acknowledgements

We wish to thank Dr. Louis Vermeulen (AMC, Amsterdam, The Netherlands) for providing us with several of the cell lines used in this study. We are grateful to Dr. Erik A.C. Wiemer and Dr. Robbert J. Rottier (Erasmus MC, Rotterdam, The Netherlands) for providing us, respectively with H1299 and A549 cell lines. We gratefully acknowledge Dr. Aart G. Jochemsen (LUMC, Leiden, The Netherlands) for helpful discussion on p53-related works and providing us the p53 related plasmids. We thank Dr. Markus Gerhard (Technische Universität München, Munich, Germany) for kindly providing HA-tagged RNF43 plasmid. We thank Natascha Nieuwenhuijze (Erasmus MC, Rotterdam, The Netherlands) for help with confocal microscopy. We would like to thank Dr. Feng Cong (Novartis Institutes for Biomedical Research, Cambridge, MA, USA) for kindly giving us positive responses and the detailed protocol of surface SNAP-labeling experiments. This work was supported by the China Scholarship Council PhD fellowship (grant numbers 201408060053 to S.L., and 201408220029 to P.L.); and an Erasmus MC Grant; M.M. received financial support from National Science Centre Poland (UMO-2015/16/T/NZ4/00176).

## References

1. H. X. Hao *et al.*, ZNRF3 promotes Wnt receptor turnover in an R-spondin-sensitive manner. *Nature* **485**, 195-200 (2012).
2. B. K. Koo *et al.*, Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors. *Nature* **488**, 665-669 (2012).
3. H. X. Hao, X. Jiang, F. Cong, Control of Wnt receptor turnover by R-spondin-ZNRF3/RNF43 signaling module and its dysregulation in cancer. *Cancers (Basel)* **8**, 54 (2016).
4. A. Loregger *et al.*, The E3 ligase RNF43 inhibits Wnt signaling downstream of mutated  $\beta$ -catenin by sequestering TCF4 to the nuclear membrane. *Science Signaling* **8**, ra90 (2015).
5. M. Giannakis *et al.*, *RNF43* is frequently mutated in colorectal and endometrial cancers. *Nat. Genet.* **46**, 1264-1266 (2014).
6. B. H. Min *et al.*, Dysregulated Wnt signalling and recurrent mutations of the tumour suppressor RNF43 in early gastric carcinogenesis. *J. Pathol.* **240**, 304-314 (2016).
7. G. L. Ryland *et al.*, *RNF43* is a tumour suppressor gene mutated in mucinous tumours of the ovary. *J. Pathol.* **229**, 469-476 (2013).
8. G. L. Ryland *et al.*, Mutational landscape of mucinous ovarian carcinoma and its neoplastic precursors. *Genome Med.* **7**, 87 (2015).
9. K. Wang *et al.*, Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer. *Nat. Genet.* **46**, 573-582 (2014).
10. Y. Jiao *et al.*, Whole-exome sequencing of pancreatic neoplasms with acinar differentiation. *J. Pathol.* **232**, 428-435 (2014).
11. N. Waddell *et al.*, Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature* **518**, 495-501 (2015).
12. Y. Zou *et al.*, RNF43 mutations are recurrent in Chinese patients with mucinous ovarian carcinoma but absent in other subtypes of ovarian cancer. *Gene* **531**, 112-116 (2013).
13. T. Eto *et al.*, Impact of loss-of-function mutations at the RNF43 locus on colorectal cancer development and progression. *J. Pathol.* **245**, 445-455 (2018).
14. H. H. N. Yan *et al.*, RNF43 germline and somatic mutation in serrated neoplasia pathway and its association with BRAF mutation. *Gut* **66**, 1645-1656 (2017).
15. X. Jiang *et al.*, Inactivating mutations of *RNF43* confer Wnt dependency in pancreatic ductal adenocarcinoma. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 12649-12654 (2013).
16. Z. Steinhart *et al.*, Genome-wide CRISPR screens reveal a Wnt-FZD5 signaling circuit as a druggable vulnerability of *RNF43*-mutant pancreatic tumors. *Nat. Med.* **23**, 60-68 (2017).
17. B. K. Koo, J. H. van Es, M. van den Born, H. Clevers, Porcupine inhibitor suppresses paracrine Wnt-driven growth of *Rnf43*/*Znrf3*-mutant neoplasia. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7548-7550 (2015).
18. N. Cancer Genome Atlas Research, Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* **513**, 202-209 (2014).
19. M. Giannakis *et al.*, Genomic correlates of immune-cell infiltrates in colorectal carcinoma. *Cell Rep.* **15**, 857-865 (2016).
20. R. Yaeger *et al.*, Clinical sequencing defines the genomic landscape of metastatic colorectal cancer. *Cancer Cell* **33**, 125-136 (2018).
21. S. Seshagiri *et al.*, Recurrent R-spondin fusions in colon cancer. *Nature* **488**, 660-664 (2012).
22. T. Hashimoto *et al.*, WNT pathway gene mutations are associated with the presence of dysplasia in colorectal sessile serrated adenoma/polyps. *Am. J. Surg. Pathol.* **41**, 1188-1197 (2017).
23. S. Sekine *et al.*, Frequent *PTPRK-RSPO3* fusions and *RNF43* mutations in colorectal traditional serrated adenoma. *J. Pathol.* **239**, 133-138 (2016).

24. T. E. Soumerai *et al.*, Clinical utility of prospective molecular characterization in advanced endometrial cancer *Clin. Cancer Res.* **24**, 5939-5947 (2018).
25. W. J. Gibson *et al.*, The genomic landscape and evolution of endometrial carcinoma progression and abdominopelvic metastasis. *Nat. Genet.* **48**, 848-855 (2016).
26. D. A. Levine *et al.*, Integrated genomic characterization of endometrial carcinoma. *Nature* **497**, 67-73 (2013).
27. K. Wang *et al.*, Exome sequencing identifies frequent mutation of *ARID1A* in molecular subtypes of gastric cancer. *Nat. Genet.* **43**, 1219-1223 (2011).
28. N. Cancer Genome Atlas, Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330-337 (2012).
29. X. Jiang, O. Charlat, R. Zamponi, Y. Yang, F. Cong, Dishevelled promotes Wnt receptor degradation through recruitment of ZNRF3/RNF43 E3 ubiquitin ligases. *Mol. Cell* **58**, 522-533 (2015).
30. A. Vanderwalde, D. Spetzler, N. Xiao, Z. Gatalica, J. Marshall, Microsatellite instability status determined by next-generation sequencing and compared with PD-L1 and tumor mutational burden in 11,348 patients. *Cancer Medicine* **7**, 746-756 (2018).
31. P. Bailey *et al.*, Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* **531**, 47-52 (2016).
32. A. K. Witkiewicz *et al.*, Whole-exome sequencing of pancreatic cancer defines genetic diversity and therapeutic targets. *Nat. Commun.* **6**, 6744 (2015).
33. J. Wu *et al.*, Whole-exome sequencing of neoplastic cysts of the pancreas reveals recurrent mutations in components of ubiquitin-dependent pathways. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 21188-21193 (2011).
34. H. Sakamoto *et al.*, Clinicopathological significance of somatic *RNF43* mutation and aberrant expression of ring finger protein 43 in intraductal papillary mucinous neoplasms of the pancreas. *Mod. Pathol.* **28**, 261-267 (2015).
35. J. H. Tsai *et al.*, *RNF43* mutation frequently occurs with *GNAS* mutation and mucin hypersecretion in intraductal papillary neoplasms of the bile duct. *Histopathology* **70**, 756-765 (2017).
36. K. Shinada *et al.*, RNF43 interacts with NEDL1 and regulates p53-mediated transcription. *Biochem. Biophys. Res. Commun.* **404**, 143-147 (2011).
37. H. Nailwal, S. Sharma, A. K. Mayank, S. K. Lal, The nucleoprotein of influenza A virus induces p53 signaling and apoptosis via attenuation of host ubiquitin ligase RNF43. *Cell death and disease* **6**, e1768 (2015).
38. L. Niu *et al.*, RNF43 inhibits cancer cell proliferation and could be a potential prognostic factor for human gastric carcinoma. *Cell. Physiol. Biochem.* **36**, 1835-1846 (2015).
39. H. Xie *et al.*, Association of RNF43 with cell cycle proteins involved in p53 pathway. *Int. J. Clin. Exp. Pathol.* **8**, 14995-15000 (2015).
40. C. Xing *et al.*, Reversing effect of ring finger protein 43 inhibition on malignant phenotypes of human hepatocellular carcinoma. *Mol. Cancer Ther.* **12**, 94-103 (2013).
41. Y. Haupt, R. Maya, A. Kazaz, M. Oren, Mdm2 promotes the rapid degradation of p53. *Nature* **387**, 296-299 (1997).
42. Y. Zhang *et al.*, RNF43 ubiquitinates and degrades phosphorylated E-cadherin by c-Src to facilitate epithelial-mesenchymal transition in lung adenocarcinoma. *BMC Cancer* **19**, 670 (2019).
43. M. Zebisch *et al.*, Structural and molecular basis of ZNRF3/RNF43 transmembrane ubiquitin ligase inhibition by the Wnt agonist R-spondin. *Nat. Commun.* **4**, 2787 (2013).
44. Y. Xie *et al.*, Interaction with both ZNRF3 and LGR4 is required for the signalling activity of R-spondin. *EMBO Reports* **14**, 1120-1126 (2013).
45. K. S. Carmon, X. Gong, J. Yi, A. Thomas, Q. Liu, RSPO-LGR4 functions via IQGAP1 to potentiate Wnt signaling. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E1221-1229 (2014).



46. K. S. Carmon, X. Gong, Q. Lin, A. Thomas, Q. Liu, R-spondins function as ligands of the orphan receptors LGR4 and LGR5 to regulate Wnt/ $\beta$ -catenin signaling. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 11452-11457 (2011).
47. J. M. Gonzalez-Sancho, K. R. Brennan, L. A. Castelo-Soccio, A. M. Brown, Wnt proteins induce dishevelled phosphorylation via an LRP5/6-independent mechanism, irrespective of their ability to stabilize  $\beta$ -catenin. *Mol. Cell. Biol.* **24**, 4757-4768 (2004).
48. H. J. Lee, D. L. Shi, J. J. Zheng, Conformational change of Dishevelled plays a key regulatory role in the Wnt signaling pathways. *eLife* **4**, e08142 (2015).
49. I. Behm-Ansmant *et al.*, mRNA quality control: an ancient machinery recognizes and degrades mRNAs with nonsense codons. *FEBS Lett.* **581**, 2845-2853 (2007).
50. B. Madan *et al.*, Wnt addiction of genetically defined cancers reversed by PORCN inhibition. *Oncogene* **35**, 2197-2207 (2016).
51. G. Picco *et al.*, Loss of AXIN1 drives acquired resistance to WNT pathway blockade in colorectal cancer cells carrying RSPO3 fusions. *EMBO Mol. Med.* **9**, 293-303 (2017).
52. T. Tsukiyama *et al.*, Molecular role of RNF43 in canonical and noncanonical wnt signaling. *Mol. Cell. Biol.* **35**, 2007-2023 (2015).
53. S. M. Mazzone, E. R. Fearon, AXIN1 and AXIN2 variants in gastrointestinal cancers. *Cancer Lett.* **355**, 1-8 (2014).
54. M. van de Wetering *et al.*, Prospective derivation of a living organoid biobank of colorectal cancer patients. *Cell* **161**, 933-945 (2015).
55. M. Fujii *et al.*, A colorectal tumor organoid library demonstrates progressive loss of niche factor requirements during tumorigenesis. *Cell Stem Cell* **18**, 827-838 (2016).
56. H. Ashkenazy *et al.*, ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344-350 (2016).
57. V. Neumeyer *et al.*, Loss of endogenous RNF43 function enhances proliferation and tumour growth of intestinal and gastric cells. *Carcinogenesis* **40**, 551-559 (2019).
58. S. A. Forbes *et al.*, COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res.* **43**, D805-811 (2015).
59. W. Wang *et al.*, Blocking Wnt secretion reduces growth of hepatocellular carcinoma cell lines mostly independent of  $\beta$ -catenin signaling. *Neoplasia* **18**, 711-723 (2016).
60. F. A. Ran *et al.*, Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281-2308 (2013).

# **Chapter 7**

## **Summary**



It has become well established that alterations in the abundance and diversity of bacterial species present in the gastrointestinal tract, can lead to the development of premalignant and malignant lesions in the human stomach, colon, and rectum (1–3). However, detection of microbes or microbial components in pancreatic neoplasms has not yet been performed, let alone microbiome profiling. Despite the well-known critical roles of gut bacterial translocation and its involvement in aggravation of pancreatitis, it has remained unclear whether the pancreas contains bacterial contaminations that may contribute to cystogenesis or tumorigenesis. In **Chapter 2**, we show that pancreatic cyst fluid (PCF) harbors abundant bacteria, which are not introduced by the endoscopic sampling procedure. To this aim, we first performed PCR amplification of the bacterial 16S ribosomal RNA (16S rRNA) gene and observed that the great majority (92.8%) of the PCF samples are positive for bacterial DNA detection. We next investigated the PCF microbiome with the help of amplicon-based bacterial 16S rDNA Sanger sequencing and next-generation sequencing (NGS). The local microbial signature identified by NGS analysis showed the presence of 408 bacterial genera in this specific niche. We then integrated the PCF microbiota fingerprint with the microbiomes observed in 13 other body sites, which were retrieved from the Human Microbiome Project (HMP) database. Interestingly, the PCF microbiome uniquely comprised 17 predominant bacteria. In contrast, all the 13 other body sites were dominated by 15 bacterial genera that however were least abundant in the PCF microbiota, suggesting that a unique microbial ecosystem is present in the pancreatic cystic neoplasms. Importantly, of all the uniquely abundant bacteria in the PCF microbiome, *Bacteroides* spp., *Escherichia/Shigella* spp., and *Acidaminococcus* spp. were prominent, and a considerable abundance of *Staphylococcus* spp. and *Fusobacterium* spp. was also present. All these bacterial species are potential pro-oncogenic (1–4), suggesting a possible link of the microbiome to pancreatic carcinogenesis.

Increased abundance of pathogens attributable to bacterial dysbiosis is also present during development of colorectal cancer (CRC). However, the mere presence of specific pathogens in the gut cannot fully explain the cancer-promoting characteristics of gut microbiota. Hence, there might be an additional mechanism by

which the resident gut microbes in the intestinal lumen and mucus layer cooperatively induce cancer development. As such, intestinal biofilms formed by highly-ordered structural amalgamation of microbes are emerging as an important causative factor for human CRC (5–9). In **Chapter 3**, we comprehensively reviewed the cancer-promoting effects and underlying mechanisms of bacterial biofilm involved in CRC development. We first introduced the important roles of gut microbiota in tumorigenesis. The link of polymicrobial infections to carcinogenesis was then highlighted, because unlike gastric or liver cancers, which are linked to *Helicobacter pylori*, and chronic hepatitis B or C virus, respectively, most other cancers are not attributable to a single pathogen. Evidence suggests that colonic mucosal biofilms promote bacteria-driven CRC through compromising intestinal epithelial barrier function, increasing genotoxicity, exacerbating inflammation, modulating host metabolisms, and promoting epithelial cell proliferation (5–7). We argue that biofilm communities that include bacteria possessing epithelial adherence capacity and coaggregation properties, such as *Fusobacteria*, may be essential for bacteria-induced CRC initiation. We postulate that the interplay between intestinal barrier loss and bacterial dysbiosis are boosted by bacterial biofilms in CRC. Taken together, the formation of bacterial biofilms in the gut enhances most of the tumor-promoting properties that have previously been linked to individual bacterial species.

By considering the clinicopathological characteristics of biofilm-promoted CRC (e.g., preference of proximal location in colon, and inflammatory context) (6), it is reasonable to expect that bacterial biofilms may not only promote carcinogenesis, but also have the potential to induce a specific subtype of CRC. In **Chapter 4**, we propose that these biofilms possess mucus stimulating abilities and thereby contribute to mucinous CRC formation. In the normal colon, mucus forms a firm and a loose layer that jointly serves as the first line of defense against colonization of pathogens or microbiota with potentially deleterious effects, thereby impeding bacteria-driven cancer initiation. There is a balance between microbiota-mediated mucus degradation and mucus replenishment of Goblet cells. In this respect, bacterial colonization and invasion could induce elevated mucus production. We also briefly review the forms of genetic instability observed in CRC, which are chromosomal instability (CIN),

microsatellite instability (MSI), also known as DNA mismatch repair (MMR) deficiency, and the CpG island methylator phenotype (CIMP). Of these subtypes of CRC, the ones with MSI and/or CIMP are predominantly observed in the proximal (right-sided) colon, which is also the site where most mucinous CRCs are detected. Interestingly, this is also the preferential location of cancer-associated biofilms, suggesting a mechanistic link between these biofilms and mucinous CRCs. We propose three potential explanations for this proximal preference. A thinner mucus layer, colonization of specific microbiota (e.g., *Fusobacteria*-dominant), and the more fluidic luminal content in the proximal colon, all appear to favor a more efficient formation and adherence of bacterial biofilm on this side. Finally, we delineate a possible molecular basis for the contribution of bacterial biofilm to mucinous differentiation. We argue that the mucus-promoting effects of biofilms are probably weak in tumors that do not express ATOH1, and are preferably present in tumors with a moderate activation of oncogenic  $\beta$ -catenin signaling.

As stated above, biofilms may indirectly result in ATOH1 activation leading to mucinous differentiation. The molecular mechanisms underlying modulation of ATOH1 activity in CRC are however still unclear. In **Chapter 5**, we have evaluated previously reported mechanisms and identified new players to stabilize ATOH1 protein, one of the main transcriptional factors to regulate mucinous differentiation in the gastrointestinal tract. First, we show that ATOH1 positively correlates with expression of secreted mucins. Moreover, we observed that a non-mucinous CRC xenograft expressing ATOH1, phenocopied mucinous histology. Using CRISPR/Cas9-mediated knockout and various biochemical studies, we demonstrate that the APC protein may be involved in stabilizing ATOH1. This stabilization process does however not depend on either GSK3 $\beta$  or  $\beta$ -catenin activity, as was previously suggested. SHH inhibition led to increased ATOH1 levels, opposing the effects seen in medulloblastoma. The previously reported JAK2-mediated phosphorylation of Y80 also moderately increased ATOH1 levels in CRC cell lines. Next, through mass-spectrometry analysis and co-immunoprecipitation, we found that the E3 ubiquitin ligase UBR5 can bind to ATOH1. UBR5 functions via its ubiquitin ligase activity to

stabilize ATOH1 by preventing proteasomal degradation in a subset of CRC cell lines. UBR5 exhibits this activity in a complex with or independent of DYRK proteins, for which we have observed to also be involved in ATOH1 stabilization using the DYRK2 and DYRK4 inhibitor ID-8. Attempts to link all these mechanisms failed or were inconclusive, likely showing that ATOH1 protein regulation is a multifactorial process. Lastly, we propose a potential biomarker signature in which *DYRK4* and *MUC2* jointly allow to molecularly distinguish CRC with a mucinous subtype from its non-mucinous counterpart.

In **chapter 6**, we investigate the relevance of specific truncating *RNF43* mutations for the tumorigenic process. Most CRCs carry inactivating *APC* mutations, which is generally associated with a strong  $\beta$ -catenin activation, thereby making a prominent mucinous differentiation less likely. In contrast, there is an alternative route, in which inactivating *RNF43* mutations result in a moderate  $\beta$ -catenin activation in a Wnt-dependent manner (10,11). These mutations can lead to an increase in Wnt-receptor levels at the cell surface. Such mutations could render cancer cells sensitive to therapeutics targeting Wnt-ligand secretion and the Wnt-receptor itself (12–14). Moreover, given that *RNF43* is frequently mutated in various human cancers, such as colorectal and pancreatic cancers, its mutation has gained substantial attention (15–17). However, very few studies have shown the extent to which *RNF43* mutations activate  $\beta$ -catenin signaling. In **Chapter 6**, we conclude that commonly observed C-terminal truncations of *RNF43* retain their function to downregulate  $\beta$ -catenin signaling, and therefore do not confer Wnt-dependency onto colorectal cancers. This implies that Wnt-based therapeutics may not benefit a subset of cancer patients harboring these C-terminal *RNF43* mutations. In this study, we first performed analyses of cancer genomics datasets and observed that tumors with a functional MMR system show a preferential 5'-location of *RNF43* truncating mutations, suggesting that C-terminal truncations fail to impact oncogenic  $\beta$ -catenin signaling. Extensive molecular investigations of the  $\beta$ -catenin regulating properties of *RNF43* mutants were then performed. Overexpression of C-terminal truncation mutants and wild-type *RNF43* exhibited comparable effects on  $\beta$ -catenin signaling,

FZD-receptor turnover and DVL-binding. These observations suggest that C-terminal truncations of RNF43 retain basically the same level of  $\beta$ -catenin regulation as the wild-type protein. These findings were confirmed at endogenous levels in cell lines using siRNA-mediated knockdown and CRISPR/Cas9-mediated knockout and knockin. We also evaluated other previously reported interactors of RNF43. We found that the RNF43-mediated nuclear membrane sequestration of TCF7L2 is less prominent than its Wnt-receptor turnover function in CRC cells. In addition, we could not reproduce previous results that suggested that RNF43 associates with p53 and E-cadherin breakdown. Phosphorylation levels of DVL2 in a panel of 17 CRC lines were evaluated for their responsiveness to the Wnt-secretion inhibitor IWP12, showing that only CRC cells carrying a N-terminal RNF43 truncation respond to Wnt inhibition. In parallel, we show that only CRC cells harboring N-terminal truncations of RNF43 are impaired in their growth by blocking Wnt secretion. The findings in this chapter are not only relevant for understanding RNF43-driven tumorigenesis, but also have important implications for the large number of upstream Wnt inhibitors in development and in clinical trials.



## References

1. S. R. Konstantinov, E. J. Kuipers, M. P. Peppelenbosch, Functional genomic analyses of the gut microbiota for CRC screening. *Nat. Rev. Gastroenterol. Hepatol.* **10**, 741-745 (2013).
2. V. Gopalakrishnan, B. A. Helmink, C. N. Spencer, A. Reuben, J. A. Wargo, The influence of the gut microbiome on cancer, immunity, and cancer immunotherapy. *Cancer Cell* **33**, 570-580 (2018).
3. J. Chen, J. C. Domingue, C. L. Sears, Microbiota dysbiosis in select human cancers: Evidence of association and causality. *Semin. Immunol.* **32**, 25-34 (2017).
4. R. F. Schwabe, C. Jobin, The microbiome and cancer. *Nat. Rev. Cancer* **13**, 800-812 (2013).
5. C. M. Dejea *et al.*, Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria. *Science* **359**, 592-597 (2018).
6. C. M. Dejea *et al.*, Microbiota organization is a distinct feature of proximal colorectal cancers. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18321-18326 (2014).
7. C. H. Johnson *et al.*, Metabolism links bacterial biofilms and colon carcinogenesis. *Cell Metab.* **21**, 891-897 (2015).
8. J. Yu *et al.*, Invasive *Fusobacterium nucleatum* may play a role in the carcinogenesis of proximal colon cancer through the serrated neoplasia pathway. *Int. J. Cancer* **139**, 1318-1326 (2016).
9. J. L. Drewes *et al.*, High-resolution bacterial 16S rRNA gene profile meta-analysis and biofilm status reveal common colorectal cancer consortia. *NPJ Biofilms Microbiomes* **3**, (2017).
10. B. K. Koo *et al.*, Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors. *Nature* **488**, 665-669 (2012).
11. H. X. Hao, X. Jiang, F. Cong, Control of Wnt receptor turnover by R-spondin-ZNRF3/RNF43 signaling module and its dysregulation in cancer. *Cancers (Basel)* **8**, 54 (2016).
12. B. K. Koo, J. H. van Es, M. van den Born, H. Clevers, Porcupine inhibitor suppresses paracrine Wnt-driven growth of *Rnf43;Znrf3*-mutant neoplasia. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7548-7550 (2015).
13. X. Jiang *et al.*, Inactivating mutations of *RNF43* confer Wnt dependency in pancreatic ductal adenocarcinoma. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 12649-12654 (2013).
14. Z. Steinhart *et al.*, Genome-wide CRISPR screens reveal a Wnt-FZD5 signaling circuit as a druggable vulnerability of *RNF43*-mutant pancreatic tumors. *Nat. Med.* **23**, 60-68 (2017).
15. M. Giannakis *et al.*, *RNF43* is frequently mutated in colorectal and endometrial cancers. *Nat. Genet.* **46**, 1264-1266 (2014).
16. Y. Jiao *et al.*, Whole-exome sequencing of pancreatic neoplasms with acinar differentiation. *J. Pathol.* **232**, 428-435 (2014).
17. N. Waddell *et al.*, Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature* **518**, 495-501 (2015).

# **Chapter 8**

## **Discussion**



During the last decades, remarkable advances in unveiling tumorigenesis of digestive system cancers have been made. However, current knowledge on genetic alterations, epigenetic changes, and environmental exposure to carcinogens in combination cannot fully recapitulate the pathogenesis of most malignancies. The gut microbiome has been attracting intense attention with respect to human physiology and pathology in the last few years. It is becoming clear that interaction of the gut microbiota with the host tissue acts as an important contributor to carcinogenesis. This thesis has attempted to demonstrate microbiome and molecular characterizations of neoplasms in the pancreas and intestine. These findings may help shed light into the etiology of digestive cancer. Inspired by a research approach called “microbiological-molecular-pathological-epidemiology” (1), here I discuss the interplay between the microbiome, host genetic variations and DNA mutations acquired during tumorigenesis.

### **Microbiota induces inflammation, activates oncogenic signaling, and leads to genetic alterations**

In agreement with more recent reports (2–4), we show a tantalizing possibility that pancreatic neoplasms, like *Helicobacter Pylori*-induced gastric neoplasms, are probably a bacteria-induced disease. Potential pathogens such as *Bacteroides* spp. and *Escherichia/Shigella* spp. were highly abundant in the pancreatic cyst fluid (PCF) samples. In particular, a considerable amount of *Fusobacteria* was detected (**Chapter 2**). *Fusobacterium* species are considered to be a highly invasive pathogen in a variety of human diseases, including CRC (5–11). They have been found to be able to attach to surfaces of the intestinal mucosa (12) and localize within the crypts of colon and rectum (13). Causal evidence supports that *Fusobacteria* regulate tumor-eliciting immune responses, which induce over-expression of cytokines, and create a supportive microenvironment for CRC development (6). More importantly, their critical role in oncogenic signaling pathways should be underlined, in particular  $\beta$ -catenin signaling that is one of the key pathways involved in pancreatic and colorectal cancer. For instance, *Fusobacteria* can release a virulence adhesion molecule FadA to bind with E-cadherin and subsequently activate  $\beta$ -catenin signaling (14). Several

studies have established that loss of or decreased levels of E-cadherin in epithelial cells is associated with cancer-eliciting inflammation and neoplasms (15,16). Reduction of E-cadherin in colonic epithelium conducts signaling cascades that trigger carcinogenesis and tumor invasion (17,18). Moreover, one well-known pathogenic strain of *Bacteroides* spp. that is highly abundant in the PCF microbiome (**Chapter 2**), the so-called Enterotoxigenic *Bacteroides fragilis* (ETBF), could generate a toxin (*B. fragilis* toxin, BFT) (19). This BFT toxin is also able to induce E-cadherin cleavage and thereby lead to activation of  $\beta$ -catenin signaling (19). This toxin also possesses genotoxic ability that promotes production of reactive oxygen species (ROS) resulting in DNA damage and genomic instability (20) (reviewed in **Chapter 3**).

In addition, there is a strong association of spatial organization of microbial communities, the so-called biofilm, with CRC initiation and development (5,21–23) (reviewed in **Chapter 3**). The microbial pathogens engaging in this high order structure not only offer protection against host defences but also exert enhanced pathogenic functions to their host such as excessive inflammatory responses that have the potency to impose a higher mutation burden onto the epithelial cells, thereby increasing the chance of tumor formation. Mucosal biofilm is also associated with impairment or redistribution of E-cadherin in epithelial cells (5). This markedly enhances microbial pathogens adherence to exposed mucosal surfaces, contributing to pro-cancerous inflammation and activating Wnt/ $\beta$ -catenin-related signaling (5). Moreover, bacterial biofilm may promote early colorectal tumorigenesis through acquisition of the “second hit” *Apc* mutation, as exemplified by the use of *Apc*<sup>Min/+</sup> mice model carrying a single truncated allele of *Apc* (21). For instance, co-colonization of polyketide synthase (*pks*)-positive *Escherichia coli* and ETBF in *Apc*-deficient mice led to a much earlier development of neoplasia as compared with sham-control or mice colonized by a single species alone (21). In this case, such carcinogenic microbial organization promotes inactivation of the second *Apc* allele by loss-of-heterozygosity (LOH) or increases the chance that APC-mutant cells grow into a visible tumor. Overall, these observations point out that gut microbiota not only induces

inflammation, but also leads to oncogenic activation of signaling pathways and genetic alterations. Such events orchestrate cancer initiation and development.

### **Genetic alteration influences composition and spatial organization of microbiota**

Tumors having different genetic backgrounds can create distinct microenvironments favorable for specific species of bacteria. DNA mutations may affect the composition of the microbial community that is present in the lumen, mucosa, or inside tumors. For example, a previous report suggested that the *KRAS* p.G12D mutation, a commonly observed oncogenic driver mutation in pancreatic cancer, could contribute to an increase in translocation of gut microbiota to the pancreas. It led to distinct differences in diversity and composition of the intestinal and intra-pancreatic microbiota, as exemplified by a higher translocation capability of fecal microbiota in *Kras* mutated mice than in their wild-type littermates (3). Another example of significant differences in intratumoral microbial abundances were found in an *APC* mutation-specific manner, e.g., *Christensenellaceae* was less abundant in human CRCs carrying *APC* loss-of-function (LoF) mutation (24). Importantly, *APC* LoF mutations could increase the capability of bacteria to adhere to and colonize intestinal mucosa, thereby modulating the host-microbe interplay. For instance, *Bacteroides* and *Enterobacteriaceae* in colonic mucosa of *Apc*<sup>Min/+</sup> mice were significantly higher enriched and more cultivable when compared with that of its wild-type littermates (21).

Moreover, the appearance and bacterial composition of biofilms is also influenced by different genomic changes existing or arising within hosts. *Fusobacteria*-dominant bacterial biofilms develop predominantly in the proximal (right-sided) colon of sporadic CRC (5) (**Chapter 4**). Interestingly, it is noteworthy that biofilms are not dominated by *Fusobacteria* when intestinal tumors arise in patients or mice carrying a germline *APC/Apc* mutation. For instance, bacterial biofilms consisting predominantly of *E. coli* and *B. fragilis* were detected in colonic mucosa samples from most patients with familial adenomatous polyposis (FAP), which is characterized by the development of precancerous colorectal lesions (polyps)

driven by bi-allelic *APC* loss (21). Moreover, the biofilm formed upon this genetic condition is patchy distributed along the whole length of the large intestine, in contrast to the continuous presence of *Fusobacteria*-dominant biofilm in the proximal colon (5,21). Such FAP biofilm organization was detected prior to early-stage colonic neoplasia, suggesting that this biofilm develops at early times preceding tumor formation. In contrast, the *Fusobacteria*-dominant biofilm preferentially located in the proximal colon in sporadic cancer may develop as a result of the emerging tumors (5,21). It is therefore conceivable that microbial composition and dominant species of biofilm are determined under conditions relevant to the APC mutation status.

Therefore, genetic alterations might be capable of determining composition of microbiota and higher-order organization thereof. Further investigations are thus required for underlying the molecular mechanisms that regulate biofilm formation associated with cancer development and progression.

### **Microbial regulation of cancer cell proliferation and differentiation may depend on genetic context**

As discussed in **Chapter 4**, aggregation and collaboration of multiple microbial species in biofilms appear to have an increased potential to promote mucus secretion in colonic tumor cells, which may in part also be attributable to the exacerbation of inflammation induced by these biofilms. An important prerequisite for the increased mucus production by tumor cells that we propose, is that the genetic alterations present in the cancer cells still allow for sufficient differentiation towards the mucinous direction.

Interestingly, our comprehensive molecular analysis showing that ATOH1 is more stable when functional APC protein is expressed, whereas ATOH1 protein levels are more rapidly reduced upon APC loss. Expression levels of secreted mucin genes are also decreased in APC-deficient tumors (**Chapter 5**). *APC* mutations generally associate with higher levels of  $\beta$ -catenin signaling, which likely impedes an efficient differentiation towards the mucinous direction. On the other hand, N-terminal truncating *RNF43* mutations lead to a modest level of  $\beta$ -catenin signaling activation

(**Chapter 6**). These observations infer that *RNF43* inactivating mutations may be more favorable for cancer cells to induce mucinous differentiation than APC loss (25). A second important difference between APC- and RNF43-mutant cancers, is that the latter remain dependent on extracellular Wnt-ligand exposure to induce  $\beta$ -catenin signaling. Given the extensive presence of intra-tumor heterogeneity in various tumors types (26,27), it is reasonable to expect that not all RNF43-mutant cancer cells are exposed to the same level of Wnt-ligands. Cancer cells within the tumor mass that receive high Wnt levels, induce higher levels of  $\beta$ -catenin signaling and are more likely to show a stem-cell and/or proliferative phenotype. On the other hand, with insufficient exposure of Wnt-ligands, these same RNF43-mutant cancer cells may be allowed to induce a mucinous differentiation. This hypothetical scenario may explain why proliferative and differentiated compartments co-occur in mucinous tumors. Thus, inactivating *RNF43* mutations may simultaneously induce tumorigenesis and mucinous differentiation as present in mucinous neoplasms such as MCC or IPMN.

Specific bacteria in the microbiome may induce a different mutational pattern to the host. A recent study indicated that polyketide synthase-positive (*pks*<sup>+</sup>) *E. coli* strains secrete the genotoxin colibactin that could directly induce a specific mutational signature, comprised of increased numbers of single-base substitutions and indels, into wild-type organoids derived from the healthy human intestine. This unique mutational signature is present in 9.4% and 16.3% of CRC samples in a UK cohort and a Dutch cohort, respectively (28). Moreover, it has been known that enrichment of *Fusobacterium* or ETBF is related to specific CRC molecular subtypes (9,29,30). Elevated levels of *Fusobacteria* may cause hypermethylation of MLH1 leading to a mismatch repair (MMR) defect, probably owing to increased release of reactive oxygen species (ROS) or hydrogen sulfide (31,32). ETBF infection could lead to promoter hypermethylation of CpG islands during inflammation-induced tumorigenesis in a MMR protein MSH2-dependent manner (30). However, it should be noted that many of the mutations induced by these bacteria themselves or in combination with a defective MMR, might be “bystander mutations”. In **Chapter 6**, we show one such example, i.e. the C-terminal truncating *RNF43* mutations that



positively correlate with MMR-deficient tumors, but nevertheless do not represent a true “driver” mutation in carcinogenesis. This calls attention to a proper functional evaluation of mutations in future investigations, combined with an integrative analysis of the microbiome and human cancer genome.

In conclusion, recapitulating the findings in this thesis, one can envision that integrative analyses of the microbiome and molecular characterizations, holds great promise for better understanding the pathogenesis of neoplasms in the pancreas and intestine. Moreover, there has been a strong requirement regarding precision medicine strategies for pancreatic and colorectal cancer, whereas the information provided by previous studies remains limited. Therefore, integration of more detailed analyses of genetic, molecular, microbiome as well as epidemiological features in prospective studies is warranted, since diagnostically and therapeutically relevant (1).

## References

1. T. Hamada, J. A. Nowak, D. A. Milner, Jr., M. Song, S. Ogino, Integration of microbiology, molecular pathology, and epidemiology: a new paradigm to explore the pathogenesis of microbiome-driven neoplasms. *J. Pathol.* **247**, 615-628 (2019).
2. R. Mendez *et al.*, Microbial dysbiosis and polyamine metabolism as predictive markers for early detection of pancreatic cancer. *Carcinogenesis*, 561-570 (2019).
3. S. Pushalkar *et al.*, The Pancreatic Cancer Microbiome Promotes Oncogenesis by Induction of Innate and Adaptive Immune Suppression. *Cancer Discov.* **8**, 403-416 (2018).
4. E. Del Castillo *et al.*, The Microbiomes of Pancreatic and Duodenum Tissue Overlap and Are Highly Subject Specific but Differ between Pancreatic Cancer and Noncancer Subjects. *Cancer Epidemiol. Biomarkers Prev.* **28**, 370-383 (2019).
5. C. M. Dejea *et al.*, Microbiota organization is a distinct feature of proximal colorectal cancers. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18321-18326 (2014).
6. A. D. Kostic *et al.*, *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe.* **14**, 207-215 (2013).
7. A. D. Kostic *et al.*, Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome Res.* **22**, 292-298 (2012).
8. K. Mima *et al.*, *Fusobacterium nucleatum* in colorectal carcinoma tissue according to tumor location. *Clin. Transl. Gastroenterol.* **7**, e200 (2016).
9. K. Mima *et al.*, *Fusobacterium nucleatum* in colorectal carcinoma tissue and patient prognosis. *Gut* **65**, 1973-1980 (2016).
10. K. Mima *et al.*, *Fusobacterium nucleatum* and T cells in colorectal carcinoma. *JAMA Oncol.* **1**, 653-661 (2015).
11. M. Castellarin *et al.*, *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome Res.* **22**, 299-306 (2012).
12. A. Swidsinski *et al.*, Acute appendicitis is characterised by local invasion with *Fusobacterium nucleatum*/necrophorum. *Gut* **60**, 34-40 (2011).
13. A. N. McCoy *et al.*, *Fusobacterium* Is Associated with Colorectal Adenomas. *PLoS One* **8**, e53653 (2013).
14. M. R. Rubinstein *et al.*, *Fusobacterium nucleatum* promotes colorectal carcinogenesis by modulating E-cadherin/ $\beta$ -catenin signaling via its FadA adhesin. *Cell Host Microbe.* **14**, 195-206 (2013).
15. N. Gassler *et al.*, Inflammatory bowel disease is associated with changes of enterocytic junctions. *Am. J. Physiol. Gastrointest. Liver Physiol.* **281**, G216-228 (2001).
16. T. Kucharzik, S. V. Walsh, J. Chen, C. A. Parkos, A. Nusrat, Neutrophil transmigration in inflammatory bowel disease is associated with differential expression of epithelial intercellular junction proteins. *Am. J. Pathol.* **159**, 2001-2009 (2001).
17. H. Semb, G. Christofori, The tumor-suppressor function of E-cadherin. *Am. J. Hum. Genet.* **63**, 1588-1593 (1998).
18. S. H. M. Wong, C. M. Fang, L. H. Chuah, C. O. Leong, S. C. Ngai, E-cadherin: Its dysregulation in carcinogenesis and clinical implications. *Crit. Rev. Oncol. Hematol.* **121**, 11-22 (2018).
19. S. Wu, P. J. Morin, D. Maouyo, C. L. Sears, *Bacteroides fragilis* enterotoxin induces c-Myc expression and cellular proliferation. *Gastroenterology* **124**, 392-400 (2003).
20. A. C. Goodwin *et al.*, Polyamine catabolism contributes to enterotoxigenic *Bacteroides fragilis*-induced colon tumorigenesis. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15354-15359 (2011).
21. C. M. Dejea *et al.*, Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria. *Science* **359**, 592-597 (2018).
22. J. L. Drewes *et al.*, High-resolution bacterial 16S rRNA gene profile meta-analysis and biofilm status reveal common colorectal cancer consortia. *NPJ Biofilms Microbiomes.* **3**, 34 (2017).

23. C. H. Johnson *et al.*, Metabolism links bacterial biofilms and colon carcinogenesis. *Cell Metab.* **21**, 891-897 (2015).
24. M. B. Burns *et al.*, Colorectal cancer mutational profiles correlate with defined microbial communities in the tumor microenvironment. *PLoS Genet.* **14**, e1007376 (2018).
25. S. O. Kleeman *et al.*, Exploiting differential Wnt target gene expression to generate a molecular biomarker for colorectal cancer stratification. *Gut* **69**, 1092-1103 (2020).
26. N. Andor *et al.*, Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. *Nat. Med.* **22**, 105-113 (2016).
27. B. Losic *et al.*, Intratumoral heterogeneity and clonal evolution in liver cancer. *Nat. Commun.* **11**, 291 (2020).
28. C. Pleguezuelos-Manzano *et al.*, Mutational signature in colorectal cancer caused by genotoxic pks. *Nature* **580**, 269-273 (2020).
29. T. Tahara *et al.*, *Fusobacterium* in colonic flora and molecular features of colorectal carcinoma. *Cancer Res.* **74**, 1311-1318 (2014).
30. A. R. Maiuri *et al.*, Mismatch Repair Proteins Initiate Epigenetic Alterations during Inflammation-Driven Tumorigenesis. *Cancer Res. Treat.* **77**, 3467-3478 (2017).
31. M. Koi, Y. Okita, J. M. Carethers, *Fusobacterium nucleatum* infection in Colorectal Cancer: Linking Inflammation, DNA Mismatch Repair and Genetic and Epigenetic Alterations. *Journal of the Anus, Rectum and Colon* **2**, 37-46 (2018).
32. V. L. Hale *et al.*, Distinct microbes, metabolites, and ecologies define the microbiome in deficient and proficient mismatch repair colorectal cancers. *Genome Med.* **10**, 78 (2018).

# **Chapter 9**

## **Dutch Summary (Nederlandse Samenvatting)**



Het is tegenwoordig algemeen geaccepteerd dat veranderingen in diversiteit en aantallen bacteriesoorten in het maagdarmkanaal kunnen leiden tot de ontwikkeling van premaligne en maligne laesies in de menselijke maag en dikke darm (1–3). Detectie van microben of microbiële componenten is in pancreas neoplasmata echter nog niet uitgevoerd, en zeker niet een gedetailleerde profilering van het microbioom. Ondanks de welbekende cruciale rol van bacteriële translocatie vanuit de darm en de betrokkenheid hiervan bij een verergering van pancreatitis, is het onduidelijk gebleven of de pancreas bacteriële verontreinigingen bevat die kunnen bijdragen aan cystogenese of tumorigenese. In **hoofdstuk 2** laten we zien dat pancreatische cyste-vloeistof een grote hoeveelheid bacteriën bevat, die niet door de endoscopische procedure geïntroduceerd worden. Hiervoor hebben we allereerst een PCR uitgevoerd voor het bacteriële 16S ribosomale RNA (16s rRNA) en zagen dat de overgrote meerderheid (92.8%) van de cysten positief zijn voor bacterieel DNA. Vervolgens is het microbioom onderzocht met behulp van Sanger sequenzen en next-generation sequencing (NGS) van bacterieel 16S rDNA. De microbiële signatuur die geïdentificeerd werd met deze analyse, toonde de aanwezigheid aan van 408 bacteriële genera in deze specifieke niche. Daarna hebben we het microbioom van de pancreascysten vergeleken met die van 13 andere organen, zoals verkregen uit het Human Microbiome Project (HMP). Het microbioom van de pancreascysten bestond overwegend uit 17 specifieke bacteriële genera. De 13 andere lichaamsplekken werden echter gedomineerd door 15 bacteriële genera die juist weinig aanwezig waren in het microbioom van de pancreascysten, wat suggereert dat er een uniek microbiel ecosysteem aanwezig is in de pancreascysten. Belangrijk is dat van alle overvloedig aanwezige bacteriën in het cysten-microbiom, vooral *Bacteroides* spp., *Escherichia/Shigella* spp. en *Acidaminococcus* spp. duidelijk aanwezig waren, naast een aanzienlijke hoeveelheid van *Staphylococcus* spp. en *Fusobacterium* spp. Al deze bacteriesoorten zijn potentieel pro-oncogeen (1–4), wat duidt op een mogelijk verband tussen het microbiom en pancreas-carcinogenese.

Een verhoogde aanwezigheid van pathogenen ontstaan door een bacteriële dysbiose, is ook aanwezig tijdens de ontwikkeling van darmkanker. Echter kan de aanwezigheid van specifieke ziekteverwekkers in de darmen de kankerbevorderende eigenschappen van de darmmicrobiota niet volledig verklaren. Daarom wordt er een aanvullend mechanisme verondersteld waarmee darmmicroben, aanwezig in zowel het darmlumen alsook de slijmlaag, gezamenlijk een bijdrage leveren aan kankerontwikkeling. De vorming van

zogenaamde bacteriële biofilms, bestaande uit een sterk geordende structuur van microben, komt de laatste jaren naar voren als een belangrijker veroorzaker van darmkanker (5–9). In **hoofdstuk 3** bediscussiëren we uitgebreid de kankerbevorderende effecten en onderliggende mechanismen van deze bacteriële biofilms. Allereerst introduceren we de belangrijke rol van specifieke darmmicrobiota betrokken bij tumorvorming. Daarna benadrukken we de relatie tussen polymicrobiële infecties en carcinogenese, want in tegenstelling tot maag- en leverkankers, die respectievelijk geassocieerd zijn met *Helicobacter pylori* en een chronische hepatitis B of C virus infectie, worden de meeste tumoren niet veroorzaakt door één specifieke pathogeen. Er zijn verschillende bewijzen dat mucosale biofilms darmkankervorming kunnen bevorderen door de barrièrefunctie van het darmepitheel te verslechteren, meer DNA-schade te induceren, ontstekingen te verergeren, het metabolisme van de gastheer te veranderen en de proliferatie van epitheelcellen te bevorderen (5–7). We beargumenteren dat biofilms die bacteriën bevatten met zowel een hechtingscapaciteit aan het onderliggende epitheel alsook coaggregatie-eigenschappen, zoals *Fusobacteria*, essentieel kunnen zijn voor door bacteriën geïnduceerde darmkanker. Samenvattend versterkt de vorming van bacteriële biofilms in de darmen de meeste tumorbevorderende eigenschappen die eerder in verband zijn gebracht met individuele bacteriesoorten.

Door rekening te houden met de klinisch-pathologische kenmerken van biofilm-geassocieerde darmkanker (bijv. proximale voorkeurslocatie in de colon en inflammatoire context) (6), kan redelijkerwijs verwacht worden dat biofilms niet alleen tumorvorming bevorderen, maar ook de potentie hebben om een specifiek subtype darmkanker te induceren. In **hoofdstuk 4** stellen we voor dat biofilms slijmstimulerende eigenschappen bezitten en daarmee kunnen bijdragen aan de vorming van slijmvormende darmkankers. In de normale dikke darm vormt het slijm een stevige en losse laag die samen dient als de eerste verdedigingslinie tegen kolonisatie van pathogenen of microbiota met mogelijke schadelijke effecten. Hierdoor wordt de bijdrage van bacteriën aan kankerinitiatie verminderd. Er is een evenwicht tussen door microbiota gemedieerde slijmafbraak en aanvulling vanuit slijmbekercellen. In dit opzicht kunnen bacteriële kolonisatie en invasie een verhoogde slijmproductie bewerkstelligen. We bespreken ook kort de vormen van genetische instabiliteit die worden waargenomen bij darmkanker, namelijk chromosomale instabiliteit (CIN), microsatellietinstabiliteit (MSI), ook wel bekend als DNA mismatch

repair (MMR) -deficiëntie, en het CpG-eiland methylator fenotype (CIMP). Van deze subtypes van darmkanker worden degenen met MSI en/of CIMP voornamelijk waargenomen in de proximale (rechtszijdige) dikke darm, wat ook de plaats is waar de meeste slijmvormende darmkankers worden waargenomen. Interessant is dat dit ook de voorkeurslocatie is van met kanker geassocieerde biofilms, wat een mechanistisch verband suggereert tussen biofilms en slijmvormende darmkankers. We stellen drie mogelijke verklaringen voor deze proximale voorkeur voor. Een dunnere slijmlaag, kolonisatie van specifieke microbiota (bijvoorbeeld *Fusobacteria*-dominant) en de vloeibaardere darminhoud in de proximale dikke darm, lijken allemaal een efficiëntere vorming en hechting van bacteriële biofilms aan deze kant te bevorderen. Tenslotte schetsen we een mogelijke moleculaire basis voor de bijdrage van bacteriële biofilms aan mucineuze differentiatie. We stellen dat de slijmbevorderende effecten van biofilms waarschijnlijk zwak zijn bij tumoren die geen ATOH1 tot expressie brengen, en bij voorkeur aanwezig zijn in tumoren met een matige activering van oncogene  $\beta$ -catenine signalering.

Zoals hierboven vermeld, kunnen biofilms indirect leiden tot activering van ATOH1, wat leidt tot mucineuze differentiatie. De moleculaire mechanismen die betrokken zijn bij de regulatie van ATOH1-activiteit in darmkanker zijn echter nog steeds onduidelijk. In **hoofdstuk 5** hebben we eerder gerapporteerde mechanismen geëvalueerd en nieuwe spelers geïdentificeerd om het ATOH1 eiwit te stabiliseren, één van de belangrijkste transcriptie-factoren om de mucineuze differentiatie in het maagdarmkanaal te reguleren. Allereerst laten we zien dat ATOH1 positief correleert met de expressie van uitgescheiden mucines. Bovendien zien we dat een xenotransplantaat van een niet-slijmvormende darmkankercellijn die ATOH1 tot expressie brengt, een histologie laat zien met slijmvormende kenmerken. Met behulp van CRISPR/Cas9-gemedieerde knock-out en verschillende biochemische onderzoeken, tonen we aan dat het APC-eiwit mogelijk betrokken is bij het stabiliseren van ATOH1. Dit stabilisatieproces is echter niet afhankelijk van de activiteit van GSK3 $\beta$  of  $\beta$ -catenine, zoals eerder werd gesuggereerd. SHH-remming leidde tot verhoogde ATOH1-spiegels, in tegenstelling tot de effecten die worden gezien bij medulloblastoom. De eerder gerapporteerde JAK2-gemedieerde fosforylering van Y80 in het ATOH1 eiwit, leidde ook in darmkanker cellijnen tot een milde verhoging van ATOH1-niveaus. Vervolgens ontdekten we door middel van massaspectrometrie en co-immunoprecipitatie dat de E3 ubiquitinligase UBR5 kan binden aan ATOH1. UBR5



functioneert via zijn ubiquitineligase-activiteit om ATOH1 te stabiliseren door proteasomale afbraak te voorkomen in een deel van de darmkankercellijnen. UBR5 vertoont deze activiteit in een complex met of onafhankelijk van DYRK-eiwitten, waarvoor we hebben waargenomen dat ze ook betrokken zijn bij ATOH1-stabilisatie met behulp van de DYRK2 en DYRK4-remmer ID-8. Pogingen om al deze mechanismen te koppelen faalden of waren niet sluitend, wat waarschijnlijk aantoont dat ATOH1-eiwitregulatie een multifactoriaal proces is. Ten slotte stellen we een mogelijke biomarkerhandtekening voor, waarin *DYRK4* en *MUC2* het gezamenlijk mogelijk maken om CRC met een mucineus subtype moleculair te onderscheiden van zijn niet-mucineuze tegenhanger.

In **hoofdstuk 6** onderzoeken we de relevantie van specifieke truncerende *RNF43*-mutaties voor het tumorigene proces. De meeste darmkankers bevatten inactiverende *APC*-mutaties, die over het algemeen worden geassocieerd met een sterke  $\beta$ -catenine-activering, waardoor een prominente slijmproductie minder waarschijnlijk wordt. Daarentegen is er een alternatieve route, waarbij inactiverende *RNF43*-mutaties resulteren in een matige  $\beta$ -catenine-activering op een Wnt-afhankelijke manier (10,11). Deze mutaties kunnen leiden tot een verhoging van de Wnt-receptorniveaus aan het celoppervlak. Dergelijke mutaties kunnen kankercellen gevoelig maken voor therapeutica die gericht zijn op Wnt-ligandsecretie en de Wnt-receptor zelf (12–14). Aangezien *RNF43* vaak gemuteerd is in verschillende vormen van kanker bij de mens, zoals colorectale en pancreaskanker, heeft de mutatie ervan aanzienlijke aandacht gekregen (15–17). Er zijn echter maar weinig onderzoeken die hebben aangetoond in welke mate *RNF43*-mutaties echt tot activatie van  $\beta$ -catenine signalering leiden. In **hoofdstuk 6** concluderen we dat algemeen waargenomen C-terminale truncaties van *RNF43* hun functie behouden om  $\beta$ -catenine-signalering te downreguleren, en daardoor niet leiden tot Wnt-afhankelijkheid van darmkankers. Dit impliceert dat Wnt-gebaseerde therapeutica mogelijk weinig effect hebben bij een deel van de kankerpatiënten die deze C-terminale *RNF43*-mutaties dragen. Om tot deze conclusie te komen, hebben we allereerst datasets geanalyseerd van kankergenomen en vastgesteld dat tumoren met een functioneel MMR-systeem voornamelijk *RNF43*-truncerende mutaties vertonen aan de 5'-kant van het gen. Dit suggereert dat C-terminale truncaties mogelijk geen effect hebben op oncogene  $\beta$ -catenine-signalering. Vervolgens werden uitgebreide moleculaire onderzoeken naar de  $\beta$ -catenine-regulerende eigenschappen van *RNF43*-mutanten uitgevoerd. Overexpressie van

C-terminale truncatiemutanten en wildtype RNF43 vertoonde vergelijkbare effecten op  $\beta$ -catenine-signalering, afbraak van de FZD-receptor en DVL-binding. Deze waarnemingen suggereren dat C-terminale truncaties van RNF43 in principe hetzelfde niveau van  $\beta$ -catenine-regulering behouden als het wildtype eiwit. Deze bevindingen werden bevestigd op endogene niveaus in cellijnen met behulp van door siRNA gemedieerde knockdown en door CRISPR/Cas9 gemedieerde knock-out en knock-in. We hebben ook andere eerder gemelde interactoren van RNF43 geëvalueerd. We ontdekten dat in darmkankercellen de RNF43-gemedieerde binding van TCF7L2 aan de nucleaire membraan minder belangrijk is dan de omzetting van de Wnt-receptor op de celmembraan. Bovendien konden we eerdere resultaten niet reproduceren die suggereerden dat RNF43 betrokken is bij de binding en afbraak van P53 en E-cadherine. In een panel van 17 darmkanker cellijnen werden de fosforyleringsniveaus van DVL2 beoordeeld op hun respons op de Wnt-secretieremmer IWP12. Dit toonde aan dat alleen darmkankercellen die een N-terminale RNF43-truncatie dragen, reageren op Wnt-remming. Tegelijkertijd laten we zien dat alleen darmkankercellen met deze N-terminale RNF43 truncaties in hun groei worden belemmerd door Wnt-secretie te blokkeren. De bevindingen in dit hoofdstuk zijn niet alleen relevant voor het begrijpen van door RNF43 aangestuurde tumorigenese, maar hebben ook belangrijke implicaties voor het grote aantal Wnt-remmers in ontwikkeling en in klinische onderzoeken.

## Referenties

1. S. R. Konstantinov, E. J. Kuipers, M. P. Peppelenbosch, Functional genomic analyses of the gut microbiota for CRC screening. *Nat. Rev. Gastroenterol. Hepatol.* **10**, 741-745 (2013).
2. V. Gopalakrishnan, B. A. Helmink, C. N. Spencer, A. Reuben, J. A. Wargo, The influence of the gut microbiome on cancer, immunity, and cancer immunotherapy. *Cancer Cell* **33**, 570-580 (2018).
3. J. Chen, J. C. Domingue, C. L. Sears, Microbiota dysbiosis in select human cancers: Evidence of association and causality. *Semin. Immunol.* **32**, 25-34 (2017).
4. R. F. Schwabe, C. Jobin, The microbiome and cancer. *Nat. Rev. Cancer* **13**, 800-812 (2013).
5. C. M. Dejea *et al.*, Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria. *Science* **359**, 592-597 (2018).
6. C. M. Dejea *et al.*, Microbiota organization is a distinct feature of proximal colorectal cancers. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18321-18326 (2014).
7. C. H. Johnson *et al.*, Metabolism links bacterial biofilms and colon carcinogenesis. *Cell Metab.* **21**, 891-897 (2015).
8. J. Yu *et al.*, Invasive *Fusobacterium nucleatum* may play a role in the carcinogenesis of proximal colon cancer through the serrated neoplasia pathway. *Int. J. Cancer* **139**, 1318-1326 (2016).
9. J. L. Drewes *et al.*, High-resolution bacterial 16S rRNA gene profile meta-analysis and biofilm status reveal common colorectal cancer consortia. *NPJ Biofilms Microbiomes* **3**, (2017).
10. B. K. Koo *et al.*, Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors. *Nature* **488**, 665-669 (2012).
11. H. X. Hao, X. Jiang, F. Cong, Control of Wnt receptor turnover by R-spondin-ZNRF3/RNF43 signaling module and its dysregulation in cancer. *Cancers (Basel)* **8**, 54 (2016).
12. B. K. Koo, J. H. van Es, M. van den Born, H. Clevers, Porcupine inhibitor suppresses paracrine Wnt-driven growth of *Rnf43;Znrf3*-mutant neoplasia. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7548-7550 (2015).
13. X. Jiang *et al.*, Inactivating mutations of *RNF43* confer Wnt dependency in pancreatic ductal adenocarcinoma. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 12649-12654 (2013).
14. Z. Steinhart *et al.*, Genome-wide CRISPR screens reveal a Wnt-FZD5 signaling circuit as a druggable vulnerability of *RNF43*-mutant pancreatic tumors. *Nat. Med.* **23**, 60-68 (2017).
15. M. Giannakis *et al.*, *RNF43* is frequently mutated in colorectal and endometrial cancers. *Nat. Genet.* **46**, 1264-1266 (2014).
16. Y. Jiao *et al.*, Whole-exome sequencing of pancreatic neoplasms with acinar differentiation. *J. Pathol.* **232**, 428-435 (2014).
17. N. Waddell *et al.*, Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature* **518**, 495-501 (2015).

# **Appendices**

## **List of Publications**

## **PhD Portfolio**

## **Curriculum Vitae**

## **Acknowledgements**



## List of Publications

1. **Shan Li**, Marla Lavrijsen, Aron Bakker, Marcin Magierowski, Katarzyna Magierowska, Pengyu Liu, Wenhui Wang, Maikel P. Peppelenbosch, and Ron Smits. Commonly observed *RNF43* mutations retain functionality in attenuating Wnt/ $\beta$ -catenin signaling and unlikely confer Wnt-dependency onto colorectal cancers. *Oncogene*. 2020, 39:3458-3472.
2. **Shan Li**, Maikel P. Peppelenbosch, and Ron Smits. Bacterial biofilms as a potential contributor to mucinous colorectal cancer formation. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*. 2019, 1872(1):74-79.
3. **Shan Li**, Gwenny M. Fuhler, Nahush BN, Tony Jose, Marco J. Bruno, Maikel P. Peppelenbosch, and Sergey R. Konstantinov. Pancreatic cyst fluid harbors a unique microbiome. *Microbiome*. 2017, 5:147.
4. **Shan Li**, Sergey R. Konstantinov, Ron Smits, and Maikel P. Peppelenbosch. Bacterial biofilms in colorectal cancer initiation and progression. *Trends in Molecular Medicine*. 2017, 23(1):18-30.
5. **Shan Li**, Werner Helvensteijn, Raymond Poot, Jeroen Demmers, Wenhui Wang, Zhijiang Miao, Menggang Liu, Bingting Yu, Marla Lavrijsen, Buyun Ma, Shanshan Li, Pengyu Liu, Maikel P. Peppelenbosch, and Ron Smits. Unravelling posttranslational modifications of ATOH1 driving colorectal cancer towards mucinous differentiation. *Manuscript in Preparation*.
6. Pengyu Liu, Binyong Liang, Menggang Liu, Joyce HG Lebbink, **Shan Li**, Manning Qian, Marla Lavrijsen, Maikel P. Peppelenbosch, Xin Chen, and Ron Smits. Oncogenic mutations in Armadillo repeats 5 and 6 of  $\beta$ -catenin reduce binding to APC, increasing signaling and transcription of target genes. *Gastroenterology*. 2020, 158(4):1029-1043.
7. Wanlu Cao, Meng Li, Jiaye Liu, Shaoshi Zhang, Lisanne Noordam, Monique M. A. Verstegen, Ling Wang, Buyun Ma, **Shan Li**, Wenshi Wang, Michiel Bolkestein, Michael Doukas, Kan Chen, Zhongren Ma, Marco Bruno, Dave Sprengers, Jaap Kwekkeboom, Luc J. W. van der Laan, Ron Smits, Maikel P. Peppelenbosch, and Qiuwei Pan. LGR5 marks targetable tumor-initiating cells in mouse liver cancer. *Nature Communications*. 2020, 11:1961.

8. Zhijiang Miao, Shaoshi Zhang, Xumin Ou, **Shan Li**, Zhongren Ma, Wenshi Wang, Maikel P. Peppelenbosch, Jiaye Liu, and Qiuwei Pan. Estimating the global prevalence, disease progression, and clinical outcome of hepatitis delta virus infection. *Journal of infectious diseases*. 2019, doi.org/10.1093/infdis/jiz633
9. Wenhui Wang, **Shan Li**<sup>#</sup>, Pengyu Liu<sup>#</sup>, Kostandinos Sideras, Harmen J.G. van de Werken, Marieke van der Heide, Wanlu Cao, Marla Lavrijsen, Maikel P. Peppelenbosch, Marco Bruno, Qiuwei Pan, and Ron Smits. Oncogenic STRAP supports hepatocellular carcinoma growth by enhancing Wnt/ $\beta$ -catenin signaling. *Molecular Cancer Research*. 2019, 17(2):521-531. (<sup>#</sup> equal contribution)
10. Bin Xie, Renxin Xing, Ping Chen, Yuanbin Gou, **Shan Li**, Jing Xiao, and Jiahong Dong. Down-regulation of c-Met expression inhibits human HCC cells growth and invasion by RNA interference. *Journal of Surgical Research*. 2010, 162(2):231-238.
11. 谢斌, **李杉**, 刘宏鸣. 浅谈肝胆外科介入治疗技术多媒体课件的制作. 医学教育探索. 2010, 9(4):566-568.
12. **李杉**, 陈平, 金世龙, 皮儒先, 彭莉娜. 十二指肠憩室手术疗效的Meta分析. 中华消化外科杂志. 2009, 8(5):350-352.
13. 谢斌, 刘宏鸣, **李杉**. 肝胆外科介入治疗技术带教中存在的问题及对策. 西北医学教育. 2009, 17(5):1028-1029.
14. **李杉**. 胰腺实性假乳头状瘤致区域性门静脉高压一例. 中国现代普通外科进展. 2009, 12(12):1103-1104.
15. 皮儒先, **李杉**. 小儿巨大横结肠系膜囊肿1例. 中国医药导报. 2007, 4(36):154-155.
16. **李杉**. 巨大腹主动脉瘤破裂一例误诊. 临床误诊误治. 2007, 20(12):89-90.
17. **李杉**, 罗东林, 皮儒先. 十二指肠穿透性溃疡表现为反复右下腹痛1例. 中国医药导报. 2007, 4(32):79.
18. **李杉**, 皮儒先. 腹腔及胰头结核术前误诊为壶腹周围肿瘤1例. 肝胆胰外科杂志. 2007, 19(4):223.
19. **李杉**. 端粒酶抑制剂与肿瘤基因治疗的研究进展. 实用预防医学. 2005, 12(6):1513-1516.
20. **李杉**, 何念海. 新生儿痫性发作的诊断及治疗. 中国全科医学. 2005, 8(24):2073-2075.
21. **李杉**. 端粒G-四联体及其在肿瘤治疗中的应用. 生物学杂志. 2005, 22(4):4-7.

# PhD Portfolio

<b>Name PhD Student</b>	Shan Li
<b>Erasmus MC Department</b>	Gastroenterology and Hepatology
<b>PhD Period</b>	January 2015 – August 2020
<b>Promotor</b>	Prof.dr. Maikel P. Peppelenbosch
<b>Copromotor</b>	Dr. Ron Smits

## Seminars

---

2015-2020, Weekly MDL seminar program in experimental gastroenterology and hepatology (attending): 42 weeks/year; 1.5h (ETCS, 11.25)

2015-2020, Weekly MDL seminar program in experimental gastroenterology and hepatology (presenting): 2 times/year; preparation time 16h (ETCS, 5.70)

2015-2020, Weekly research group education (attending): 42 weeks/year; 1.5h (ETCS, 11.25)

2015-2020, Weekly research group education (presenting): 8 times/year; preparation time 8h (ETCS, 11.45)

## Conferences

---

2018, Dutch Digestive Disease Days, Veldhoven, the Netherlands (Oral presentation): preparation time 8h (ETCS, 0.30)

2017, 1<sup>st</sup> ACE-SCORE Day Symposium, Rotterdam, the Netherlands (Oral presentation): preparation time 8h (ETCS, 0.30)

2017, Promega Discover Glo Tour Symposium, Rotterdam, the Netherlands (attending) (ETCS, 0.30)

2015, 17<sup>th</sup> Gut Day Symposium, Rotterdam, the Netherlands (Oral presentation): preparation time 8h (ETCS, 0.30)



## **General Courses**

---

2016, Biomedical English Writing Course (ETCS, 2.00)

2016, Microbiomics I (ETCS, 0.60)

2016, Galaxy for NGS (ETCS, 1.00)

2015, CPO Course - Patient Oriented Research: design, conduct, analysis and clinical implications (ETCS, 0.30)

## **Reviewing for Scientific Journals**

---

Scientific Report: one paper (ETCS, 0.30)

Medical Science Monitor: one paper (ETCS, 0.30)

Canadian Journal of Gastroenterology and Hepatology: two papers (ETCS, 0.60)

Annals of Translational Medicine: four papers (ETCS, 1.20)

## **Academic Awards**

---

China Scholarship Council (CSC) Scholarship (File No. 201408060053)

## Curriculum Vitae

Shan Li was born on July 17, 1984, in Kaixian (now called Kaizhou District), Chongqing, China. He grew up and attended his primary, middle and high school in his hometown.

Between 2001 and 2006, he received his medical education and intern training at the Third Military Medical University (now called Army Medical University), China. In August 2006, he started work as a resident at the Department of Hepatobiliary Surgery of Daping Hospital, Army Medical University, China. In 2012, he qualified as an attending surgeon. In 2013, he came to the United Kingdom and joined the MSc Molecular Medicine (Cancer pathway) program at the University of Sheffield. His master thesis was done under the supervision of Dr. François Guesdon, and he obtained his MSc degree (with Merit) in October 2014.

In January 2015, with the support of the China Scholarship Council (CSC) scholarship, he started his PhD fellowship at the Department of Gastroenterology and Hepatology of Erasmus MC. Under the supervision of Prof. dr. Maikel P. Peppelenbosch, Dr. Sergey R. Konstantinov, and Dr. Ron Smits, he mainly focused on the microbiome and molecular characterizations of cancers in the digestive system. In January 2020, he was appointed Section Editor of the *Annals of Translational Medicine*. After his PhD, he is going to continue conducting research and perform clinical activities at the Department of Hepatobiliary Surgery of Daping Hospital, Army Medical University, China.



## Acknowledgements

The five years have elapsed in a flash since I started my PhD. I would like to thank everyone who helped me during the past five years.

I want to express my sincere gratitude to my promotor Prof. dr. Maikel P. Peppelenbosch, for whom I have great respect and admiration. Many thanks for providing me with the wonderful opportunity to learn and conduct research in the Erasmus MC, the largest and one of the most influential university medical centers in Europe, where I have gotten a good start on my research career. My warm-hearted thanks also to his insightful supervision and team leadership.

I want to express my heartfelt gratitude and appreciation to my copromotor Dr. Ron Smits, my respected supervisor, for continual guidance and perceptive advice throughout my PhD period. He enlarged my research vision, and I learned the scientific attitude and the spirit of a scientist from working with him. I am grateful for his constant support that develops my scientific thinking and laboratory skills. I appreciate him for giving me invaluable comments on writing scientific manuscripts. I also thank him for sharing his wisdom of life that inspires me.

I want to give sincere thanks to my previous supervisor Dr. Sergey R. Konstantinov for leading me into the fantastic field of microbiome research. I must acknowledge him for providing me the ideas of starting up microbiome works.

I am really grateful for Dr. Gwenny M. Fuhler. She taught me many useful tips that strengthen my presentation skills before my first oral presentation at a formal meeting. I also thank her for giving me the suggestions to improve my manuscripts.

I would like to thank Prof. dr. Luc van der Laan, Dr. Jaap Kwekkeboom, and Dr. Qiuwei Pan for their suggestions about my projects on the Gastro-oncology Friday meetings and the MDL seminars.

I am very thankful to Marla, Marcin Magierowski, and Katarzyna Magierowska for valuable efforts in my projects. I also show appreciation to all those who helped me to make these projects more successful. I wish you a bright and colorful future.

I would like to thank Leonie, Raymond, Marieke, Auke, Henk, Jan, Anthonie, Kelly, Natasha, Sharida, Paula, Gertine, Petra, and many others, for giving me valuable assistance with work-related support during the past five years. I wish you all the best.

I wish to thank my previous and present office-mates, Patrick, Lucia, Monique, Janine, Lisanne, Gulce, Aafke, Michiel, Rik, Elmer, Wesley, Guoying, Xiaolei, et al., for always being so friendly and having many fun times to erase the tedium of routine work. I wish you all the best for the future.

I would also like to thank previous and current colleagues Suk Yee, Yunpeng, Yining, Yingying, Junhong, Shaojun, Kairong, Manzhi, Shaoshi, Shihao, et al. for their help. My thanks also go to all members of staff in the Department of Gastroenterology and Hepatology. I wish you all the best.

I want to give special thanks to Wen, Buyun, Pengyu, Wenshi, Wanlu, Meng, Sunrui, Lei, Yuebang, Yijin, Xinying, Zhouhong, Wang Ling, Pengfei, Bingting, Shanshan, Li Yang, Peifa, Changbo, Kan, Yang Qin, Xumin, Yunlong, Ruyi. Not only for your help on my research work, but also for the companionship over my PhD life. Besides the academic discussions that really help me a lot, numerous funs we experienced together, e.g., traveling, cooking, drinking, having dinner, and BBQ. I cannot list all of these delightful things because there is so much to say, but I will never forget the wonderful time. Thanks for being present in one of the best periods of my life. I wish you all the best for your bright future.

To Zhijiang, I regret we didn't meet sooner! I really enjoy the time we spent together. We shared and doubled the joy we felt. We cheered up each other when we got disappointed. Wish you all the best of luck for a bright and prosperous future.

To Wenhui, thanks for your help with many experiments in the beginning of my works. Many thanks to you for always encouraging me. I wish you all the best and lots of good fortune to you and your family.

To Jiaye, many thanks for providing valuable information and tips on research work and various aspects of life including, but not limited to, travel and shopping. I am also thankful for sharing me the traditional foods you cooked and the snack foods you bought from our hometown, which made me feeling like I was back home. I wish you all the success and every happiness in the future.

To Menggang, we are very familiar with each other for 14 years to date in Chongqing. It was also a pleasure to work with and have a lot of wonderful memories

in the leisure time with you in Rotterdam. I wish all the best to you, your wife Huiling, and your daughter Xuanxuan.

I want to thank Zhihao, Wenqiu, Guannan, Lu Tao, Changbin, Wu Bin, Gao Ya, Ping Zhen, Zhanmin, Huang Ling, Zhou You, Zhang Shuang, Wenhao, Yao Yao, Liu Jun, Jinluan, Zhangling, Yan Lin, Sun Wei, et al., all my friends in the Erasmus MC. All the best for a bright future.

I also want to thank all my friends, Feng Xu, Xu Man, Gong Qiong, Jia Fan, Qu Tingting, and many others, who work outside of the Erasmus MC. Tons of best wishes to you.

To Li Ming. Your optimism made me laugh a lot. I appreciate when I finished up lab works late at night you often cook me the supper. I wish you a future filled with success and prosperity.

To Qiyu and Jing, Duanyang and Yanmin, I am grateful for the intriguing time our families spent together. All the wishes to you and your families.

To Xiangyu and Xuejin, thanks for taking a lot of time to have fun together since the beginning of my Rotterdam life. I consider myself very fortunate to have had the experience that we were having dinner, having BBQ, drinking, talking, and so on. Good luck and all the best to you and little Yiyi.

I am grateful to my friends in China. Yuan Lin, Cai Dong, Zhou Lin, Zhou Xiaobing, Hu Hao, Liao Wenle, Yuan Ye, Zhang Xiaohua, Cao Xiaofu, Yang Bin, Cheng Shiguang, Pi Ruxian, Zhou Yuyang, Chen Yin, Chen Hongxu, He Xu, Wang Wuxi, Lan Fujin, Yu Peng, Zhang Hong, Bai Fan, Li Wanfei, Li Shiming, et al. You make me feel the loss when you are not around me. However, many thanks for having lots of fun together when I was spending vacation time in China over the last five years. I will never forget the treasured time, nor will I forget our friendship. I wish nothing but the best for you and your family.

To my respected mentor, Prof. dr. Liu Hongming, thanks for the help during the period since I graduated from university to the present. My heart is fulfilled with gratefulness, and I consider myself very lucky to have the encouragement, inspiration, and influence from you. I wish all the best to you and your family.

I also appreciate several others whose names are too numerous to mention.

Lastly, I have to give my special thanks to my parents for the sacrifices they made for me throughout my life, and I appreciate their unconditional support. I also want to thank my wife Fang and my little princess Jiayue for making my life a very colorful one. I appreciate my parents-in-law for their support and understanding. I also want to acknowledge the rest of my family for their support and encouragement.