Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Medicine and Health Sciences
Department of Clinical and Molecular Medicine

Vibeke Sognnes

Fusobacterium nucleatum and the Epstein-Barr virus-encoded microRNA miR-BART10-3p in colorectal cancer

- validating the role of the genes *CSF2, CCL20, MAT2B and ELL2*

Master's thesis in Molecular Medicine

Supervisor: Robin Mjelle

May 2020

NTNU
Norwegian University of
Science and Technology

Vibeke Sognnes

Fusobacterium nucleatum and the Epstein-Barr virus-encoded microRNA miR-BART10-3p in colorectal cancer

- validating the role of the genes *CSF2, CCL20, MAT2B and ELL2*

Master's thesis in Molecular Medicine
Supervisor: Robin Mjelle
May 2020

Norwegian University of Science and Technology
Faculty of Medicine and Health Sciences
Department of Clinical and Molecular Medicine

NTNU
Norwegian University of
Science and Technology

Abstract

Colorectal cancer (CRC) is the third most commonly diagnosed and second most deadly form of cancer worldwide with more than 1.8 million new cases and 800,000 cancer deaths in 2018. Recent studies suggest that *Fusobacterium nucleatum* (*F. nucleatum*), an opportunistic anaerobe in the oral cavity, has a potential role in the development of CRC. Furthermore, recent analyses have revealed an enrichment of miR-BART10-3p, an Epstein-Barr-virus miRNA, in colorectal tissue. The present study aimed to investigate the role of *F. nucleatum* and miR-BART10-3p in CRC and determine if and how they contribute to tumour development.

The localization of *F. nucleatum* in the colon cancer cell line DLD-1 was examined by confocal microscopy. Furthermore, the direct effect of *F. nucleatum* was evaluated by co-culturing the bacteria with DLD-1 and look for changes in proliferation and migration. To identify genes that respond to *F. nucleatum*, gene expression analysis of DLD-1 co-cultured with *F. nucleatum* was performed and validated by RT-qPCR and ELISA. To identify human targets of miR-BART10-3p, transient transfection of the miRNA mimic in the colon cancer cell line SW620 was performed, and the regulatory effect of the miRNA on target genes was investigated by RT-qPCR and Luciferase Assays. Lastly, an *in vitro* procedure for establishing patient-derived CRC spheroids was developed for future studies of *F. nucleatum* and miR-BART10-3p in a physiological microenvironment closely resembling the *in vivo* conditions of a solid tumour.

F. nucleatum was confirmed to have an intracellular localization, but did, however, not promote cell migration or cell proliferation in DLD-1. Furthermore, RNA seq of *F. nucleatum*-treated DLD-1 led to the identification of *CCL20* and *CSF2*, which are important cytokines in the regulation of inflammation. A time-dependent upregulation of *CCL20* and *CSF2* mRNA, as well as a dose-dependent upregulation of CCL20 protein was observed. Furthermore, RNA seq of miR-BART10-3p-treated SW620 led to the identification of the cancer-related genes *MAT2B* and *ELL2*, both of which could be confirmed as potential targets of this miRNA using *in silico* prediction. Using *in vitro* methods, the gene *MAT2B* was validated to be a direct target of miR-BART10-3p. These results suggest that miR-BART10-3p may function as a tumor suppressor by downregulating *MAT2B*, a gene that has been shown to activate the ERK/AKT pathways in CRC, as well as an oncomiR by downregulating *ELL2*, a gene with tumor suppressor functions in prostate epithelial cells. In conclusion, the present study provides further insight into the role of *F. nucleatum* and miR-BART10-3p in CRC, and the developed procedure for making CRC spheroids enables further investigation of their underlying mechanisms in a colorectal tumour.

Sammendrag

Kreft i tykktarm- og endetarm, kolorektal kreft (CRC), er den tredje hyppigste diagnostiserte og den nest mest dødelige kreftformen i verden med over 1.8 millioner nye tilfeller og 800,000 nye dødsfall i 2018. Nyere forskning tyder på at *Fusobacterium nucleatum* (*F. nucleatum*), en opportunistisk, anaerob bakterie i munnhulen, har en potensiell rolle i utviklingen av CRC. Videre har nyere forskning avdekket økt uttrykk av miR-BART10-3p, et Epstein-Barr-virus miRNA, i tykktarmsvev. Målet med denne studien var å undersøke rollen til *F. nucleatum* og miR-BART10-3p i CRC og bestemme om og hvordan de bidrar til utviklingen av kreft.

Lokalisering av *F. nucleatum* i koloncellelinjen DLD-1 ble undersøkt med et konfokalmikroskop. Videre ble den direkte effekten av *F. nucleatum* i CRC undersøkt ved å dyrke bakterien med DLD-1 og se etter endringer i cellevekst og migrasjon. For å identifisere gener som responderer på *F. nucleatum* ble en genekspresjonsanalyse av *F. nucleatum*-behandlet DLD-1 utført og validert ved hjelp av RT-qPCR og ELISA. For å identifisere humane mål-mRNA for miR-BART10-3p ble kunstig syntetisert miR-BART10-3p gitt til koloncellelinjen SW620, og den regulatoriske effekten av miRNA på målgenene ble undersøkt ved RT-qPCR og luciferase-baserte assays. Til slutt ble en *in vitro*-prosedyre for å etablere pasient-deriverte CRC-sfæroider utviklet for fremtidige studier av *F. nucleatum* og miR-BART10-3p i et fysiologisk, mikrobielt miljø som ligner *in vivo*-forholdene til en solid svulst.

F. nucleatum ble bekreftet å ha en intracellulær lokalisering i DLD-1, men fremmet imidlertid ikke cellevekst eller migrasjon. Videre førte RNA-seq av *F. nucleatum*-behandlet DLD-1 til identifisering av *CCL20* og *CSF2*, to viktige cytokiner i reguleringen av infeksjoner. Det ble observert en tidsavhengig oppregulering av *CCL20*- og *CSF2*-mRNA, i tillegg til en tids- og doseavhengig oppregulering av CCL20-protein. Ved hjelp av RNA-seq av miR-BART10-3p-behandlet SW620 identifiserte vi nedregulering av to gener, *MAT2B* og *ELL2*, som begge i tillegg kunne bekreftes som potensielle mål-mRNA for miR-BART10-3p ved bruk av *in silico* prediksjon. *In vitro* RT-qPCR validering av disse genene bekreftet at miR-BART10-3p nedregulerer *MAT2B* og *ELL2*, og luciferase-baserte assays bekreftet at *MAT2B* er et direkte mål-mRNA. Disse resultatene antyder at miR-BART10-3p kan fungere som en tumorsuppressor ved å nedregulere *MAT2B*, et gen som har vist seg å aktivere ERK/AKT-signalveiene i CRC, i tillegg til et onkomiR ved å nedregulere *ELL2*, et gen med tumor suppressor aktivitet i epitelcellene i prostata. Alt i alt gir den nåværende studien bedre innsikt i rollen til *F. nucleatum* og miR-BART10-3p i CRC, og den utarbeidede prosedyren for å lage sfæroider muliggjør ytterligere utredning av deres underliggende mekanismer i tykktarmen.

Acknowledgements

This master's thesis is part of a Master of Science in Molecular Medicine at the Faculty of Medicine and Health Sciences (MH) at the Norwegian University of Science and Technology (NTNU) in Trondheim, Norway. The laboratory work was conducted at the Department of Cancer Research and Molecular Medicine, Faculty of Medicine, Laboratory Centre, 5th floor.

First, I would like to express my sincere gratitude to my main supervisor, Postdoctoral Researcher Robin Mjelle for giving me the opportunity to work on this interesting project, for teaching me laboratory techniques, and for giving me valuable feedback, support and guidance throughout the writing of my thesis. I am also grateful to my co-supervisors Eva Hofsli, for giving me feedback during the writing period, and to Pål Sætrom for giving me insight into this field of research. Many thanks to Per Arne Aas for always being available to assist in the lab, to Kjartan Wøllo Egeberg and Bjørnar Sporsheim for giving me training on Zeiss LSM 510 META Confocal Microscope and Invitrogen™ EVOS™ FL Auto Imaging System, and to the Genomics Core Facility in Trondheim and GATC biotech in Konstanz, Germany for their sequencing services.

I would also like to thank the Systems Biology for Oncology group at the Department of Clinical and Molecular Medicine at NTNU for giving me the opportunity to be a part of the ongoing spheroid project. In particular, thanks to Evelina Folkesson for help with experiments, for taking the time to answer my questions and for continuous feedback during the writing period.

At last I want to thank my fellow student Julie Jerud for cooperation on laboratory experiments, making me laugh on a daily basis, and for giving me motivation on occasions when I felt I had hit a brick wall. Without you, this process would have been much more difficult, and I am grateful I got to share this experience with you.

Trondheim, May 2020

Vibeke Sognnes

Table of contents

List of Figures

List of Tables

List of Abbreviations

APC	Adenomatous polyposis coli
AGO2	Argonaute 2
BART	BamHI-A region rightward transcript
BHRF	BamHI fragment H rightward open-reading frame
BRAF	B-Raf proto-oncogene serine/threonine kinase
CCL20	Chemokine (C-C motif) ligand 20
CCR6	Chemokine (C-C motif) receptor type 6
CRC	Colorectal cancer
CIMP	CpG island methylator phenotype
CIN	Chromosomal instability
CSF2	Colony stimulating factor 2
CTOS	Cancer Tissue-Originated Spheroid
DC	Dendritic cell
DCC	Deleted in colon cancer
EBV	Epstein Barr virus
ECM	Extracellular matrix
E. coli	*Escherichia coli*
ELL2	Elongation Factor for RNA Polymerase II 2
F. nucleatum	*Fusobacterium nucleatum*
GC	Gastric carcinoma
GCF	Genomics Core Facility
H. pylori	*Helicobacter pylori*
KRAS	KRAS proto-oncogene GTPase
LPS	Lipopolysaccharides
MAT2B	Methionine adenosyltransferase 2 subunit beta
MCS	Multicellular cancer spheroid
miRNA	microRNA
MLL	Mixed-lineage leukemia
MM	Multiple myeloma

MOI	Multiplicity of infection
mRNA	Messenger RNA
MSI	Microsatellite instability
NPC	Nasopharyngeal carcinoma
NK cells	Natural killer cells
piRNA	Piwi-interacting RNA
Pre-miRNA	Precursor miRNA
pri-miRNA	Primary miRNA
RB	Reconstitution buffer
RISC	RNA-induced silencing complex
RNAi	RNA interference
RT-qPCR	Reverse transcription-quantitative polymerase chain reaction
SAM	S-adenosylmethionine
SEC	Super elongation complex
siRNA	Short interfering RNA
sncRNA	Small non-coding RNA
Th17	IL17 producing helper T cell
TNF	Tumour necrosis factor
TP53	Tumour protein p53
Treg	Regulatory T cell
UTR	Untranslated region

1. Introduction

1.1. Colorectal cancer – incidence, survival, treatment and risk factors

Colorectal cancer (CRC), a collective term that includes both colon cancer and rectal cancer, is a major clinical and public health concern[1]. In Norway, CRC is the second most diagnosed cancer for both sexes, after breast cancer in women and prostate cancer in men[2]. In 2018, 4,428 new cases were diagnosed, in which 69% were diagnosed with colon cancer and 31% were diagnosed with rectal cancer[2]. Both men and women are equally at risk for colon cancer, whereas the risk of rectal cancer is slightly higher in men[2]. CRC caused 779 deaths among women and 822 deaths among men, representing 14% of cancer deaths in Norway and the second and third most frequent cause of cancer death among each sex, respectively[2]. Worldwide, CRC was the third most commonly diagnosed form of cancer in 2018 with about 1,800,977 new cases, representing 10% of all cancer diagnosis[3]. Among these, 61% were diagnosed with colon cancer and 39% were diagnosed with rectal cancer[3]. Furthermore, CRC caused 474,606 deaths among men and 387,057 deaths among women, representing the second most deadly cancer worldwide[3].

Colon cancer and rectal cancer share most risk factors, and there is a strong support that high consumption of red and processed meat, heavy alcohol use, smoking and being overweight increases the risk of these cancers[4,5]. Non-modifiable risk factors associated with higher CRC risk include inflammatory bowel disease, a family history of CRC, and increasing age[5,6]. Approximately 50% of cases of CRC are diagnosed in patients 70 years or older[2,7]. However, CRC is a complex and multifactorial disease resulting from multiple interactions between lifestyle, hereditary, genetic, epigenetic and environmental factors[8].

The primary symptoms of CRC are abdominal pain, rectal bleeding and changes in bowel habits[9]. Later symptoms may include anemia and intestinal obstruction[6]. However, the symptoms in CRC are vague which often result in late-stage detection[6,9]. The American Joint Committee on Cancer (AJCC) has divided CRC into four distinct stages (Stage I-IV) based on the TNM staging system, i.e., the size and extent of the tumour (T), the spread to nearby lymph nodes (N), and the spread (metastasis) to distant organs (M), in which increasing stage corresponds to a more advanced disease[10]. About 39% of cases of CRC are detected at an early, localized stage (Stage 0-I), for which the 5-year relative survival is about 90%[11,12]. If the cancer has spread to nearby tissue (Stage II) or to the nearby lymph nodes (Stage III), the 5-year relative survival is about 71%[11,12]. If the cancer has spread to distant organs (Stage IV), the 5-

year relative survival is about 14%[11,12]. However, for patients with metastasis localized to only a few sites and with few lesions, radical surgery is possible[11]. Therefore, the 5-year relative survival for these patients have improved considerably[11].

Surgery is often sufficient for tumours that have not spread to distant sites (Stage 0-III), but if the tumour is large or has spread to the lymph nodes, chemotherapy is usually given after surgery to suppress secondary tumour formation (adjuvant chemotherapy)[13]. Chemotherapy treatment involves the use of drugs to stop cancer cells from continuing to divide uncontrollably[14]. A drug called 5-fluorouracil (5-FU) is the most common CRC chemotherapy treatment; for decades following the discovery of chemotherapy, 5-FU was the only chemotherapeutic agent available to successfully improve 12-month survival in CRC patients[14]. Advanced CRC (Stage IV) typically require chemotherapy and/or targeted therapy to control the cancer[13]. The aim of targeted therapy is to attack specific genes or proteins that contribute to cancer growth and survival[15]. Furthermore, radiation therapy is often given at the same time as chemotherapy to increase the effectiveness of the therapy or to relieve symptoms such as pain[13]. Radiation therapy uses ionizing radiation to exterminate malignant cells and to shrink tumours before surgery[16]. Surgery is unlikely to cure advanced CRC but might still be needed to relive a blockage in the colon or for other conditions to improve the symptoms[13].

1.2. Molecular basis of colorectal cancer

CRC is a heterogenous disease originating from the epithelial cells lining the colon or rectum of the gastrointestinal tract[1]. Genetic and epigenetic factors involved in the progression of CRC are caused by three major pathways: (I) Chromosomal instability (CIN) which is recognized by accumulation of mutations in specific genes, (II) CpG island methylator phenotype (CIMP) which is characterized by simultaneous hypermethylation of numerous promoter CpG island sites, resulting in inactivation of tumor suppressor genes, and (III) microsatellite instability (MSI) which is mutations in DNA mismatch repair genes[17]. CIN accounts for 85% of cases of CRC and includes mutations in the tumour suppressor gene adenomatous polyposis coli (*APC*), which promotes the onset of CRC by activating the Wnt/β-catenin signalling pathway[18]. The Wnt/β-catenin signalling pathway is highly conserved and plays a critical role in regulating cell proliferation, differentiation, migration, genetic stability and apoptosis, transforming normal colorectal epithelium to early adenoma[19]. This is followed by mutations in KRAS proto-oncogene GTPase (*KRAS*) and B-Raf proto-oncogene serine/threonine kinase (*BRAF*) in the adenomatous stage, resulting in a deregulated RAS/MAPK signalling pathway[18]. Abnormal

RAS/MAPK signalling may lead to increased or uncontrolled cell proliferation and resistance to apoptosis[20]. Eventually, inactivation of the tumour suppressor genes tumour protein p53 (*TP53*) and deleted in colon cancer (*DCC*) occur during the transition to malignancy[18,21,22]. Mutations of the *DCC* gene result in the absence of the netrin-1 receptor. As a result, the netrin-1 receptor is not available to induce apoptosis or cell cycle arrest, resulting in uncontrolled cell proliferation[23]. *TP53* is activated in response to cellular stress, and leads to cell cycle arrest, apoptosis or DNA repair. Upon mutation, more mutations will be accumulated in the cell and eventually lead to cancer development[24]. This is a well-established multistep genetic model presented by Fearon and Vogelstein[25] based on the understanding that CRC is the result of sequential accumulations in epigenetic and genetic changes, resulting in the progression from a normal cell to CRC (Figure 1). In the model of Fearon and Vogelstein, CRC is developing due to alterations of multiple genes in different pathways, such as *TP53*, *APC*, *KRAS*, and many more genes, which act in different molecular processes and together promote CRC carcinogenesis[26].

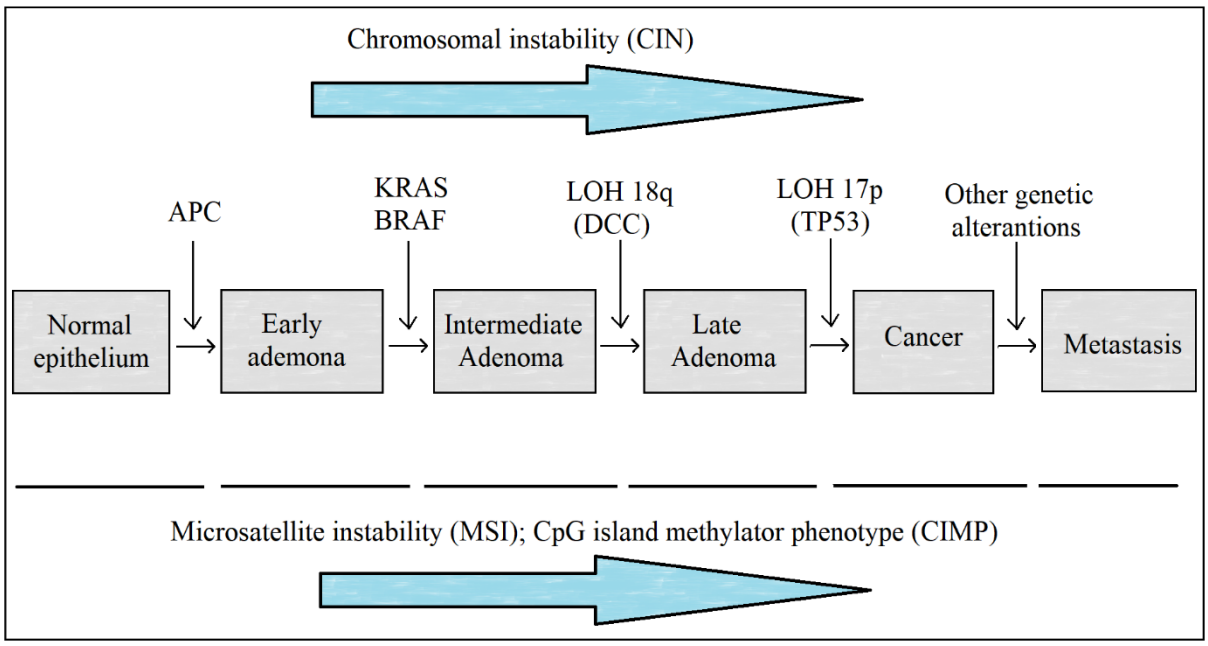


Figure 1. *Adenoma-carcinoma sequence proposed by Fearon and Vogelstein[22] in 1990. The adenoma-carcinoma sequence is caused by three major pathways: Chromosomal instability (CIN), CpG island methylator phenotype (CIMP), and microsatellite instability (MSI). CIN is recognized by the accumulation of mutations in specific genes. APC mutation is generally considered as the initial event transforming normal epithelium to adenoma. This is followed by sequential mutations in other genes, eventually resulting in the progression from a normal cell to a metastatic tumour. APC, adenomatous polyposis; KRAS, KRAS proto-oncogene GTPase; BRAF, B-Raf proto-oncogene serine/threonine kinase; LOH, loss of heterozygosity; DCC, DCC netrin 1 receptor; TP53, tumour protein 53. Adapted from Nguyen, H. T., and Duong, H., 2018[18] and Martínez, J. D. et al., 2003[27].*

1.3. Infectious agents in cancer

Approximately 20% of human cancers are caused by infectious agents, including bacteria, viruses and parasites[28]. Some infectious agents can disrupt signalling that normally keeps cells from growing and proliferating in an uncontrolled way[29]. Also, some infectious agents, such as the human immunodeficiency virus (HIV), weaken the immune system, making it harder for the body to fight off other cancer-causing infections[28,29]. A more direct mechanism involves expression of oncogenes, such as the human papillomavirus (HPV) and its expression of E6 and E7 in cervical cancer[30]. Additionally, some infectious agents cause chronic inflammation, such as *Helicobacter pylori (H. pylori)* in the development of gastric cancer[28,29]. Infectious agents have been acknowledged and listed as risk factors in a number of cancers, but despite the extensive amount of research, the association between certain bacterial and viral infections and the risk of CRC is not well described. For this reason, the present study focused on infectious agents which previously has been associated with CRC, namely *Fusobacterium nucleatum (F. nucleatum)*[43,44] and the Epstein-Barr virus (EBV)[86], and their potential role in the development of CRC.

1.4. Gut microbiota and colorectal carcinogenesis

Microbial communities are established at birth, and a lifelong symbiotic and mutualistic relationship is formed[31]. The gastrointestinal tract is colonized by more than 100 trillion bacteria, and the gut, particularly the colon, is the host of approximately 10^{14} bacteria and more than 1000 bacterial species[32,33]. At first, the belief that the microbial community was solely beneficial and not harmful to the host was a generally accepted and well-established scientific knowledge[33]. For instance, the gut microbiota is involved in food metabolism, vitamin production and waste processing[34,35]. Furthermore, the gut microbiota is involved in the development and function of the mucosal immune response, prevents colonization with pathogenic microbes, and helps maintain the physiological microenvironment[33]. In return, the human provides a place to live and feed[33-35]. However, a contributory finding which led researchers to suggest a potential link between certain bacteria and CRC was the significant 12-fold higher risk of developing colon cancer compared to cancer in the small intestine. This risk was related to the significant differences in the amount of bacteria in colon versus the intestine, in which the colon contained 10^{12} bacteria per mL and the intestine 10^{2} bacteria per mL[33]. Over the last decades, a large number of clinical trials have been presented, indicating that gut microbiota enhances tumour growth via various mechanisms and thereby disrupting the

homeostatic balance[36-38]. Several bacteria, including *Bacteroides fragilis*[39], *Streptococcus bovis*[40], *H. pylori*[41], *Enterococcus faecalis*[42], *F. nucleatum*[43], and *Streptococcus gallolyticus*[40] are reported to have increased expression in CRC tissue. It is therefore important to understand the interactions between the gut microbiota and the host to provide personalized therapy and increase the efficacy of current treatment.

1.4.1. *Fusobacterium nucleatum (F. nucleatum)*

F. nucleatum is a gram-negative obligate anaerobe bacterium in the oral cavity and plays a role in several oral diseases, including periodontitis and gingivitis[38]. Recent analyses have revealed an enrichment of *F. nucleatum* in human CRCs and adenomas compared with adjacent normal tissue[43,44]. Furthermore, increased levels of *F. nucleatum* correlate with CIMP, MSI, and mutations in *KRAS, BRAF* and *TP53*[38,45]. Attachment of *F. nucleatum* to cell surfaces is mediated by several mechanisms (Figure 2). Abed et al. has identified that Fap2, an outer membrane protein on *F. nucleatum*, mediates adenocarcinoma-specific binding through attachment to the host polysaccharide Gal-GalNAc, which is overexpressed in CRC[46]. In addition, binding between the fusobacterial adhesin FadA and host epithelial E-cadherin may enable fusobacterial attachment, and lead to activation of the Wnt/β-catenin signalling pathway, thus promoting cell proliferation and oncogenic responses[47]. It has been reported that *F. nucleatum* can inhibit tumour killing by natural killer (NK) cells via their Fap2 protein which is able to interact with TIGIT (T cell immunoreceptor with immunoglobulin (Ig) and ITIM domains), leading to the inhibition of NK cell cytotoxicity[48]. In addition, it has been shown that *F. nucleatum* promotes CRC resistance to chemotherapy by targeting Toll-like receptor 4 (TLR4) with bacterial lipopolysaccharides (LPS), followed by MYD88 innate immune signalling and specific microRNAs (miRNAs) to activate the autophagy pathway[49]. LPS also breaks the intestinal barrier and facilitates the entry into epithelial cells[50].

F. nucleatum also display an immunosuppressive effect, such as its significant positive correlation with the mucosal proinflammatory cytokines IL-6, IL-12, IL-17 and TNF-α[51], which is consistent with increased activation of the cancer-associated nuclear factor kappa B (NF-κB)[52]. Furthermore, binding and cellular invasion of *F. nucleatum* to CRC cells selectively induces the secretion of the proinflammatory and metastatic cytokines IL-8 and CXCL1, which further induces migration of CRC cells[53]. Finally, it has been reported that *F. nucleatum* can release short-peptides and short-chain fatty acids which selectively attracts myeloid-derived suppressor cells (MDSCs) and suppress T-cell activity through multiple mechanisms[38]. Based

on all these findings, *F. nucleatum* not only localizes to and is enriched in CRC but may also directly and indirectly modulates immune and cancer cell signalling and migration.

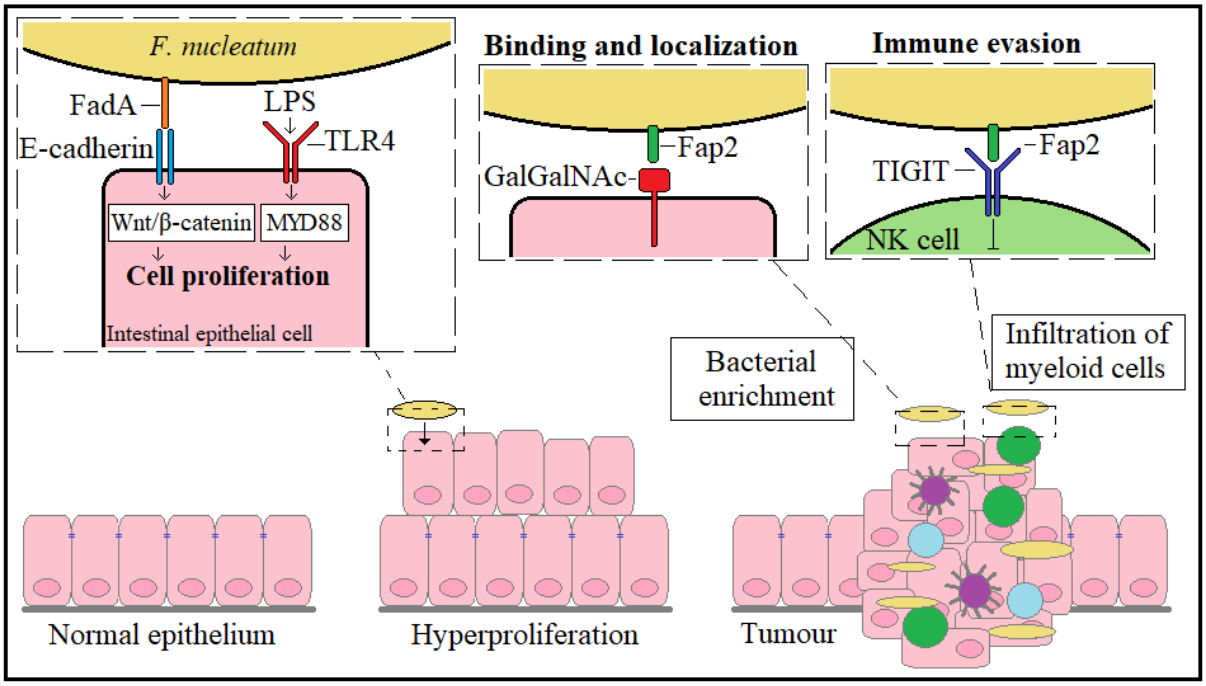


Figure 2. *Mechanisms of F. nucleatum surface proteins in CRC. FadA activates the E-catenin/β-catenin signalling pathway, which contributes to cell proliferation. Lipopolysaccharides (LPS) activate the MYD88/NF-κβ pathway through toll-like receptor 4 (TLR4), leading to an increase in several inflammatory factors. Gal-GalNAc is the receptor of Fap2 and recruits the bacteria to the tumour site. In addition, Fap2 binds TIGIT leading to the inhibition of NK cell cytotoxicity. Adapted from Brennan and Garrett, 2019*[54].

1.5. RNA interference

RNA interference (RNAi) is a process of gene silencing mediated by small non-coding RNAs (sncRNAs)[55]. Several types of small RNA molecules function in RNAi. The first type is the double-stranded short interfering RNA (siRNA) that cleaves messenger RNAs (mRNAs), thereby effectively silencing the expression of its target genes[56]. Short interfering RNAs are not encoded by the human genome, but are common in other species, in particular plants[57]. A second type of sncRNA is the Piwi-interacting RNA (piRNA). This RNA is specialized in silencing transposable elements and plays an important role in the testes[58]. The third one, which is the primarily focus in the present study, is the miRNA. MicroRNA molecules are naturally occurring single-stranded RNAs and come from endogenous non-coding RNAs found within the introns of larger RNA molecules[56]. MicroRNAs regulate gene expression by base-pairing to miRNA recognition elements (MREs) located on the mRNA 3' untranslated regions (UTR). Thus, it triggers translational inhibition and degradation of the target gene[55].

The major steps in the RNAi pathway are similar for both siRNAs and miRNAs, however, the biogenesis of siRNAs is generally more complex than that of miRNAs, in particular in plants, involving several different proteins[59]. In contrast, the germ-line specific piRNA pathway differs considerably from that of siRNA and miRNA, and little is known about piRNA biogenesis and their mode of action[60]. MicroRNAs are first produced as long primary miRNA (pri-miRNA) molecules by RNA polymerase II in the nucleus[61] (Figure 3). Before leaving the nucleus, the single-stranded ends of the pri-miRNA are cleaved by the nuclear microprocessor complex formed by the RNase III enzyme Drosha and the DiGeorge critical region 8 (DGCR8) protein,

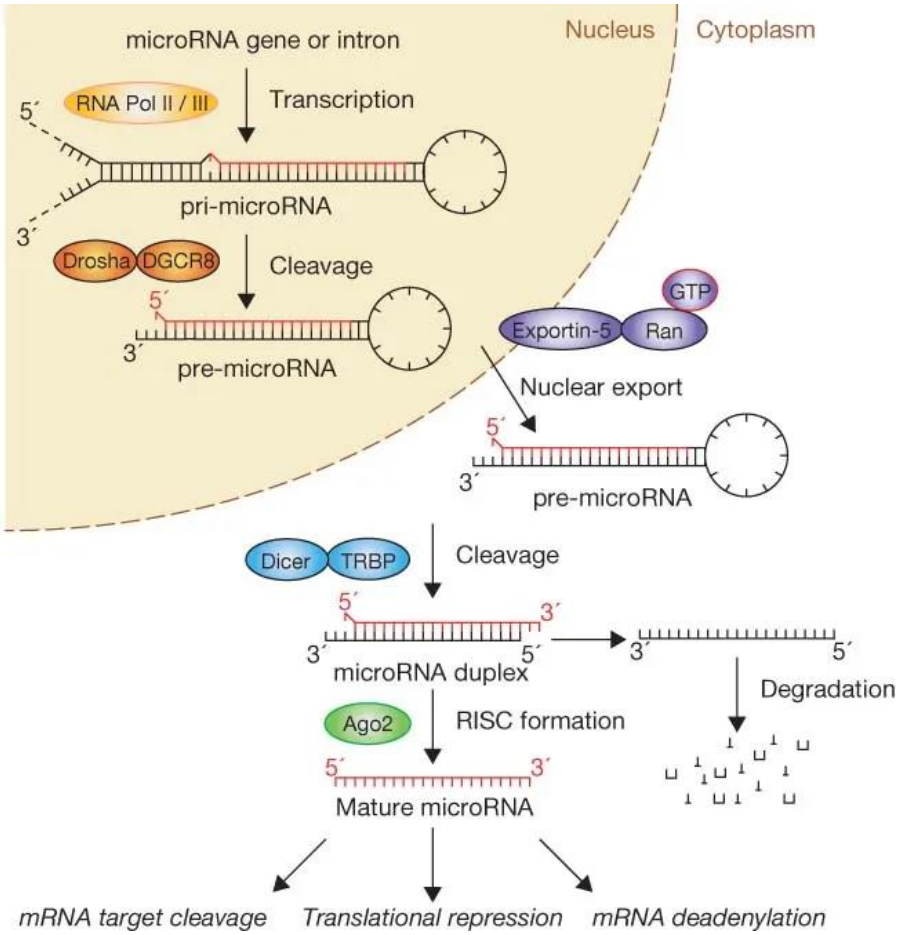


Figure 3. *RNA interference pathway. Long primary miRNAs (pri-miRNAs) are synthesized in the nucleus. Pri-miRNAs are processed by the microprocessor complex Drosha–DGCR8 to form precursor miRNAs (pre-miRNAs). The pre-miRNAs are then exported to the cytoplasm by Exportin-5–Ran-GTP, and further processed by the RNase Dicer in complex with the double-stranded RNA-binding protein TRBP to produce microRNAs (miRNAs). miRNAs are incorporated into the RISC complex and the "passenger" strand is cleaved and released by AGO2. The retained strand is used as a template by the RISC complex and binds to complementary mRNA sequences, resulting in the enzymatic cleavage or translational inhibition of the complementary target mRNA. Figure from Winter et al., 2009[61].*

resulting in a precursor miRNA (pre-miRNA)[62]. The pre-miRNA is exported to the cytoplasm by the small nucleic acid exporter Exportin 5 in complex with Ran-GTP[61]. After export, the pre-miRNA is processed by an RNase called Dicer in complex with the double-stranded RNA-binding protein TRBP, which cleaves the terminal loop of the pre-miRNA, leaving a double stranded RNA duplex[61]. One of the strands of the duplex, often referred to as the guide strand, is incorporated into the RNA-induced silencing complex (RISC), and the other strand, often referred to as the passenger strand, is usually degraded. The strand which is loaded into RISC is referred to as the mature miRNA[62]. The human genome encodes four different argonaute proteins which can bind miRNAs and form a RISC[63]. Argonaute 2 (Ago2) is the most highly expressed argonaute protein in humans and is the only human argonaute with an active catalytic domain for cleavage activity[61]. The function of the mature RISC complex is to bind complementary mRNA sequences, usually in the 3'UTR of the mRNAs, resulting in enzymatic cleavage, degradations or translational inhibition of the complementary target mRNA[61]. In contrast to siRNAs which show perfect complementarity and immediate cleavage of the target mRNAs[59], the miRNA binding does not need to involve perfect complementarity. Instead, the extent of complementarity influences how the subsequent RNA interference will transpire[61].

1.5.1. MicroRNA and its function in colorectal cancer

MicroRNAs are non-coding, highly conserved, single-stranded RNA molecules about 22 nucleotides in length[64]. Single miRNAs may regulate multiple targets, and single targets may be regulated by multiple miRNAs[65]. Because miRNAs can inhibit gene expression, they play important roles in human cancers[66,67]. For example, they may act as oncomiRs and promote cancer development by downregulating tumour suppressor genes or other genes involved in cell differentiation. Similarly, they may act as tumour suppressor miRNAs and downregulate different proteins with oncogenic activity[66,67].

MicroRNAs have altered expression profiles in CRC and are therefore promising to use as biomarkers[68]. In CRC, altered miRNA expression is shown especially in MSI tumours, which accounts for 15% of all cases[69]. Additionally, specific functions of miRNAs in CRC-associated pathways have been identified, such as inactivation of KRAS[70,71] and APC[71,72] which are major initiating events in colorectal carcinogenesis, resulting in a deregulated RAS/MAPK- and Wnt/β-catenin pathway. Furthermore, there has been identified a miRNA profile that can predict and differentiate among CRC metastasis[73]. By targeting cellular or viral genes, these miRNAs are involved in the regulation of multiple cellular responses such as host cell

proliferation, apoptosis, and immune escape. Thus, miRNAs are thought to contribute to the development of CRC.

1.6. Epstein-Barr virus (EBV)

EBV was the first identified human oncovirus discovered in 1964 in Burkitt's Lymphoma, and since then in several other types of cancers[74-76]. EBV is a double-stranded DNA virus belonging to the family *Herpesviridae*, subfamily *Gamaherpesvirinae*, genus *Lymphocryptovirus* and species *Human herpesvirus 4*[77]. The virus is the primary cause of infectious mononucleosis and has been estimated to infect more that 90% of the world's population by adulthood[78]. However, if the initial infection occurs during childhood, it is often asymptomatic[78]. Transmission of EBV is primarily via saliva but may also be spread through blood transfusion or as a result of organ transplantation[79]. Upon initial infection, EBV infects oropharyngeal epithelial cells by direct fusion of the viral envelope with the cell plasma membrane[80]. The virus enters lytic replication and spreads through the epithelium[81]. Following the initial infection, EBV infects circulating B cells through the interaction of the viral glycoprotein gp350 to CD21 (cluster of differentiation 21). Then, the viral glycoprotein gp42 interacts with the major histocompatibility complex (MHC) class II to initiate entry[80]. Interaction of gp42 and MHC class II triggers fusion of the core herpes-virus fusion machinery consisting of gB and the heterodimer gH/gL with the endosomal membrane[80]. Ultimately, the virus persists in a latency state for the lifetime of the host[81]. During latent infection, the EBV genome exists as a circular non-integrated episome and replicates by recruiting the cellular replication machinery[76]. For lytic reactivation to occur, the viral genome must linearize[76]. Reactivation and production of new viral particles can be induced when B cells differentiate into proliferating blasts[82]. The viral particles are epitheliotropic and can infect epithelial cells to establish latency[82]. As the epithelial cells differentiate, viral particles replicate and are released into saliva for transmission to a new host[82].

1.6.1. Epstein-Barr virus-encoded microRNAs

EBV miRNAs are expressed in all phases of the viral life cycle and are transcribed and generated in the same way as cellular miRNAs[83]. As many as 25 EBV miRNA precursors and 44 mature EBV miRNAs have been identified, four of which are encoded from the BamHI fragment H rightward open-reading frame (BHRF) region and the remainders are from the BamHI-A region rightward transcript (BART) region[84]. Expression of EBV miRNAs differs depending on three types of viral latency (I, II and III)[83]. BHRF miRNAs are highly expressed

in latency III and lytic replication-infected cells, such as B lymphoma cells, but are almost undetectable in cells under latency I and II, such as nasopharyngeal carcinoma (NPC) cells and gastric carcinoma (GC) cells[84,85]. BART miRNAs occur in all types of latency and are abundantly expressed in epithelial tumour cells harbouring the virus in type I and II latency, suggesting that EBV-BART miRNAs may contribute to the development of epithelial malignancies[84,85]. A recent study shows that EBV miRNAs, most of which are encoded from the BART loci, target host mRNAs and genes involved in CRC development[86]. For instance, miR-BART10-3p, which previously has been associated with NPC[87] and GC[88], was significantly elevated in CRC tissue.

1.7. Three-dimensional cultures and patient-derived colorectal cancer spheroids

The extracellular matrix (ECM) environment *in vivo* provides a variety of biophysical properties and biochemical cues that are essential in the tumour environment[89]. For instance, ECM topography, composition, permeability, mechanical rigidity and spatial organization affect cancer cell proliferation, differentiation, invasion, and metastasis, as well as tumour response to therapy[89]. In a normal colon, epithelial cells receive important survival signals from the ECM and undergo rapid apoptosis and clearance as soon as they lose their cell-matrix interaction (anoikis)[90]. Although resistance to anoikis is a crucial step during tumorigenesis and in particular during the metastatic spreading of cancer cells, most of the epithelial cells in solid tumours depend on cell-matrix interactions for their survival[91,92]. In addition, it is reported that E-cadherin mediated cell-cell contact, accompanied by AKT activation, is crucial for the survival of cancer cells[92].

Two-dimensional (2D) monolayer cultures are the standard *in vitro* model to study cancer[93]. However, cells inhabiting a flat solid surface are stretched and undergo cytoskeletal adjustments, because they lack exposure to the ECM, which may produce artificial polarity and cause abnormal gene and protein expression[93]. In contrast, a three-dimensional (3D) *in vitro* cancer model allows cell-cell and cell-matrix interactions that closely mimic the environmental conditions of the original tumour (Figure 4A), and has given great interest for a wide variety of diagnostic and therapeutic applications[94,95]. Figure 4B shows the processes of spheroid formation beginning with an interplay of the transmembrane receptor integrin with ECM,

leading to cell aggregation and later formation of compact spheroids through cadherin-cadherin interactions.

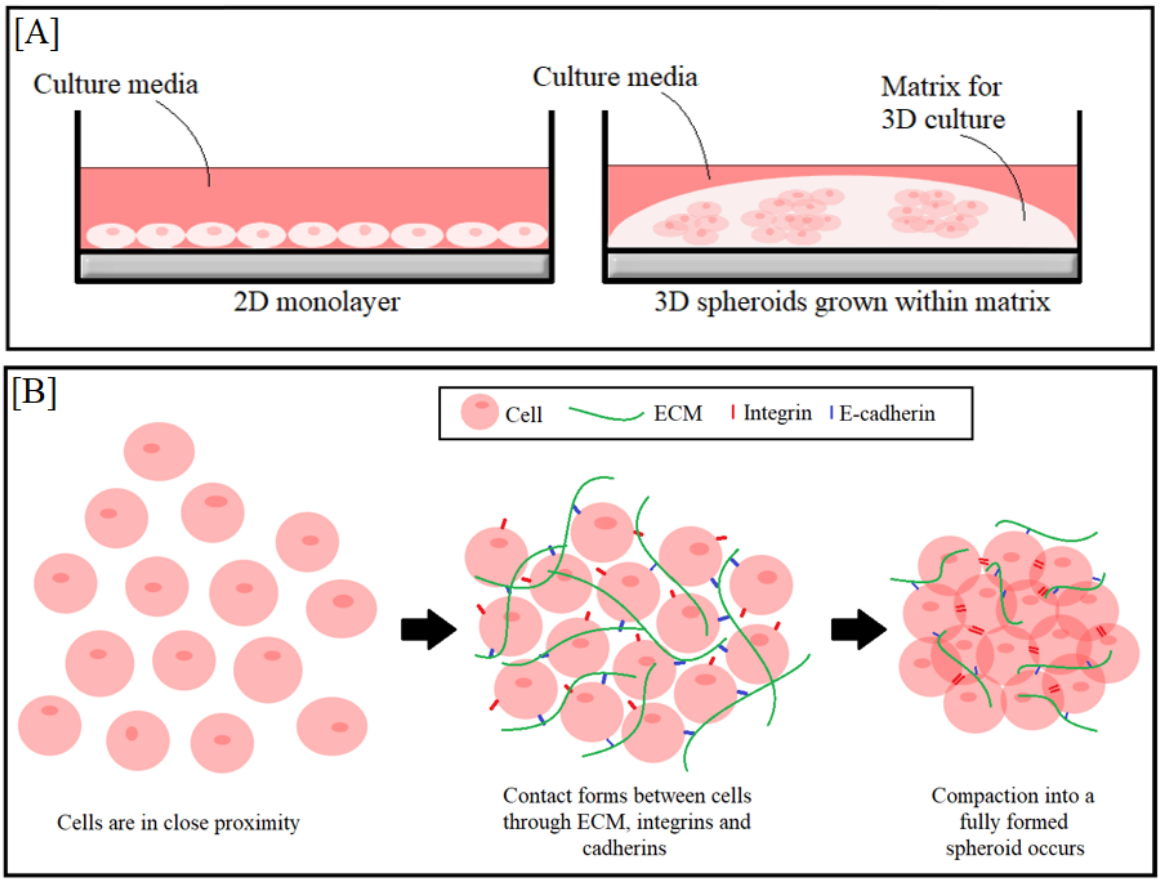


Figure 4. *Three-dimensional spheroid culture.* ***[A]*** *3D cell culture application of spheroids grown within a matrix compared to the traditional 2D monolayer.* ***[B]*** *The processes of spheroid formation starting with an interplay of integrin with ECM (extracellular matrix), leading to cell aggregation and later compaction into fully formed spheroids through cadherin-cadherin interactions. Adapted from Ibidi, n.d.*[96] *and Gionet-Gonzales1 and Leach, 2018*[97]*.*

Multicellular cancer spheroids (MCSs) are micro-sized cellular aggregates derived from individual patients and have been shown to successfully recapitulate the architectures and distinctive functions of the original tumour, even after long-term expansion[98,99]. MCSs are carried out from cancer cell lines or primary cells, which may be combined with fibroblasts, endothelial, or immune cells[89]. Recently, cancer cells derived from several different tumour types, such as colon[100], breast[101], pancreas[102], liver[103] and prostate[104], have been cultured in 3D conditions using a collagen-based matrix or other extracellular matrix components, such as Matrigel, which support attachment, survival and *in vivo*-like 3D growth[105]. In the present study, a method for establishing patient-derived CRC spheroids are being developed based on

a culture method developed by Kondo et al[92] for primary CRC cells, in which cell-cell interactions were maintained throughout the process.

Spheroids are a powerful tool to predict patient drug responses and support the personalization of treatment. In addition, spheroids can be applied to study gastrointestinal diseases and host-microbe interactions in the intestine[106,107]. Furthermore, the spheroids provide a model to investigate the response of the intestine to the presence of bacteria and to study the bacteria during the early steps of pathogen invasion[32].

1.8. Identification of candidate genes affected by *F. nucleatum* and EBV miRNA

Mjelle et al.[86] has detected increased expression of small RNAs from *F. nucleatum* and miR-BART10-3p in CRC tissue compared to normal tissue in three different datasets: Neerincx et al.[108], Sun et al.[109], and Mjelle et al.[86] (Figure 5).

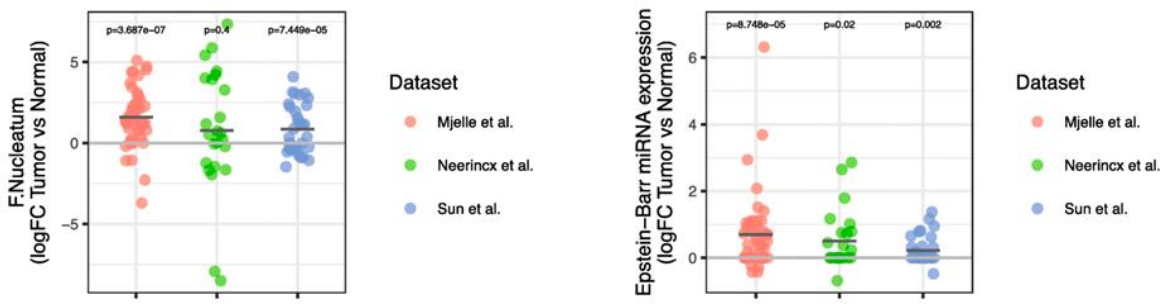


Figure 5. *Expression of RNAs in CRC tissue treated with F. nucleatum (left) and miR-BART10-3p (right). The figures show the fold-change values between paired tumour and normal tissue for F. nucleatum and miR-BART10-3p in the Neerincx, Sun and Mjelle datasets. Figure from Mjelle et al., 2019[86].*

To identify genes that respond to *F. nucleatum* in CRC, gene expression analysis of CRC cell lines co-cultured with *F. nucleatum* were performed. To identify human targets of miR-BART10-3p, transient transfection of miR-BART10-3p mimics in CRC cell lines were performed. Having detected major changes in gene expression, both upon *F. nucleatum* infection and miR-BART10-3p transfection, the present study looked further into four candidate genes, *CSF2* and *CCL20*, which were among the top upregulated genes in cells treated with *F. nucleatum*, and *MAT2B* and *ELL2*, which were downregulation upon miR-BART10-3p transfection and predicted targets of this miRNA using *in silico* prediction. The genes of interest were selected for further validation to investigate whether *F. nucleatum* and miR-BART10-3p are connected to the development of CRC through the regulation of these genes.

1.8.1. Chemokine (C-C motif) ligand 20 (*CCL20*)

Chemokine (C-C motif) ligand 20 (CCL20), also known as liver and activation-regulated chemokine (*LARC*), Exodus-1 or macrophage inflammatory protein-3a (*MIP-3α*), was discovered independently by three separate groups in 1997 using bioinformatic techniques[110-112]. CCL20 is a chemokine that plays an important role in the recruitment of dendritic cells (DCs), the proinflammatory IL17 producing helper T cells (Th17) and the regulatory T cells (Treg) to neoplastic lesions[113]. In normal colonic mucosa, *CCL20* mRNA is lowly expressed (Figure 6). However, in response to an inflammatory stimulus, the expression of *CCL20* is significantly increased[113]. Gene expression of *CCL20* can be stimulated by microbial factors such as LPS, and inflammatory cytokines such as tumour necrosis factor (TNF) and interferon-gamma (INF-γ)[114]. The CCL20 protein exerts all of its biological activity by binding and activating its sole high-affinity receptor chemokine (C-C motif) receptor type 6 (CCR6) and induces a strong chemotactic response by increasing the intracellular calcium ions[115]. Overexpression of CCR6/CCL20 in CRC cells has been demonstrated to increase their proliferation, migration and metastatic potential[115]. Another study has demonstrated that stimulation with CCL20 leads to activation of the ERK-1/2 kinase, MAP kinase and the AKT kinase which are major determinants in the control of diverse cellular processes such as proliferation, survival and differentiation[116].

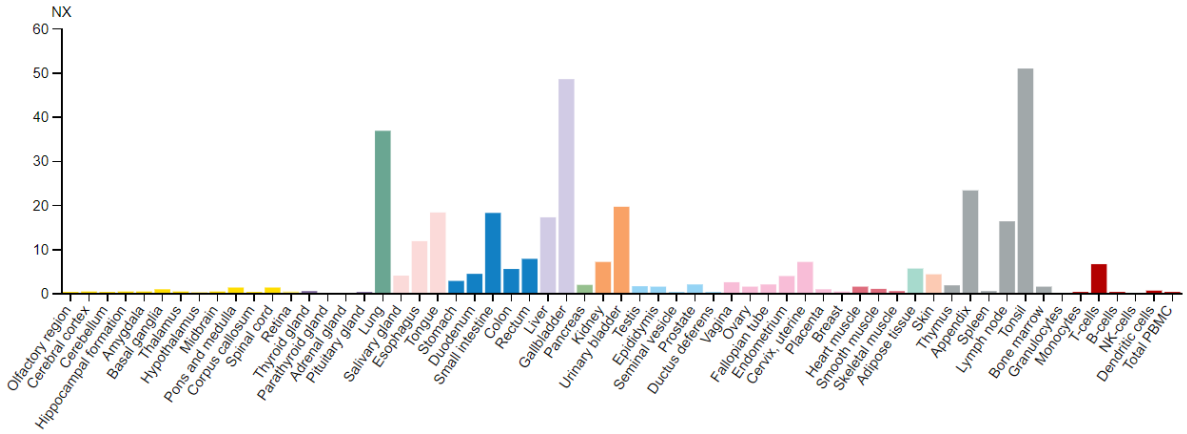


***Figure 6**. CCL20 RNA expression from consensus normalized expression (NX) data for 55 tissue types and 6 blood cell types by combining the data from three different datasets: RNA-seq data from the Human protein Atlas, RNA-seq data from the Genotype Tissue Expression project, and CAGE data from the FANTOM5 project. The colour-coding indicates the tissue groups with functional features in common. Figure from The Human Protein Atlas, n.d.[117].*

1.8.2. Colony stimulating factor 2 (*CSF2*)

The colony stimulating factor 2 (CSF2), also known as granulocyte macrophage-colony stimulating factor (GM-CSF), is a cytokine that stimulates hematopoietic stem cell differentiation into granulocytes and macrophages in the bone marrow[118]. In addition, CSF2 stimulates the recruitment, maturation, and functioning of DCs[118]. It is primarily expressed on natural killer cells, T cells, macrophages, endothelial cells, mast cells, and fibroblasts[119]. Normally, *CSF2* mRNA is lowly expressed in the colon (Figure 7). However, in response to inflammatory stimuli, such as bacterial endotoxins and local infections, the mRNA levels are significantly elevated[119]. The CSF2 protein signals through binding and activation of its cognate high-affinity receptor CD116, which is present on multiple cell types, including endothelial cells, granulocytes, lymphocytes and monocytes[119]. On the one hand, the cytokine has been considered to exert anti-tumour immune responses, mainly by the activation of DCs[120]. On the other hand, CSF2 upregulation has been shown to suppress the immune response and result in poor prognosis in multiple cancer types[118]. In has been reported that CSF2 promotes tumour progression and invasion by enhancing the expression of invasion associated MMPs, such as MMP-2, -9 and, -26,[121]. In CRC patients, demethylation and overexpression of *CSF2* mRNA is considered a potential diagnostic and prognostic marker indicating poor prognosis[118]. However, the exact role of *CSF2* mRNA in CRC remains unclear.

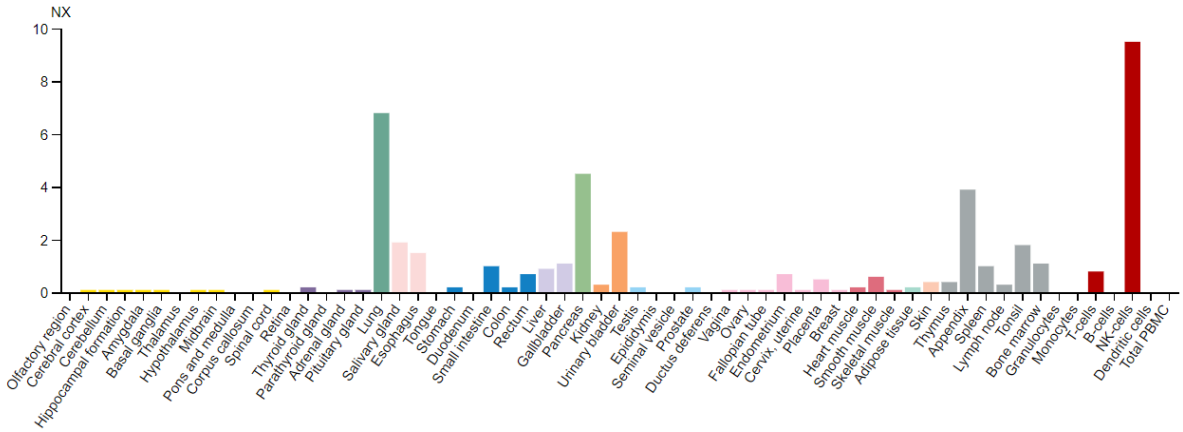


Figure 7. *CSF2 RNA expression from consensus normalized expression (NX) data for 55 tissue types and 6 blood cell types by combining the data from three different datasets: RNA-seq data from the Human protein Atlas, RNA-seq data from the Genotype Tissue Expression project, and CAGE data from the FANTOM5 project. The colour-coding indicates the tissue groups with functional features in common. Figure from The Human Protein Atlas, n.d.*[122].

1.8.3. Methionine adenosyltransferase 2 subunit beta (*MAT2B*)

Methionine adenosyltransferase 2 subunit beta (MAT2B) belongs to the methionine adenosyltransferase (MAT) family and is a critical enzyme that catalyses the biosynthesis of the methyl donor S-adenosylmethionine (SAM) from methionine and ATP[123]. *MAT2B* is synthesized in all mammalian cells (Figure 8) and encodes a regulatory subunit (β) that is physically associated with the MAT2A dimer, modulating the kinetic properties of MAT2A[124]. Interestingly, Figure 8 shows that *MAT2B* mRNA is expressed at high levels in immune cells, in particular B cells, which are the principal target cells for EBV[125]. In addition, *MAT2B* has a relatively high expression in epithelial cells of the gastrointestinal tract, which also has been found to be susceptible to EBV infection[125]. Previous studies have demonstrated that *MAT2B* may act as an oncogene in the carcinogenesis of several tumours, including CRC[126,127]. It has been reported that *MAT2B* encodes for variant proteins V1 and V2 that interacts with GIT1 (G Protein Coupled Receptor Kinase Interacting ArfGAP 1), and forms a scaffold that is essential to recruit and activate the ERK-1/2 pathway to promote cell growth and tumourigenesis[126]. In addition, downregulation of *MAT2B* has been shown to inhibit migration and activate apoptosis by inhibiting the AKT pathway[128]. These findings suggest that *MAT2B* knockdown could be efficient for halting cell proliferation through simultaneous suppression of AKT and ERK, supporting its potential as a therapeutic target.

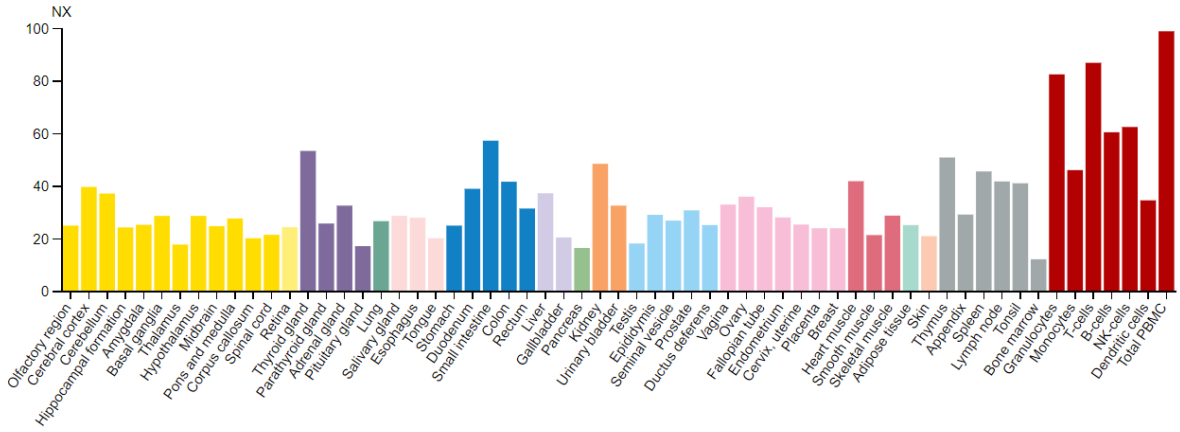


Figure 8. *MAT2B RNA expression from consensus normalized expression (NX) data for 55 tissue types and 6 blood cell types by combining the data from three different datasets: RNA-seq data from the Human protein Atlas, RNA-seq data from the Genotype Tissue Expression project, and CAGE data from the FANTOM5 project. The colour-coding indicates the tissue groups with functional features in common. Figure from The Human Protein Atlas, n.d.*[129].

1.8.4. Elongation Factor for RNA Polymerase II 2 (*ELL2*)

Elongation Factor for RNA Polymerase II 2 (ELL2) is an elongation factor component of the super elongation complex (SEC), which is required to increase the catalytic rate of RNA polymerase II (Pol II) transcription by releasing Pol II from some of the pausing sites along the DNA strand[130]. SEC has a functional role in the regulation of the transcriptional elongation checkpoint control (TECC), and misregulation of this stage is associated with carcinogenesis[131]. For instance, SEC is a frequent translocation partner of MLL (mixed-lineage leukemia), and relocalization to a MLL target such as the *HOX* gene results in the evasion of normal transcriptional controls and aberrant activation of MLL target genes involved in haematological malignancies[132]. The tissue specificity of the *ELL2* mRNA is relatively low[133] (Figure 9), but it has been reported that the RNA expression of *ELL2* varies in some human tissue[130]. For instance, *ELL2* is highly expressed in normal and Multiple myeloma (MM) plasma cells and the ELL2 protein drives secretory-specific Ig heavy chain mRNA production at a high rate via enhanced exon skipping and polyadenylation[134]. In addition, high expression of *ELL2* has been reported in the prostate, and knockdown of *ELL2* in prostate epithelial cancer cell lines has been shown to increase proliferation, migration, and invasion[135]. Figure 9 shows that *ELL2* is expressed in the tissue of the gastrointestinal tract as well, but whether *ELL2* exhibit tumour suppressive properties in colorectal tissue is yet unknown.

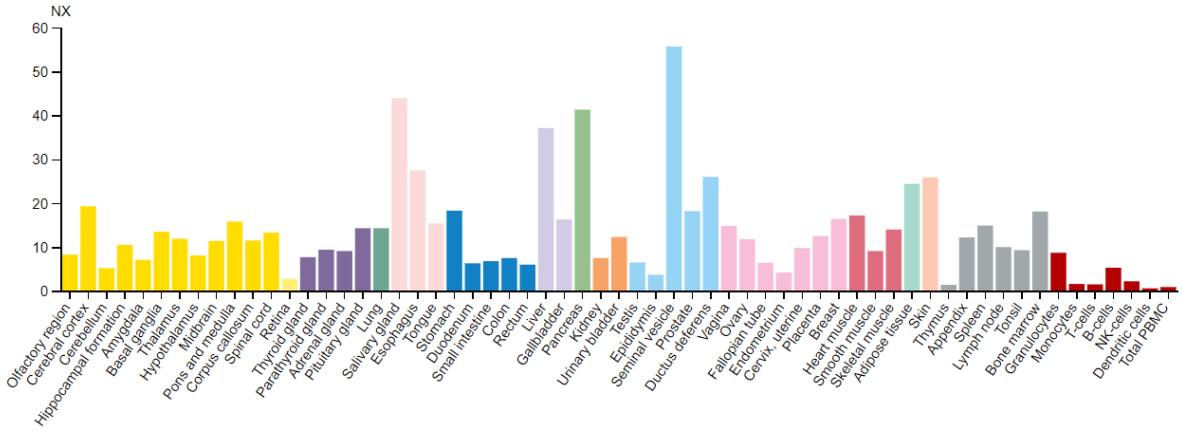


Figure 9. *ELL2 RNA expression from consensus normalized expression (NX) data for 55 tissue types and 6 blood cell types by combining the data from three different datasets: RNA-seq data from the Human protein Atlas, RNA-seq data from the Genotype Tissue Expression project, and CAGE data from the FANTOM5 project. The colour-coding indicates the tissue groups with functional features in common. Figure from The Human Protein Atlas, n.d.*[129].

2. Aim and objectives of study

The overall aim of this study was to investigate the role of *F. nucleatum* and the Epstein-Barr virus-encoded microRNA miR-BART10-3p in CRC and determine if and how they contribute to tumour development. The specific aims were:

1. Evaluate the direct effect of *F. nucleatum* in CRC by co-culturing the bacteria with CRC cell lines and look for changes in proliferation, migration and gene expression.
2. Examine the localization of *F. nucleatum* upon *F. nucleatum* co-culturing with CRC cell lines using confocal microscopy.
3. Evaluate the direct effect of miR-BART10-3p on gene expression in CRC cell lines and examine the regulatory effect of miR-BART10-3p on target genes by using 3' UTR target clones in expression vectors.
4. Establish an *in vitro* procedure for developing patient-derived CRC spheroids for future studies of *F. nucleatum* and miR-BART10-3p in a molecular and phenotypic landscape of an original tumour.

3. Methodology

In this project CRC cell lines are applied for different experiments, including co-culturing with *F. nucleatum* and transfection with miR-BART10-3p. This is followed by functional assays, including migration-, growth- and gene expression assays. The main methods used are confocal microscopy, ELISA, two-step RT-qPCR and Luciferase assays. In addition, a procedure for establishing patient-derived CRC spheroids is developed.

3.1. Applied reagents and kits

Reagents with supplier used in this assignment are listed in Table 1, except the reagents used in the spheroid optimization project which are listed in Supplementary section 4.1.

Table 1. *Reagents used in experiment listed with supplier*

Reagents	Manufacturer
LB-broth	In-house
Tryptic Soy Broth acc EP + USP	Sigma-Aldrich
RPMI 1640	Sigma-Aldrich
Dulbecco's Modified Eagle's Medium (DMEM)	Sigma-Aldrich
DMEM/Nutrient Mixture F-12 Ham	Gibco™
Dulbecco's phosphate buffered saline (PBS)	Sigma-Aldrich
Trypsin – EDTA Solution	Sigma-Aldrich
Penicillin-Streptomycin	Sigma-Aldrich
L-Glutamine solution	Sigma-Aldrich
Fetal bovine serum	Sigma-Aldrich
GelRed® Nucleic Acid Gel Stain	Biotium
GeneRuler 1 kb Plus DNA Ladder	Thermo Scientific™
GelPilot DNA Loading Dye, 5x	QIAGEN
Tris Acetate-EDTA (TAE) buffer	In-house
eBioscience™ CFSE	Invitrogen™
Rhodamine Phalloidin	Invitrogen™
DAPI	Sigma-Aldrich
Bovine serum albumin (BSA)	New England Biolabs
Lipofectamine RNAiMAX	Invitrogen™
TaqMan™ Universal Master Mix II, no UNG	Applied Biosystems™
XhoI	New England Biolabs
NotI	New England Biolabs
NEBuffer 3.1	New England Biolabs
SOC medium	In-house
Ampicillin	In-house
DharmaFECT Duo	Horizon Discovery

Kits with supplier used in this assignment are listed in Table 2.

Table 2. *Kits used in experiment listed with supplier*

Kit	Manufacturer
Q5® High-Fidelity PCR Kit	BioLabs
Total RNA Purification Kit	Norgen Biotek
High Capacity RNA to cDNA Kit	Applied Biosystems™
QIAquick Gel Extraction kit	QIAGEN
Wizard® Plus Minipreps DNA Purification System	Promega
LightSwitch Luciferase Assay System	Active Motif
Dual-Luciferase® Reporter Assay System	Promega
Human CCL20/MIP-3 alpha Quantikine ELISA Kit	R&D Systems

3.2. Bacterial strains and growth conditions

F. nucleatum (ATCC 25586) was grown at 37 °C in Tryptic Soy Broth acc EP + USP (TSB) under anaerobic conditions, and 1 mL was transferred to a new tube of TSB once a week to maintain proliferation. As presented in Figure 10, the anaerobic environment was created using an anaerobic jar with an anaerobic atmosphere generation bag (Thermo Scientific™). The anaerobic environment was controlled by an anaerobic indicator test (Sigma-Aldirch), that was placed inside of the anaerobic jar. The indicator turned pink in the presence of oxygen, and white in the absence of oxygen. Non-pathogenic *Escherichia coli (E. coli)* DH5α was used as a control. *E. coli* DH5α was thawed the day before usage and grown at 37 °C in LB broth under aerobic conditions. The multiplicity of infection (MOI) is the ratio between the number of bacteria in an infection and the number of host cells. The used MOI is indicated in each experiment.

Figure 10*. F. nucleatum in an anaerobic environment consisting of an anaerobic jar, an anaerobic atmosphere generation bag, and an anaerobic indicator test.*

3.3. Cell lines and culture conditions

Cell lines LS411N (ATCC® CRL-2159™) and DLD-1 (ATCC® CCL-221™) were grown in RPMI-1640 medium modified to contain L-glutamine and sodium bicarbonate. The medium was supplemented with 10% fetal bovine serum (FBS), 1% Penicillin-Streptomycin and L-glutamine. Cell lines SW620 (ATCC® CCL-227™) and SW420 (ATCC® CCL-228™) were grown in Dulbecco's Modified Eagle's Medium (DMEM) modified to contain 4500 mg/L glucose, L-glutamine, sodium pyruvate, and sodium bicarbonate. The medium was supplemented with 10% FBS, 1% Penicillin-Streptomycin and L-glutamine. Medium without Penicillin-Streptomycin was prepared for all cell lines during microbial assays.

The cells were maintained and split at approximately 80% confluency every 2-3 days in 75 cm^2 cell culture flasks with filter caps. All reagents were preheated in a 37 °C water bath. The cells were split by washing the cells twice with 10 mL of PBS, adding 1 mL trypsin-EDTA and incubating the cells for 3-4 minutes at 37 °C. To loosen the cells from the surface of the flask, fresh culture media (10 mL) was added to the detached cells, and the cells were split in a ratio of approximately 1:10. Fresh culture media (10 mL) was then added to the flask, and the cell culture was placed in a 5% CO_2 incubator at 37 °C.

3.4. Genes of interest

The genes of interest were selected based on initial gene expression analysis that were performed to assess *F. nucleatum* and miR-BART10-3p gene expression profiles during active invasion of CRC cell lines. We performed co-culture and transfection of *F. nucleatum* and miR-BART10-3p, respectively, and preparation of the sequencing libraries were performed by Mjelle in collaboration with the Genomics Core Facility (GCF) at NTNU. Cell lines DLD-1, SW620, LS411N and SW420 were infected with *F. nucleatum*, *E. coli* DH5α or TSB for 6 hours in antibiotic-free medium, followed by 24 hours of incubation with antibiotics to remove any extracellular bacteria. In addition, CRC cell lines SW620 and LS411N were transfected with miR-BART10-3p or a negative miRNA for 48 hours. The cells were harvested, and RNA was isolated using the Total RNA Purification Kit from Norgen Biotek. mRNA sequencing was done by using the SENSE mRNA-Seq Library Prep Kit V2 (Lexogen), and the sequencing library was sent to the GCF and sequenced in a 75 base pair single read run.

Having detected major changes in gene expression, both upon *F. nucleatum* infection and miR-BART10-3p transfection (Subheading 4.1), four candidate genes were further investigated;

CSF2 and *CCL20* in DLD-1, which were among the top upregulated genes when treated with *F. nucleatum*; and *MAT2B* and *ELL2*, which were downregulation upon miR-BART10-3p transfection and predicted targets of this miRNA using *in silico* prediction.

3.5. Proliferation and migration assay of *F. nucleatum*-treated CRC cells

An *in vitro* proliferation assay and migration assay can be used to determine whether cells are triggered to divide or migrate after exposure to specific stimulus, or to assess differences between cell populations in their ability to divide or migrate in response to the same stimulus[136,137]. As described in the introduction, *F. nucleatum* induces secretion of pro-inflammatory cytokines that drives CRC migration and proliferation *in vivo*. Here we want to test the ability of DLD-1 cells to proliferate and migrate *in vitro* with and without the stimuli of *F. nucleatum*.

3.5.1. Principle of µ-slide 8 well grid-500

Cell proliferation can be estimated by using the µ-slide 8 well grid-500 (ibidi). A certain volume of cells is placed in a well with an imprinted 500 µm cell location grid (Figure 11). When cells are attached to the surface, the cells are enumerated one by one in a certain area of known size at different time points with a phase contrast microscope. Cell proliferation can also be monitored by analysing the cell-free area of a certain area of known size at different time point with an automated imaging software[138].

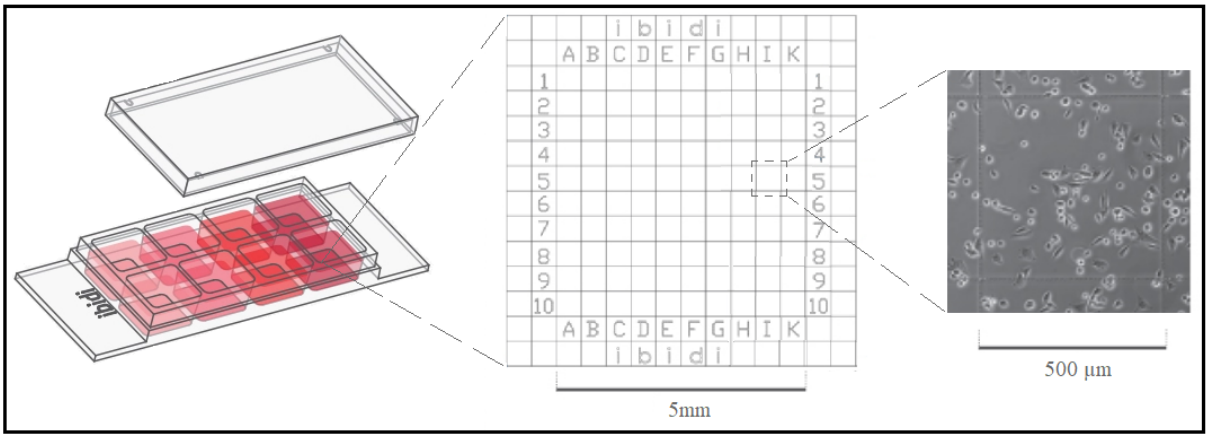


Figure 11. *Layout of the µ-Slide 8 Well Grid-500 (Ibidi). Modified figure from Ibidi, n.d.[138].*

3.5.2. Principle of Wound Healing and Migration Assay

Migration can be estimated by using the Wound Healing and Migration Assay (Ibidi). A cell-free area is created in a confluent monolayer of cells by using a Culture-Insert (Figure 12). A

Culture-Insert provides two culture reservoirs each separated by a 500 µm removable insert. Culturing cells in both reservoirs and then removing the Culture-Insert, result in two well-defined cell patches. The exposure to the cell-free area induces the cells to migrate into the gap[139]. The migration into the gap is monitored at different time points by using a phase contrast microscope, and the gap closure rate is analysed manually or by using an automated software[137].

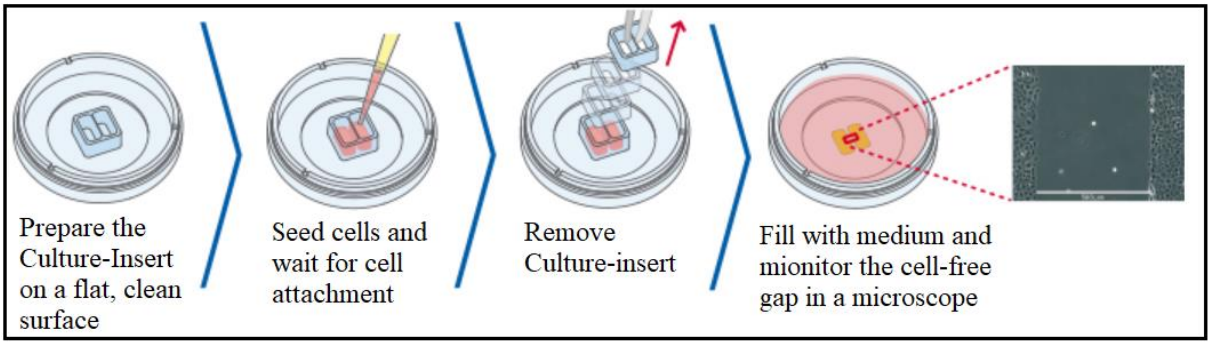


Figure 12. *Creating the Gap Using a Culture-Insert (Ibidi). The culture-Insert is prepared on a flat, clean surface. Cells are seeded in the reservoirs and cultured until the cells attach to the surface and form a monolayer. The Culture-Insert is then removed, and the dish is filled with medium. Cell migration into the gap is now monitored at different time points by using a phase contrast microscope. Modified figure from Ibidi, n.d.[140].*

3.5.3. Cell proliferation assay procedure

A cell suspension of 33 000 DLD-1 cells in 2 mL antibiotic-free medium was seeded in a 6-well plate. The cells were either untreated, incubated with *E. coli* DH5α or *F. nucleatum* at a MOI of 500. After 6 hours of incubation at 37 °C with 5% CO_2, cells were trypsinized and the medium was replaced with medium containing antibiotics to inhibit bacterial growth outside of the cells. Cells (300 µL) were seeded in 4 technical replicates in the µ–Slide 8 Well Grid–500 according to manufacturer's specifications, and each technical replicate was monitored after 0, 24 and 48 hours by using a phase contrast microscope. Cell proliferation was estimated by measuring the cell-free area in the image processing program ImageJ[141], and percent proliferation for each time point was calculated in Microsoft Excel.

3.5.4. Wound Healing and Migration Assay procedure

A cell suspension of 600 000 DLD-1 cells in 2 mL antibiotic-free medium was seeded in a 6-well plate. The cells were then untreated or incubated with *F. nucleatum* at a MOI of 300. After 6 hours of incubation at 37 °C with 5% CO_2, cells were trypsinized and medium was replaced with medium containing antibiotics. Culture-Inserts were prepared in a new 6-well plate, and the untreated and treated DLD-1 cell suspensions were transferred to the new wells. DLD-1 (70

µl) was seeded into each of the wells in the Culture-Insert, and 1.8 mL DLD-1 was seeded in the outer well. Cells were cultured until they formed an optically confluent monolayer, approximately for 24 hours. After 24 hours, the Culture-Insert was removed to create a gap. The cells were monitored by using a phase contrast microscope, and a picture of the gap was taken at 0, 24, 48 and 72 hours. The gap was estimated by measuring the gap size at the image processing program ImageJ[141], and the percent wound closure for each time point was calculated in Microsoft Excel.

3.6. Confocal microscopy of DLD-1 treated with *F. nucleatum*

To determine whether *F. nucleatum* is an extracellular or intracellular bacterium, the localization of the bacterium in co-culture with DLD-1 was visualized in a confocal microscope.

3.6.1. Principle of confocal microscopy

In contrast to wide-field microscopy that illuminates the whole sample at once, confocal microscopy enables high resolution images without any disturbing fluorescent light from the background of a diffraction limited spot at a specific depth within the sample. Furthermore, by stacking images from different optical sections, 3D structures can be analysed. It is also possible to analyse multicolour immunofluorescence staining that include several lasers and emission/excitation filters[142,143].

Figure 13 shows that a laser light is directed through a confocal pinhole to a dichroic mirror where it is reflected through the objective and focused to a diffraction limited spot in the sample. Emission light from the sample is directed back through the objective and the dichroic mirror to the light sensing detector. A pinhole inside the optical pathway cuts off signals that are out of focus, thus allowing only the fluorescence signals from the illuminated spot to enter the light detector. The proportionate voltage is produced, amplified and converted into digital levels for image display and storage[142,143].

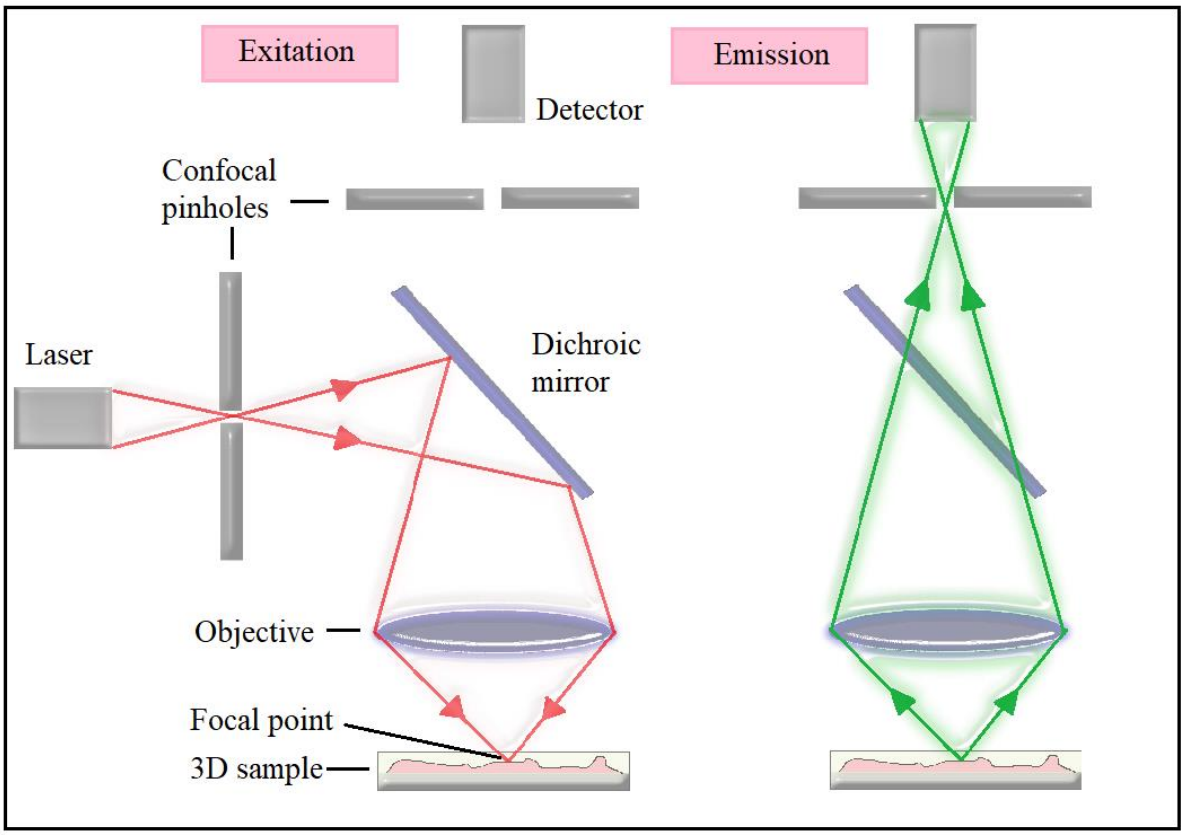


***Figure 13**. Excitation and emission light pathways in a basic confocal microscope. Laser light is sent through a confocal pinhole in order to eliminate out-of-focus light. The light is then reflected by a dichroic mirror and passes through an objective which focuses the light to a limited spot on the sample. Light is emitted at a longer wavelength which passes back through the objective and dichroic mirror, and then focuses through the upper confocal pinhole before the light hits the detector. Adapted from Ibidi, n.d.*[144].

3.6.2. Staining procedure

A cell suspension of 100 000 DLD-1 cells in 2 mL antibiotic-free medium was seeded into confocal dishes and incubated over night at 37 °C and 5% CO_2 until it reached a confluence of approximately 50%. *F. nucleatum* was labelled with eBioscience™ CFSE (Invitrogen™) by incubating at 37 °C and 300 rpm for 30 minutes followed by three washes with PBS. CFSE-labelled *F. nucleatum* was added to the cells at a MOI of 500 and incubated for 2 hours. The cells were fixed in 4% formaldehyde solution in PBS for 10 minutes at room temperature and washed in prewarmed PBS twice before and after fixation. Fixed cells were pre-incubated with PBS containing 1% bovine serum albumin (BSA) to reduce nonspecific background staining before labelling actin with the Rhodamine Phalloidin F-actin probe (Invitrogen™). The staining solution was made according to the manufacturer's protocol and subsequently added to the cells for 20 minutes at room temperature. Following two washes with PBS, the nucleus of the cells

was labelled with 0.1 µg/mL DAPI (Sigma-Aldrich) by incubating for 15 minutes in room temperature followed by two washes with PBS.

3.6.3. Visualisation using a confocal microscope

The cells were scanned under an 63X objective using a Zeiss LSM 510 META Confocal Microscope. To avoid cross talk, "Stack images" were taken at different cross sections of the sample with different lasers and detecting channels for each probe. The cells were imaged by using a laser for CFSE at 488 nm, Rhodamine Phalloidin at 561 nm, and DAPI at 405 nm.

3.7. Time course of *F. nucleatum*-treated cells with two-step RT-qPCR

To determine how fast *F. nucleatum* alters host gene expression of *CCL20* and *CSF2*, CRC cell line DLD-1 was infected with *F. nucleatum* and harvested at various time points. Gene expression was measured using two-step RT-qPCR.

3.7.1. Principle of RT-qPCR with relative quantification

In reverse transcription polymerase chain reaction (RT-qPCR), RNA is used as the starting material. RNA is first transcribed into complementary DNA (cDNA) by reverse transcriptase. The cDNA is then used as the template for the qPCR reaction. Reverse transcription can be performed separately from qPCR (two-step RT-qPCR) or directly in the qPCR mix (one-step RT-qPCR). The former is preferred when multiple runs will be made of the same starting material or when storage of cDNA is necessary.

In qPCR, a fluorescent reporter gene is used to measure the quantity of cDNA present at each cycle. At a point where the qPCR fluorescence signal is detectable above the background, a threshold cycle (Ct) value can be determined. If the Ct value is low, it means the fluorescence crosses the threshold early, meaning that the gene expression in the sample is high[145,146]. Relative quantification uses the $2^{-\Delta\Delta CT}$ (Livak) method[147] to determine changes in gene expression relative to a housekeeping gene (HG):

ΔCT (test) = CT (target, test) – CT (Reference, test)

ΔCT (HG) = CT (target, HG) – CT (Reference, HG)

$\Delta\Delta$CT = ΔCT (test) - ΔCT (HG)

Fold change = $2^{-\Delta\Delta CT}$

If the fold change is larger than 1 it means that the gene is upregulated; if the fold change is less than 1 it means that the gene is downregulated.

3.7.2. Procedure for Time course assay

A cell suspension of 250 000 DLD-1 cells in 2 mL antibiotic-free medium was seeded in three 6-well plates and cultured until they formed an optically confluent monolayer, for approximately 24 hours. The cells were then untreated, incubated with *E. coli* DH5α or *F. nucleatum* at a MOI of 500. After 0, 3, 6, 12 and 24 hours, cell culture medium was removed from the wells, centrifuged at 500 x g for 5 minutes, and 600 µL of the supernatants were stored in Eppendorf tubes at -20 °C until use for Quantikine ELISA (Subheading 3.8). For two-step RT-qPCR, cells were washed with PBS and gently scraped off the bottom of the wells into PBS, starting with the first well in the plate. Cell suspensions were centrifuged at 3000 x g for 5 minutes, and cell pellets were frozen and kept at -20 °C until use for RNA isolation.

3.7.3. Procedure for RNA isolation and two-step RT-qPCR

RNA isolation was done according to the "Total RNA Purification Kit" from Norgen Biotek Corp, except that RNA elution was done with 25 µl H_2O instead of 50 µl Elution Solution A. Preparation of lysates was done according to Step 1A(ii) "Cell Lysate Preparation from Cells Growing in Suspension and Lifted Cells" before proceeding to Step 2 "Total RNA Purification". RNA was quantified using a Nanodrop™ 1000 spectrophotometer (Thermo Fisher Scientific), and equal amounts of RNA were prepared by diluting samples in RNAse free water. Reverse transcription was done according to the "High Capacity RNA to cDNA Kit" from Thermo Fisher Scientific. The 20 µL reaction mix was prepared and loaded according to the "TaqMan® Gene Expression Assays Protocol" from Applied Biosystems using the TaqMan® Universal Master Mix, no UNG. Commercially available fluorescently labelled Taqman probes were used for *CCL20* (Hs00355476_m1), *CSF2* (Hs00929873_m1) and the housekeeping gene *ACTB* (Hs99999903_m1). All samples were loaded in triplicate. qPCR analysis was performed on the StepOne™ Real-Time PCR System (Applied Biosystems) under the following reaction conditions: 95 °C for 10 minutes for polymerase activation, followed by 40 cycles of 95 °C for 15 seconds and 60 °C for 1 minute.

3.7.4. Calculation of results from qPCR Time course assay

Results were calculated using Microsoft Excel. Triplicates were averaged for each sample. The fold change between treated and untreated DLD-1 was determined using the $2^{-\Delta\Delta Ct}$ (Livak) method[147]. For each time point, ΔCt (the expression of *CCL20* and *CSF2* in untreated or treated cells normalized to *ACTB* housekeeping gene) and the ΔΔCt (ΔCt of treated cells normalized to the ΔCt of untreated cells) were calculated, and the fold changes of the gene expression at each time point in the differentially treated cells were determined.

3.8. Quantikine ELISA of Human CCL20

To determine whether *F. nucleatum* alters host cell release of CCL20, the analyte was measured in DLD-1 cell culture supernatants with and without treatment of *F. nucleatum* using Quantikine ELISA. Cell culture supernatants collected at 0, 6, 12 and 24 hours after infection (Subheading 3.7.2.) were used to assess whether CCL20 release is upregulated in a time-dependent manner. In addition, cell culture supernatants from DLD-1 infected with increasing concentrations of *F. nucleatum* (0 MOI, 0.1 MOI, 1.0 MOI, 10 MOI and 100 MOI) were used to assess whether CCL20 release is upregulated in a dose-dependent manner.

3.8.1. Principle of Quantikine ELISA Human CCL20 Immunoassay

A capture antibody specific for human CCL20 has been immobilized onto the surface of a multiwell plate (Figure 14). Standards and samples are added to the wells and any CCL20 present is bound by the antibody. The wells are washed to remove any unbound antigen. Following the washing step, a CCL20-specific detection antibody conjugated to an enzyme is added to the wells. Following a washing step to remove any excess antibody, a substrate is added to the wells, and a colorimetric change occurs in proportion to the amount of CCL20 bound in the sample. Finally, stop solution is added, and the absorbance of the colour is measured on an ELISA plate reader at 450 nm[148].

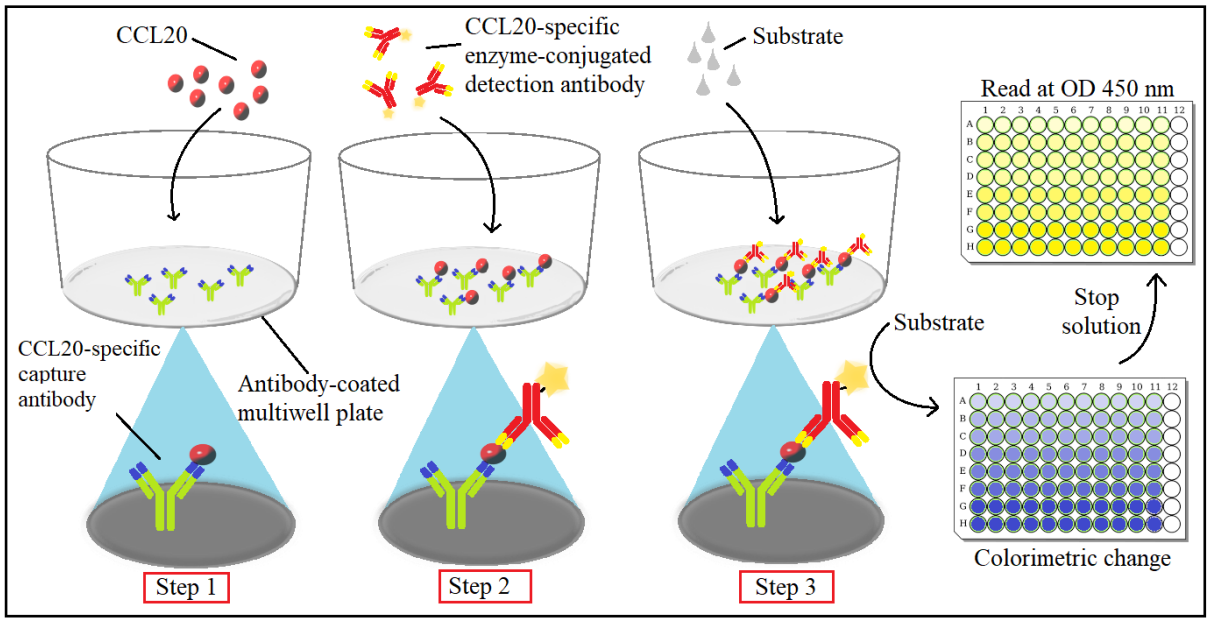


Figure 14. *Principle of Quantikine ELISA Human CCL20 Immunoassay. CCL20-specific capture antibodies are coated in a multiwell plate. Samples with CCL20 are added, and CCL20 is bound by the immobilized antibody (Step1). Enzyme-conjugated detection antibodies are then added and binds to the captured CCL20 (Step 2). Unbound materials are washed away between each step. A substrate solution is added to the wells and a blue colour develops in proportion to the amount CCL20 present in the sample (Step 3). The colour development is stopped by a stop solution, turning the colour in the wells yellow. The absorbance of the colour is measured at 450 nm. Results are expressed as Optical Density (OD). Adapted from R&D Systems, n.d.*[149].

3.8.2. Procedure for dose-dependent assay

A cell suspension of 250 000 DLD-1 cells in 2 mL antibiotic-free medium was seeded in a 6-well plate and cultured until it formed an optically confluent monolayer, approximately for 24 hours. The cells were co-cultured with *E. coli* DH5α or *F. nucleatum* at 0 MOI, 0.1 MOI, 1.0 MOI, 10 MOI and 100 MOI. After 6 hours of incubation, cell culture medium was removed from all wells, centrifuged at 500 x g for 5 minutes, and 600 µL of the supernatants were stored in Eppendorf tubes at -20 °C until use for Quantikine ELISA

3.8.3. Procedure for Quantikine ELISA

DLD-1 cell culture supernatants collected 0, 6, 12 and 24 hours after the addition of 500 MOI *F. nucleatum* or *E. coli* DH5α (Subheading 3.7.2.), as well as DLD-1 cell culture supernatants collected 6 hours after the addition of 0 MOI, 0.1 MOI, 1.0 MOI, 10 MOI and 100 MOI *F. nucleatum* or *E. coli* DH5α (Subheading 3.8.2.), were collected and thawed on ice. The ELISA assay was done according to the "Quantikine® ELISA Human CCL20/MIP-3α Immunoassay procedure" from R&D Systems. All standards and samples were assayed in duplicate. The

optical density was determined using a microplate reader set to 450 nm with correction set to 540 nm.

3.8.4. Calculation of results from Quantikine ELISA

Results were calculated using Microsoft Excel. Readings at 540 nm were subtracted from the readings at 450 nm. The mean of each duplicate was calculated, and the mean of the zero standard optical density was subtracted from each duplicate. A standard curve was created by plotting the mean optical density for each standard against the concentration, and a best fit curve was made through the points on the graph by checking the linear trendline. The equation of the trendline was further used to calculate the concentration of each duplicate.

3.9. Two-step RT-qPCR of SW620 transfected with miR-BART10-3p

Transfection is defined as the introduction of nucleic acids into the cytoplasm of eukaryotic cells. Such introductions of foreign nucleic acid can result in altered properties of the cell, allowing the study of host gene function and protein expression[150]. To determine whether miR-BART10-3p modulates gene expression of its predicted mRNA targets, SW620 cells were transfected with miR-BART10-3p or a negative miRNA control for 24 hours, following gene expression measurements using two-step RT-qPCR.

3.9.1. Transfection assay of SW620 transfected with miR-BART10-3p

A cell suspension of 500 000 SW620 cells in 2 mL culture medium was seeded in a 6-well plate. The cells were then transfected with miR-BART10-3p or a negative miRNA control by using Lipofectamine RNAiMAX (Invitrogen™) according to the manufacturer's instructions. After 24 hours of incubation, the cell suspensions were trypsinized, resuspended in a 1.5 mL tube with PBS, and then centrifuged at 4000 x g for 5 minutes. Cell pellets were frozen and kept at -20 °C until use for RNA isolation. RNA isolation and two-step RT-qPCR were performed as described in Subheading 3.7.3., except that cDNA was seeded in 6 technical replicates. Commercially available fluorescently labelled Taqman probes were used for the target genes *ELL2* (Hs00603761_g1), *MAT2B* (Hs00203231_m1) and the housekeeping gene *ACTB* (Hs99999903_m1).

3.9.2. Calculation of results from qPCR Transfection assay

Results were calculated using Microsoft Excel. The average of the technical replicates was calculated for each sample. The fold change of SW620 treated with miR-BART10-3p versus negative miRNA was determined using the $2^{-\Delta\Delta Ct}$ (Livak) method[147]. For each sample, ΔCt (gene expression of *ELL2* and *MAT2B* in cells treated with miR-BART10-3p or negative miRNA normalized to the housekeeping gene *ACTB*) and the ΔΔCt (ΔCt of miR-BART10-3p normalized to the ΔCt of negative miRNA) were calculated, and the fold change of gene expression in SW620 treated with miR-BART10-3p versus negative miRNA was determined.

3.10. Cloning vector and luciferase Assay

The regulatory effect of miR-BART10-3p on the target genes was determined by using a 3'UTR target clone in expression vector. Cells were co-transfected with a 3'UTR reporter vector and a miR-BART10-3p or a negative miRNA, followed by measurements of the luciferase activity. Two types of luciferase assay systems were used in this assay; Dual-Luciferase® Reporter Assay System (Promega) and LightSwitch™ Luciferase Assay System (Active Motif). The former assay was used for vector construction of the miR-BART10-3p 8mer target site of *MAT2B* in a psiCHECK™-2 vector, whereas the latter assay was used for transfection-ready 3´UTR Reporter GoClones for both *MAT2B* and *ELL2*.

3.10.1. Principle of DNA cloning

DNA cloning is a process of making large numbers of identical copies of a specific DNA sequence, such as a gene. A gene of interest is linked through standard 3′ → 5′ phosphodiester bonds to a vector DNA molecule, which can replicate when introduced into a host cell. The insertion is done using restriction enzymes that cut open the vector DNA molecule and the gene of interest at specific sites, using the same restriction enzyme for both the vector and the insert, and a DNA ligase which joins the vector and the insert at the specific cut sites, producing recombinant DNA. When the vector with the insert is introduced into a host cell, the inserted DNA is transcribed along with the vector using the vector's promoter, producing large numbers of gene products. The vectors most commonly used in DNA cloning naturally occur and replicate in *E. coli* DH5α. To simplify working with vectors, their length is reduced, and they contain little more than the essential nucleotide sequences required for their use in DNA cloning: a replication origin, a drug-resistance gene, a reporter gene, a promoter, and a region in which the gene of interest can be inserted. Binding of miRNAs to the target gene that is

transcribed by the vector results in translational inhibition or degradation of the targeted transcript[151].

3.10.1.1. *LightSwitch™ 3'UTR Reporter GoClone®*

The LightSwitch 3'UTR clones are pre-made 3'UTR reporter constructs that can be used to validate miRNA target sites. The LightSwitch clone contains a luciferase gene and a downstream 3'UTR for the gene of interest (Figure 15). Luciferase is an oxidative enzyme that produces bioluminescence in the presence of a substrate. The luciferase reporter gene is expresser under control of a promoter. The chimeric vector is co-transfected into a desired cell

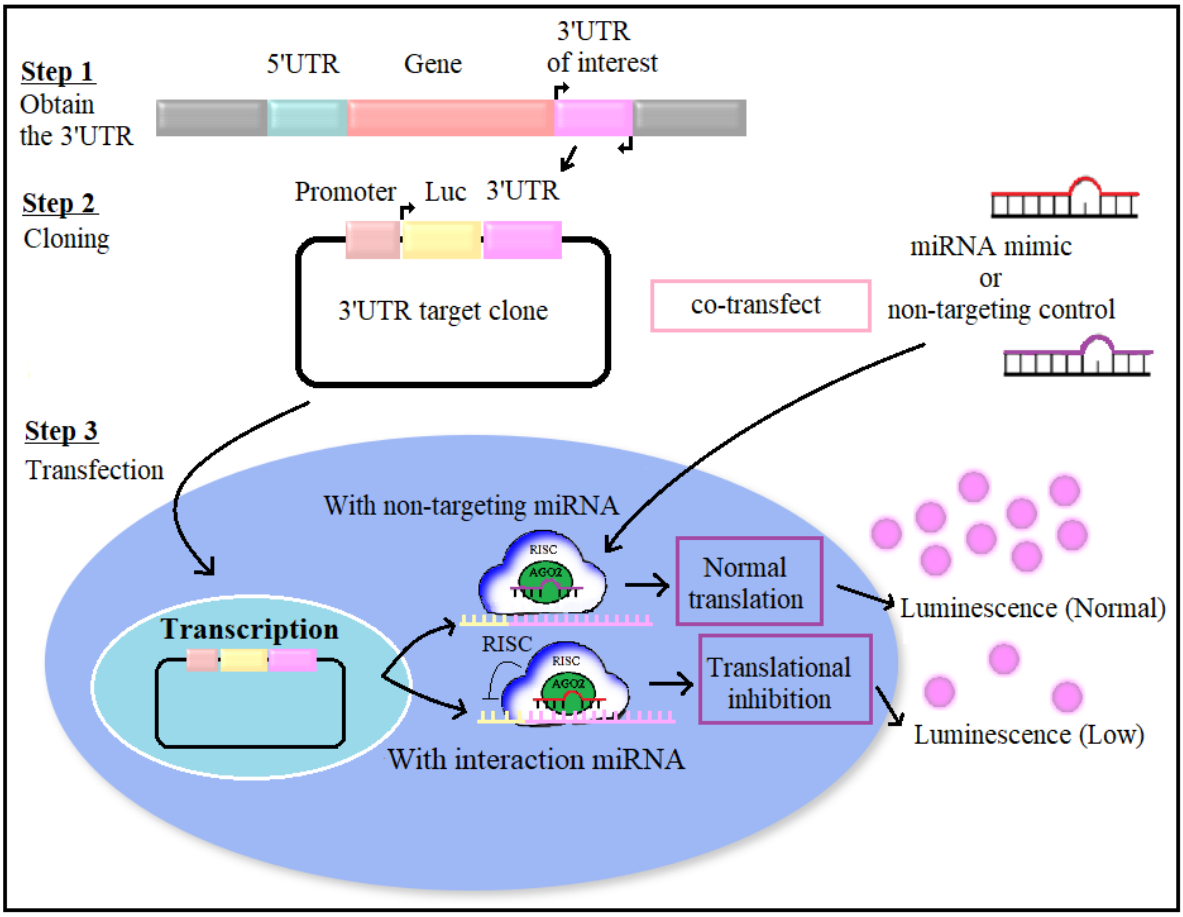


Figure 15. *How 3'UTR miRNA target clones work. Obtained 3'UTR of interest is inserted into a vector downstream of a luciferase reporter gene, which is expressed under control of a promoter. The chimeric vector is co-transfected into a desired cell line with a miRNA mimic or a negative control. The vector is transcribed, and a mRNA consisting of a luciferase reporter and a 3' UTR target sequence is exported to the cytoplasm. The RISC complex containing a negative miRNA does not affect translation of the mRNA. The RISC complex containing a miRNA mimic is complementary to the mRNA strand, resulting in translational inhibition and low luminescence. Adapted from Biocat, n.d.[152] and Switchgear Genomics, n.d.[153].*

line with a miRNA mimic or a negative control. After transcription, the mRNA, consisting of a luciferase reporter and a 3' UTR target sequence is exported to the cytoplasm. Both negative miRNAs and miRNA mimics are incorporated into the RISC complex. The negative miRNA does not contain any binding site to the target mRNA and will therefore not bind. Therefore, the RISC complex containing a negative miRNA does not affect translation or the stability of the mRNA. In contrast to the negative miRNA, the miRNA mimic has a complementary target site within the target mRNA, resulting in degradation or translational inhibition of the target mRNA. This will further lead to decreased luminescence from the reporter gene since less protein is being produced[152,153].

3.10.2. psiCHECK™ -2 vector and Dual-Luciferase® Assay Principle

The psiCHECK™ -2 vector is 6,273 bp long and uses *Renilla* luciferase as the primary reporter gene (Figure 16A). The 3'UTR region of the target gene is cloned into a multiple cloning region located downstream of the *Renilla* translational stop codon. The vector uses a second reporter gene, firefly luciferase, which allows for internal normalization of *Renilla* luciferase expression[154]. The psiCHECK™ -2 vector should be used with the Dual-Luciferase® Reporter Assay System for maximum sensitivity.

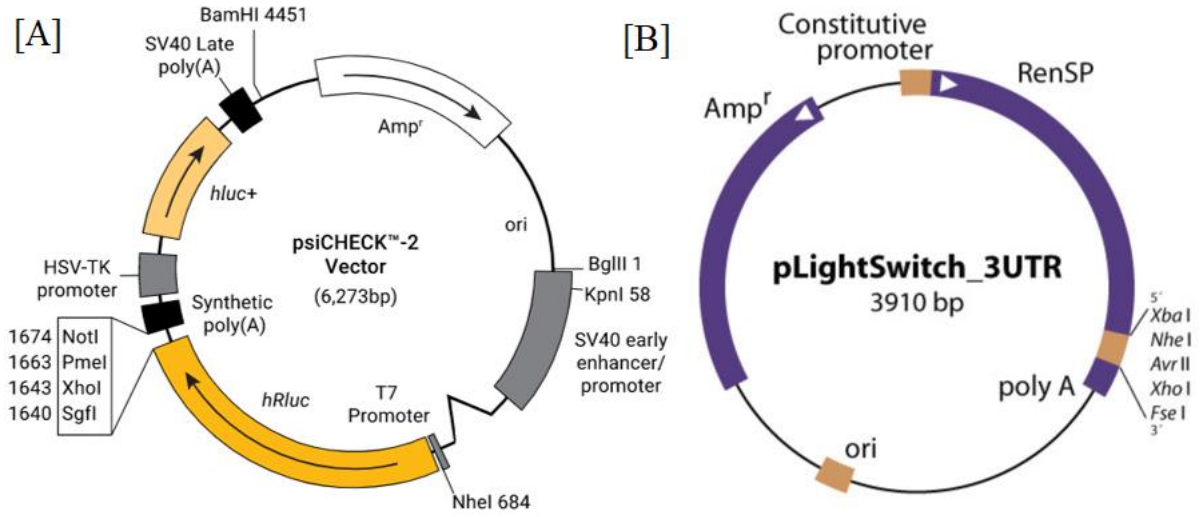


***Figure 16**. The psiCHECK™-2 vector **[A]** with Renilla luciferase (hRluc) as the primary reporter gene and firefly luciferase (hluc+) as the second reporter gene. The miRNA target site is between the restriction enzymes XhoI and NotI, where the target site of the gene of interest is inserted. The LightSwitch™ 3'UTR Reporter Vector **[B]** with RenSP, an optimized luciferase gene, as the reporter gene. This vector is purchased with transfection ready 3'UTR target sequence, which is already cloned into the vector. Figures from Promega, n.d.[154] and Active Motif, n.d.[155].*

First, cells are lysed by adding the Firefly Luciferase Reagent. The reagent contains a substrate for firefly luciferase to produce firefly luciferase luminescence. The firefly enzyme catalyses the ATP-, Mg^{2+}- and O_2-dependent oxidation of luciferin to oxyluciferin (Figure 17). After quantifying the firefly luminescence, the *Renilla* Luciferase Reagent is added. The reagent quenches the firefly luciferase luminescence and provides the substrate for *Renilla* luciferase to produce *Renilla* luciferase luminescence. *Renilla* luciferase catalyses the O_2-dependent oxidation of the substrate Coelenterazine to Coelenteramide. The light production of both reactions is measured on a luminometer. The results are then expressed as the ratio of *Renilla* to firefly luciferase activity[156].

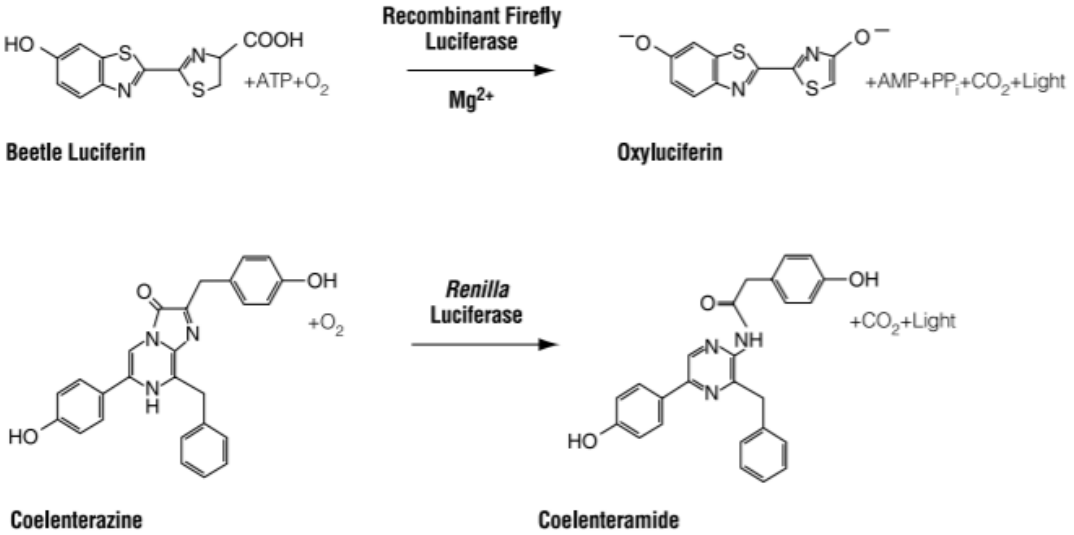


***Figure 17**. Luminescent reactions catalysed by Firefly and Renilla luciferase. Figure from Promega, n.d.[154].*

3.10.3. LightSwitch™ 3'UTR Reporter Vector and LightSwitch™ Luciferase Assay System Principle

The LightSwitch™ 3'UTR Reporter Vector should be used with LightSwitch™ Luciferase Assay Reagents for maximum sensitivity. The 3'UTR gene of interest is cloned into the vector downstream of the luciferase reporter gene. LightSwitch™ reporter vectors are transfection-ready vectors, so no cloning or DNA preparation is required. The vector is 3,910 bp long, and contains an optimized *Renilla* luciferase reporter gene, called RenSP (Figure 16B). Its overall enzymatic activity is increased, and a protein destabilization domain is added to decrease its half-life. The RenSP enzyme catalyses oxidation of coelenterazine to coelenteramide to produce light at 480 nm (Figure 18). The light production of the reaction is measured on a luminometer[157].

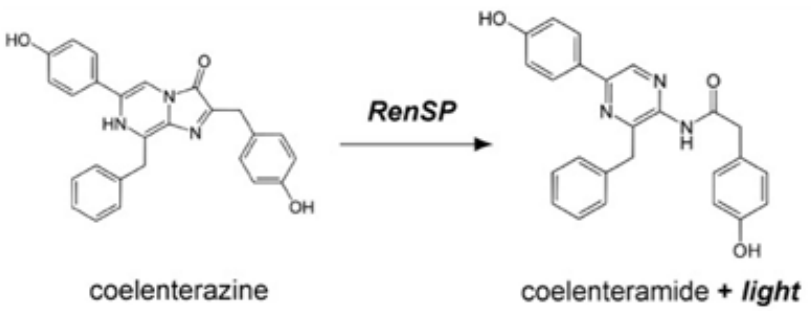


Figure 18. *Luminescent reaction catalysed by RenSP luciferase. Figure from © SwitchGear Genomics, n.d.*[157].

3.10.4. Vector construction procedure for the the psiCHECK™-2 vector

The psiCHECK™-2 vector was used to investigate the effect of specific target sites for miR-BART10-3p without the entire 3'UTR sequence context. Using only the target site and a few nucleotides upstream and downstream of the target site allows us to investigate the effect of a specific target site itself, thereby eliminating potential effects from other parts of the 3' UTR. To construct a psiCHECK™-2 vector with the *MAT2B* 8mer target site, the psiCHECK™-2 vector (Promega) was digested with restriction enzymes XhoI and NotI (New England Biolabs) for about 24 hours, as described in Supplementary section 2.1. After restriction enzyme cutting, the expected length of the vector was 6,242 bp. Gel electrophoresis was done as described in Supplementary section 2.2., using the GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific™). The gel was purified using the QIAquick Gel Extraction kit (QIAGEN) according to manufacturer's specifications. Pre-extracted and extracted gels are shown in Figure 19A and B. The *MAT2B* target-site sequence, containing the miR-BART10-3p binding site, was purchased from Integrated DNA Technologies as a pair of oligos with a premade overhang for XhoI and NotI (Table 3). The *MAT2B* oligo pair was phosphorylated, annealed and ligated into the digested psiCHECK™-2 vector as described in the Target Sequence Cloning Protocol (Addgene). The reporter vector was then used to transform *E. coli* DH5α-competent cells by heat shock transformation. For complete protocol of heat shock transformation, see Supplementary section 2.3. Following heat shock transformation, the bacteria were seeded on a LB-plate with antibiotics, cultured for 24 hours and finally purified using the Wizard® *Plus* SV Minipreps DNA Purification System (Promega) according to the manufacturer's specifications, to release the vector DNA. Vector DNA was then quantified using a Nanodrop™ 1000 spectrophotometer (Thermo Fisher Scientific).

***Table 3**. Sequences of the MAT2B insert oligo pair. The miR-BART10-3p 8mer target site is indicated in red.*

Name	Sequence
MAT2B.1	5'-TCG AGG CTT GAG ATA TTT CAA CAT **GTT ATG TA**T ATT GGA ACG C-3'
MAT2B.2	5'-GGC CGC GTT CCA ATA TAC ATA ACA TGT TGA AAT ATC TCA AGC C-3'

The reporter vectors were checked by PCR and gel electrophoresis. Primers were designed to amplify a region of the vector containing the insert. The *MAT2B* Forward primer was designed to overlap the *MAT2B* insert, meaning that the primer would only bind if the insert was present in the vector, thereby indicating a successful cloning procedure. Using this technique, we could easily screen multiple colonies from the cloning by running the PCR products on an agarose gel and identify the expected PCR product. The specific primers used to see if the *MAT2B* target sequence was successfully ligated into the vectors were as follows: 5'-ACC CTA ACC ACC GCT TAA GC -3' (psiCHECK™-2 Reverse primer) and 5'-GAT ATT TCA ACA TGT TAT GTA TAT-3' (*MAT2B* Forward primer). The predicted length of PCR product is about 230 bp. PCR was done according to the PCR Protocol for Q5® High-Fidelity DNA Polymerase (BioLabs). Annealing temperature was set to approximately 5 °C lower than the melting temperature of the primers. Gel electrophoresis was done as described in Supplementary section 2.2., and image of the gel is shown in Figure 19C, using the GeneRuler 1 kb Plus DNA Ladder from Thermo Scientific™.

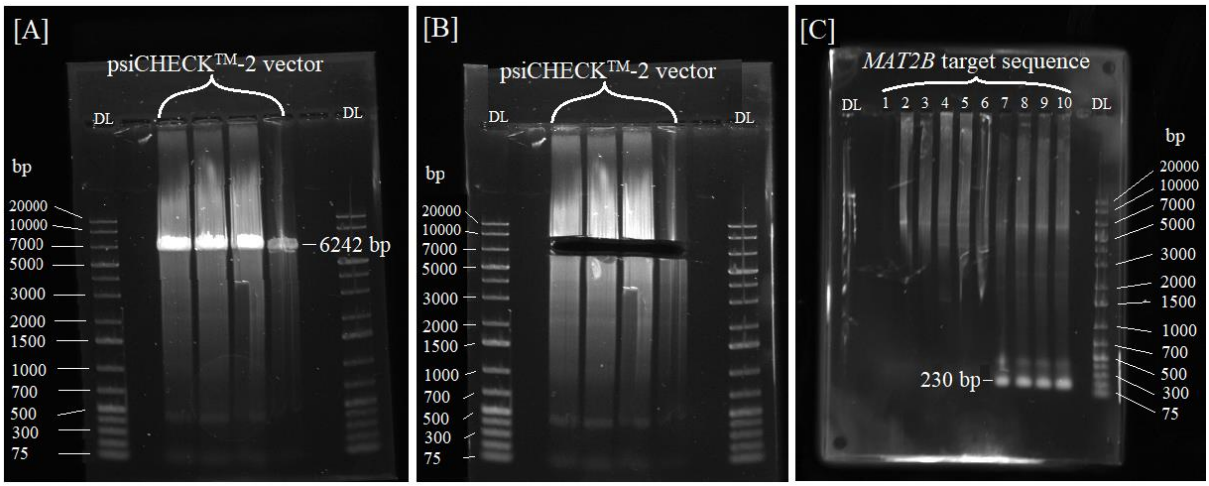


***Figure 19**. Pre-cut **[A]** and cut **[B]** gels showing a band of approximately 7,000 bp, indicating that it is the psiCHECK™-2 vector of 6,242 bp. **[C]** Gel showing ten 3'UTR target sequences of MAT2B which has been ligated into a psiCHECK™-2 vector, heat shocked and amplified using a MAT2B forward primer and psiCHECK™-2 reverse primer. Four PCR products (7-10) shows visible bands at 230 bp, which indicates that the target sequence was successfully ligated into the vector. GeneRuler 1 kb Plus DNA Ladder (DL) from Thermo Scientific was loaded into the first and last well of each gel.*

PCR products in well 7-10 seemed to have the right size of approximately 230 bp. Accurate cloning of target sequence 8 was confirmed by RNA sequencing. For RNA sequencing, 500 ng ligation product, 2.5 µL psiCHECK™-2 Reverse primer and dH_2O to a total volume of 10 µL was sent to GATC Biotech (Konstanz, Germany). Sequencing results, including FASTA sequence and chromatogram, are presented in Supplementary section 3, and accurate cloning of the *MAT2B* insert into the psiCHECK™-2 vector was confirmed by detecting the reverse complement of MAT2B.1 (Table 3) within the psiCHECK™-2 sequence. This was done by using the reverse complement tool[158].

3.10.5. Co-transfection of SW620 with *MAT2B* reporter vector and miR-BART10-3p

A cell suspension of 25 000 SW620 cells in 100 µL culture medium was seeded in a 96-well plate for 24 hours. At a confluence of approximately 80%, the 3'UTR reporter vector or an empty psi-CHECK™-2 vector (control) was co-transfected with miR-BART10-3p or a negative miRNA to the cells in 4 technical replicates. Transfection was done using DharmaFECT Duo (Horizon Discovery) according to the manufacturer's instructions. Co-transfected cells were incubated for 48 hours in a 5% CO_2 incubator at 37 °C.

3.10.6. Dual-Luciferase® Reporter Assay and calculation of results

Firefly and *Renilla* luciferase activities were measured 48 hours after transfection using the Dual-Luciferase® Reporter Assay System (Promega) according to the manufacturer's instructions, using a 2 second read in a plate luminometer. The results were calculated by first removing background noise by subtracting the mean of the untreated samples (blank) from each sample and control. Results were then expressed as the ratio of *Renilla* to firefly luciferase activity. The degree of knockdown was further measured by dividing the signal from miR-BART10-3p by negative miRNA, and then normalizing each sample with the control.

3.10.7. LightSwitch™ Luciferase Assay and calculation of results

Transfection-ready 3´UTR Reporter GoClone® constructs from Active Motif were used for *MAT2B* and *ELL2*. Cell seeding and transfection was done as described in Subheading 3.10.5., except that a Random 3'UTR control vectors (R04_3UTR) from SwitchGear Genomics was used as a negative control. Random control constructs contain non-conserved, non-genic, and non-repetitive human genomic fragments[159], and are therefore expected not to be a target for

miR-BART10-3p. A total of 48 hours after transfection, luciferase activity was measured using the LightSwitch Luciferase Assay System (Active Motif) according to the manufacturer's instructions. Each well was read for 2 seconds in a plate luminometer. The results were calculated by first removing the background noise by subtracting the mean signal from the untreated samples (blank) from each sample and control. The degree of knockdown was measured as the ratio of miR-BART10-3p divided by negative miRNA, and then normalized with the control.

3.11. Patient-derived colorectal cancer spheroids

In this experiment, we wanted to establish and introduce a new method in the laboratory with patient-derived CRC spheroids. The aim was to provide a method for future studies of *F. nucleatum* and miR-BART10-3p in a molecular and phenotypic landscape of an original tumour, and to study the native microbiome of the spheroids. Reagents used in this experiment are listed in Supplementary section 4.1.

3.11.1. Human specimens

This project was performed in cooperation with the Systems Biology for Oncology at NTNU. Primary CRC tissue from 17 donors with CRC stage I-IV were obtained during tumour resection at the Gastro Centre at St. Olavs hospital in Trondheim. Written informed consent was given by all donors. The donors have not been exposed to radiation or chemotherapy prior to surgery. Ten donor samples (CRC-01 to CRC-10) were retrieved in the first phase, and seven donor samples (CRC-11 to CRC-17) were retrieved in the second phase. Samples were confirmed as tumour or normal tissue by the gastrointestinal surgeon. The study protocol was approved by the Regional Committees for Medical and Health Research Ethics (REK 2019/246).

3.11.2. Phase 1 – Initial work

In the initial phase, cancer-tissue originated spheroids (CTOSs) were prepared based on a method published by Kondo et al.[92] to evaluate the reproducibility of the method. The CTOS preparation method is described in Subheading 3.11.2.1, and the CTOS culture and expansion method is described in Subheading 3.11.2.2. To successfully establish CTOSs, several modifications of the methods were tested during the initial phase (Subheading 3.11.2.3).

3.11.2.1. Cancer-tissue originated spheroids - preparation

Tissue samples were placed in supplemented DMEM on ice after tumour resection and transported to the laboratory for immediate handling. Samples were washed with cold HBSS, visible necrotic areas were removed with a scalpel and the tissue was sectioned into 1–2 mm pieces (Figure 20A and B). Following several washes, samples were digested with 20 mL DMEM + 0.26 U/mL Liberase DH (Dispase High) solution (Roche Diagnostics) in a 37 °C

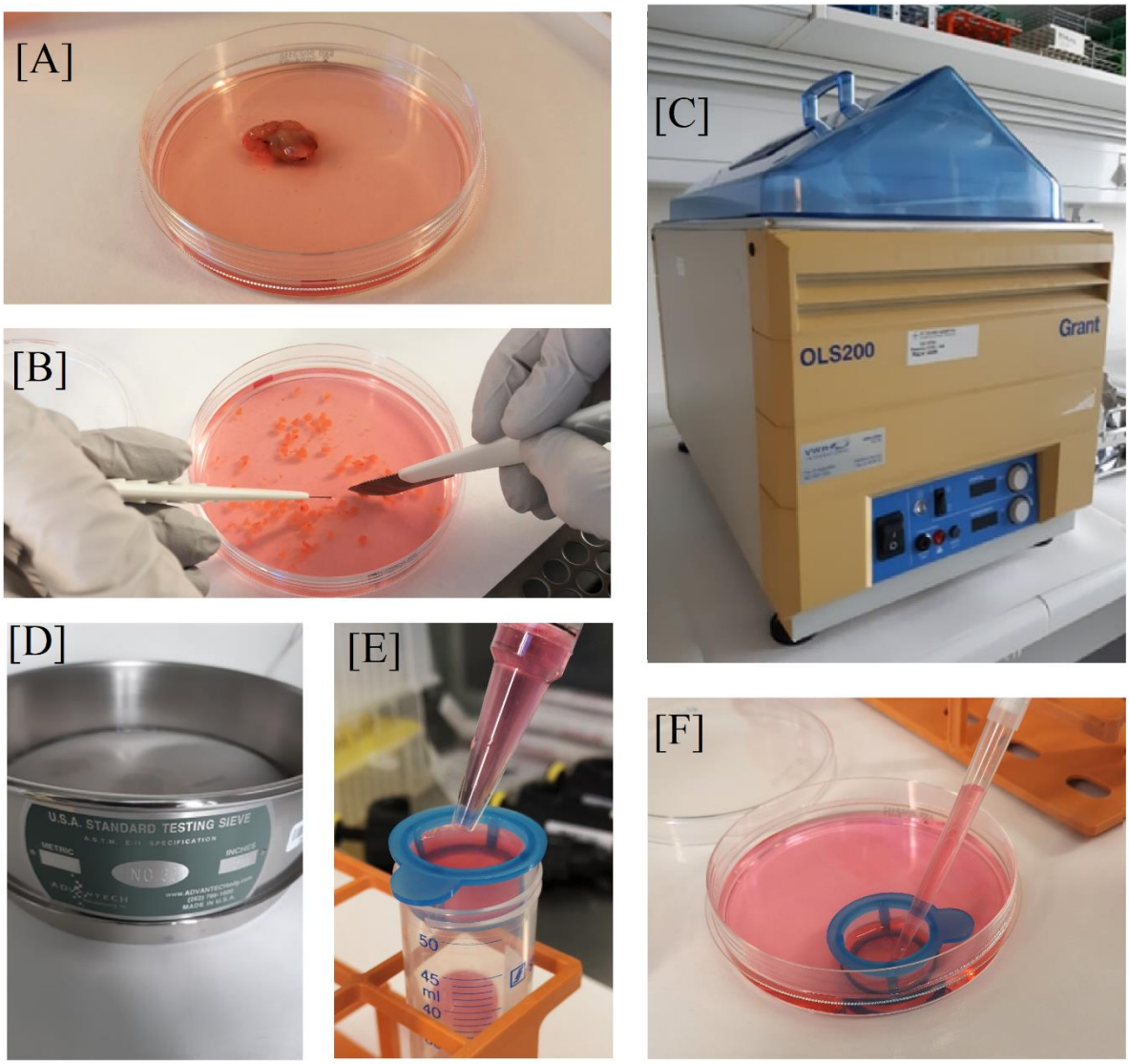


***Figure 20**. Preparation of cancer-tissue originated spheroids (CTOSs). **[A]** Tissue is transported to the laboratory in DMEM on ice for immediate handling. **[B]** Visible necrotic tissue is removed with a scalpel and the tissue is sectioned into 1–2 mm pieces. **[C]** Tissue is digested in Liberase DH solution in a 37 °C shaking water bath for 2 hours. **[D]** Digested tissue is then filtered through a 500 µm steel wire mesh, and **[E]** the filtrate is filtered through a 40 µm cell strainer. **[F]** Samples remaining in cell strainer are collected and transferred to serum-free stem cell medium. Tissue is washed with HBSS between each step.*

shaking water bath for 2 hours (Figure 20C). Liberase DH is a blend of digestion enzymes, and the enzyme concentration and the digestion time have been optimized by Kondo et al. to avoid over-digestion. The samples were then filtered through a 500 µm steel wire mesh (Sigma Aldrich, Figure 20D), and the filtrate was filtered through a 40 µm cell strainer (Corning Life Science, Figure 20E). Samples remaining in the cell strainer were collected (Figure 20F), washed with HBSS, and transferred to serum-free stem cell medium in a non-treated tissue culture dish. The samples were then incubated overnight in a 5% CO_2-humidified chamber at 37 °C. For complete protocol, see Supplementary section 4.2.1.

3.11.2.2. Cancer-tissue originated spheroids - culture and expansion

After 24 hours of incubation in a 5% CO_2 incubator at 37 °C, CTOSs were embedded in Cellmatrix droplets consisting of 70% Cellmatrix Type 1-A (Nitta Gelatin Inc.), 20% DMEM (Gibco) and 10% reconstitution buffer (50 mM NaOH, 260 mM NaHCO3, 200 mM Hepes), in a non-treated dish (Figure 21A). After adding the CTOS to the droplets, the gel was solidified for 30 minutes at 37 °C, and warm serum-free stem cell medium was added (Figure 21B). CTOSs were incubated for 2-3 weeks, and medium was changed every third day.

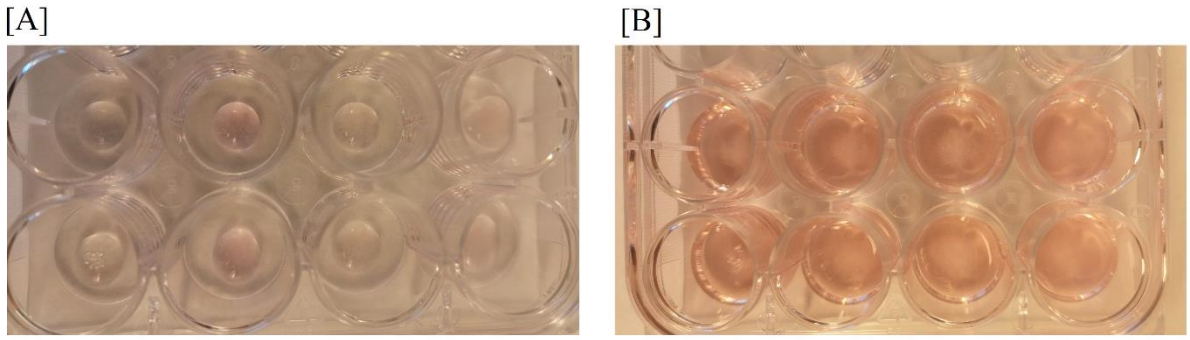


Figure 21. [A] *Cancer-tissue originated spheroids (CTOSs) embedded in Cellmatrix droplets consisting of 70% Cellmatrix Type 1-A (Nitta Gelatin Inc.), 20% DMEM (Gibco) and 10% reconstitution buffer (50 mM NaOH, 260 mM NaHCO3, 200 mM Hepes).* ***[B]*** *After30 minutes of solidification at 37 °C, warm serum-free stem cell medium is added to the droplets.*

After 2-3 weeks of cultivation, CTOSs were released from the Cellmatrix by incubating with 3 mL DMEM and 0.2 mg/mL Collagenase Type IV (Worthington) for 1 hour, followed by a washing step with PBS. For expansion, CTOSs were cut into 2-4 pieces using 23-gauge needles, and the CTOSs were transferred to serum-free stem cell medium. After 24 hours of incubation at 37 °C, the newly formed CTOSs were again embedded in Cellmatrix droplets, and culture was repeated. For complete protocol, see Supplementary section 4.2.2.

3.11.2.3. Cancer-tissue originated spheroids - modifications

A few issues were experienced in the initial phase. First of all, the digestion of samples with Liberase DH was incomplete. A magnetic stirrer, which was used by Kondo et al. was not accessible in the initial phase, and the samples were instead digested in a shaking water bath without magnetism. Furthermore, the gel matrix was fragile and tended to disintegrate. The gel was sensitive to high temperature and tended to solidify before adding it to the plates. Lastly, the manual expansion procedure using needles was laborious and not suitable for large samples.

Several modification experiments were done during the initial phase in an attempt to increase the yield of CTOS; Different volumes of CTOSs in serum-free stem cell medium (10 µL, 20 µL, 30 µL) were added to the gel to evaluate which volume resulted in the most stable gel (CRC-01); In contrast to Kondo et al., to make the gel less fragile, the gel was added in one layer instead of two by adding 130 µL gel directly to the plate (CRC-01-CRC-10); Instead of mixing Cellmatrix Type 1-A, DMEM and reconstitution buffer (RB) in a ratio of 7:2:1, the fraction of RB was increased from 10% to 50% to make the gel thicker, giving the ratio 5:0:5 (CRC-06); To make the gel less fragile, CTOS were added to the gel prior to adding the gel to each well (CRC-07); To be able to study the microbiota, samples were treated with non-antibiotic medium instead of supplementary serum-free stem cell medium (CRC-08 and CRC-09); Mechanical disruption of CTOS after digestion with Liberase DH by using a syringe (10 mL) and a 1.1 needle, as well as centrifuging the CTOSs and resuspending them in 400 µL serum-free stem cell medium was done to increase the number of CTOSs before adding them to the gel (CRC-10).

3.11.3. Phase 2 – Optimization

A modified protocol of Kondo et al. developed by Jeppesen et al.[160] uses Collagenase II (Gibco™) instead of Liberase DH as the digestion enzyme. Traditional collagenase is isolated from *Clostridium histolyticum* and has great proteolytic activity in its ability to break down collagen fibrils commonly found in connective tissue[161]. Liberase DH contains highly purified Collagenase I and Collagenase II in precise ratio for high specific activity and effective dissociation of primary tissues. In addition, Liberase DH contains a high concentration of Dispase, a non-clostridial neutral protease[161]. To investigate the effect of Liberase DH digestion on the yield and viability of the CTOS compared to the traditional Collagenase II, the tissue sample was divided into two pieces and digested with either Collagenase II or Liberase DH

(Figure 22). Each piece was then embedded into four different gels that were provided and qualified specifically for use in 3D culture studies; Cellmatrix Type 1-A, which also was used in the initial phase; Matrigel (Corning Life Science); Geltrex (Gibco™); and Cultrex (R&D Systems). Matrigel was used by Jeppesen et al. and is also one of the most commonly used gels in spheroid 3D culture experiments[162]. There are, on the other hand, relatively few published studies about the latter two gels in spheroid 3D cultures. Furthermore, as the expansion procedure with needles was very laborious and not suitable for larger samples, the sample was split using the digestion enzyme Collagenase II. In addition, a freezing and thawing procedure was established for the spheroids to be stored for longer periods of time. The aim of this optimization was to determine which cultivation condition is most suitable for the CTOS.

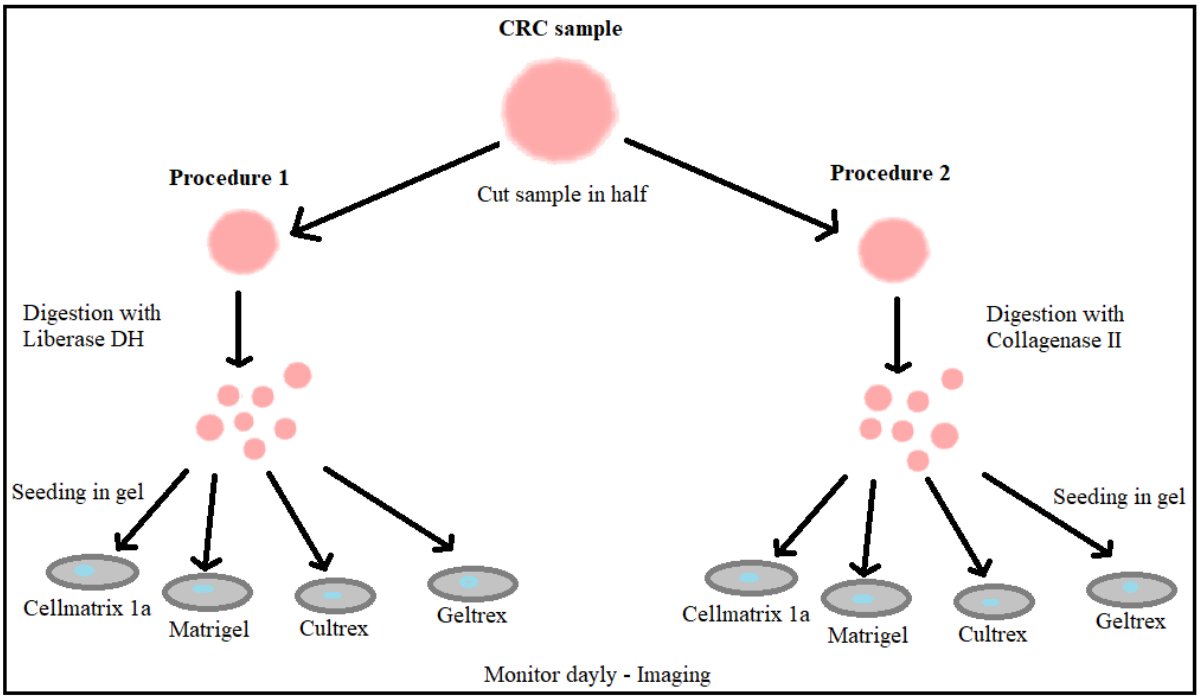


Figure 22. *Suggested optimization of the protocol developed by Kondo et al. The tissue sample is cut in half and digested with either Liberase DH or Collagenase II. The digested tissue is seeded into four types of gels; Cellmatrix Type 1-A, Matrigel, Cultrex and Geltrex. Modified from Evelina Folkesson, 2019 (unpublished work).*

3.11.3.1. Optimization procedure for spheroid preparation and culture

Tissue sample was cut in half after tumour resection and divided into two tubes labelled "L" sand "C" with cold supplemented DMEM. The CTOS preparation was done as previously described in Subheading 3.11.2.1, except that tissue sample "L" was digested with Liberase DH for 2 hours and tissue sample "C" was digested with Collagenase II for 20 minutes. Enzyme digestion was done in a water bath at 37 °C with magnetic stirring instead of a shaking water

bath to increase the CTOS yield (Figure 23A). In addition, a 500 µm pluriStrainer (pluriSelect) filter was used instead of a 500 µm steel wire mesh to make the filtration less laborious (Figure 23B). CTOS culture was done as previously described in Subheading 3.11.2.2, except that both samples were embedded in four different gels, Cellmatrix Type 1-A, Matrigel, Cultrex and Geltrex, in 2 technical replicates in a 24-well plate. For complete protocol, see Supplementary section 4.2.1 for CTOS preparation and 4.2.3 for CTOS culture.

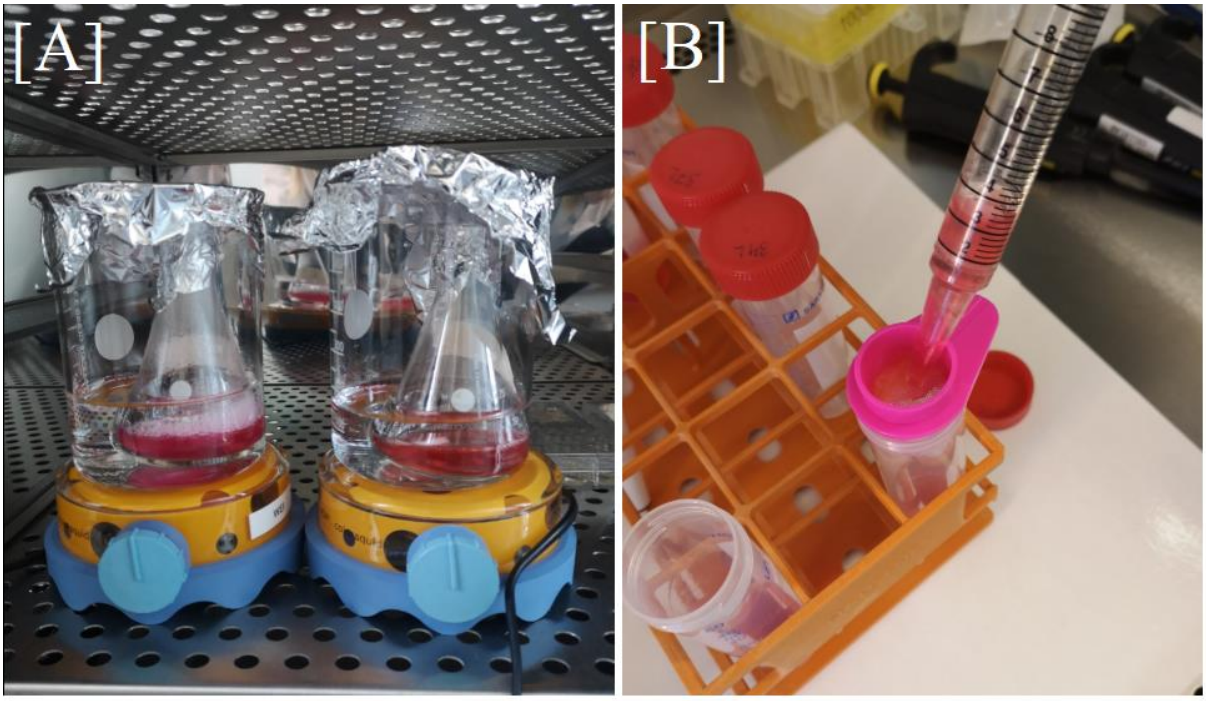


Figure 23. [A] *Tissue sample is divided in two pieces, "L" and "C". Tissue sample "L" is digested with Liberase DH for 2 hours and tissue sample "C" is digested with Collagenase II for 20 minutes. Enzyme digestion is done in a water bath at 37 °C with magnetic stirring.* ***[B]*** *Digested tissue is then filtered through a 500 µm pluriStrainer*

3.11.3.2. Optimization procedure for spheroid expansion, freezing and thawing

After cultivation, CTOSs were released from the gels by incubating with DMEM and 0.2 mg/mL Collagenase Type IV (Worthington) for 1 hour, followed by a washing step with PBS. The CTOSs were further digested (only with Collagenase type II), filtered and washed in HBSS as previously described. The sample was then resuspended in HBSS and separated into two tubes, one for expansion and one for freezing. For expansion, the CTOSs were spun down and resuspended in an appropriate volume of DMEM for the number of gels to be seeded out (20% of total volume). Then Cellmartrix type 1-A (70% of total volume) and RB (10% of total volume) was added to the CTOSs. The samples were seeded (50 µL) in a prewarmed 24-well plate and incubated for 30 minutes at 37 °C for the gel to solidify. Warm serum-free stem cell

medium was added to each well and incubated overnight. Medium was changed the following day and every third day afterwards. For freezing of the sample, CTOSs were spun down and resuspended in an appropriate volume of FCS for achieving a spheroid density of 210 spheroids/mL. Furthermore, 5% DMSO was added to the sample and immediately put in a cryovial in an isopropanol box and stored at -80 °C until use. For thawing of CTOSs, the cryovial was transferred from the -80 °C freezer and immediately placed in a 37 °C water bath for thawing. The cryovial was decontaminated with 70% ethanol and aseptically transferred to a biosafety cabinet. The entire sample was transferred to a conical tube containing serum-free stem cell medium, spun down and resuspended properly in 70% Cellmatrix Type 1-A, 20% DMEM and 10% RB. The sample was quickly plated out as gel droplets in a pre-warmed 24-well plate and incubated at 37 °C for 30 minutes. Afterwards, serum-free stem cell medium was added to the well and incubated overnight. Medium was changed the following day and every third day afterwards. For complete protocol, see Supplementary section 4.2.4.

4. Results

The results section has been divided into three parts. The first part describes *F. nucleatum* and how it promotes migration, proliferation and gene expression in DLD-1. This includes confocal microscopy, migration- and proliferation assay, ELISA and two-step RT-qPCR. The second part describes the function of miR-BART10-3p in gene silencing of SW620. This includes two-step RT-qPCR and the luciferase assays Dual Luciferase Assay and LightSwitch Luciferase Assay. Finally, the third part described the development of a procedure to make patient-derived CRC spheroids by optimizing a protocol developed by Kondo et al.[92].

Due to COVID-19 in Norway, there was limited time and limited access to the laboratory. This resulted in few or none biological replicates of the *F. nucleatum*-related qPCR experiments and the proliferation assay, of which some of them lack statistical analysis. The number of replicates is indicated for each figure. The results for each assay are presented as one representative graph, and additional raw data and images are presented in the supplementary section.

4.1. Identification of candidate genes for *F. nucleatum* and miR-BART10-3p

To identify genes that respond to *F. nucleatum* in CRC, a gene expression analysis of CRC cell lines co-cultured with *F. nucleatum* was performed. Results are presented in Supplementary section 1.1. Results include PCA plots showing differentially expressed mRNA between different CRC cell lines (DLD-1, LS411N, SW480 and SW620) and treatments (*E. coli* DH5α, *F. nucleatum* or TSB) in which there was a large natural variation in mRNA expression to treatment between cell lines (Figure S1-S2). In addition, results include a heatmap showing differentially expressed mRNA in SW620 cells treated with *E. coli* DH5α, *F. nucleatum* or TSB in which *E. coli* DH5α-treated cells clustered closer together with *F. nucleatum*-treated cells, indicating that their gene expression profiles are more similar (Figure S3). In co-culture with *F. nucleatum*, DLD-1 cells were the easiest to handle, i.e., DLD-1 had a high growth rate and a relatively low rate of cell loss, and the cell line was therefore used in subsequent experiments. Two candidate genes, *CSF2* and *CCL20*, which were among the top upregulated genes after *F. nucleatum* infection in DLD-1 compared to *E. coli DH5*α infection (Table 4) and after *F. nucleatum* infection in DLD-1 compared to TSB treatment (Table 5) were selected for further validation. Average expression shows the mean expression level for that specific gene. *CSF2* and *CCL20* had a positive log fold change, indicating that the genes are upregulated. The p-values were adjusted according to the Benjamini-Hochberg[163] method. The adjusted p-values

were larger than 0.05 for the respective genes and are therefore not significant. However, with further validation, this information is valuable.

***Table 4**. Gene expression profiles for CSF2 and CCL20 in F. nucleatum-treated DLD-1 compared to DH5α-treated DLD-1 showing the ENSEMBL ID, gene symbol, average expression (count per million, cpm), Benjamini-Hochberg adjusted p-value and log fold change.*

ENSEMBL	Gene symbol	Average expression	Adjusted p-value	Fold change (log2)
ENSG00000164400	*CSF2*	0.56	0.99893536470974	1.4330560082117
ENSG00000115009	*CCL20*	2.29	0.99956746051789	1.2310243344971

***Table 5.** Gene expression profiles for CSF2 and CCL20 in F. nucleatum-treated DLD-1 compared to TSB-treated DLD-1 showing the ENSEMBL ID, gene symbol, average expression (count per million, cpm), Benjamini-Hochberg adjusted p-value and log fold change.*

ENSEMBL	Gene symbol	Average expression	Adjusted p-value	Fold change (log2)
ENSG00000164400	*CSF2*	0.80	0.29492796350278	1.4508365673800
ENSG00000115009	*CCL20*	1.79	0.06143283668985	2.4895262318319

To identify human targets of miR-BART10-3p, transient transfection of miR-BART10-3p mimics in CRC cell lines was performed. Results are presented in Supplementary section 1.2. Results include PCA plots showing differentially expressed mRNAs in SW620 and LS411N transfected with miR-BART10-3p or negative miRNA control showing that the variation is higher for the negative control samples compared to the miR-BART10-3p transfected cells (Figure S4). In addition, the results include volcano plots showing differentially expressed mRNA in SW620 and LS411N transfected with miR-BART10-3p or negative miRNA control indicating that miR-BART10-3p target many mRNAs (Figure S5). Two candidate genes, *MAT2B* and *ELL2*, which were downregulation after miR-BART10-3p transfection of SW620 (Figure S6) and LS411N (Figure S7) were selected for further validation since they were predicted targets of miR-BART10-3p at Targetscan[164] by using their seed sequence (Table 6). SW620 cells had a high transfection efficiency and growth rate compared to LS411N and was therefore used in subsequent experiments.

Table 6. *Targetscan results showing MAT2B and ELL2 (with ENSEMBL ID) as predicted targets for miR-BART10-3p. "MSA start" and "MSA end" are the start and end positions of the target site relative to the UTR start and end position, counting alignment gaps; "UTR start" and "UTR end" are the start and end positions of the target site relative to the UTR start and end position, not counting alignment gaps; "Site type" is the type of matching between miRNA and target, i.e., "8mer" indicates an exact match to positions 2-8 of the mature miRNA (the seed + position 8) followed by an 'A', "7mer-1a" indicates that positions 2-7 of the miRNA match, followed by an 'A', "7mer-m8" indicates that there is an exact match to positions 2-8 of the miRNA, and "6mer" indicates an exact match to positions 2-7 of the mature miRNA.*

ENSEMBL	Gene symbol	miRNA	MSA start	MSA end	UTR start	UTR end	Site type
ENST00000280969	*MAT2B*	miR-BART10-3p	1190	1199	517	524	8mer
ENST00000237853	*ELL2*	miR-BART10-3p	6949	6956	3441	3446	6mer
ENST00000237853	*ELL2*	miR-BART10-3p	6434	6444	3178	3184	7mer-1a
ENST00000237853	*ELL2*	miR-BART10-3p	2518	2543	1291	1297	7mer-m8
ENST00000237853	*ELL2*	miR-BART10-3p	2713	2718	1394	1399	6mer

Gene expression profiles for *ELL2* and *MAT2B* in miR-BART10-3p transfected SW620 compared to untreated SW620 are shown in Table 7. Average expression shows the mean expression level for that specific gene. *MAT2B* and *ELL2* had a negative log fold change, indicating that the genes are downregulated. The p-value was adjusted according to the Benjamini-Hochberg method[163]. The adjusted p-value was larger than 0.05 for *ELL2* and the altered gene expression is therefore not significant. This is likely due to few replicates (n=3) in the RNA-seq experiments. However, with further validation, this information is valuable.

Table 7*. Gene expression profiles for ELL2 and MAT2B in miR-BART10-3p transfected SW620 compared to untreated SW620 showing the ENSEMBL ID, symbol, average expression, Benjamini-Hochberg adjusted p-value and log fold change.*

ENSEMBL	Gene symbol	Average expression	Adjusted p-value	Fold change (log2)
ENST00000237853	*ELL2*	4.69164201536	0.064117760663	-0.9290448566534
ENSG00000038274	*MAT2B*	6.58733332225	0.008803176410	-1.0060937477897

4.2. The effect of *F. nucleatum* on CRC cell lines

4.2.1. *F. nucleatum* did not induce proliferation and migration in DLD-1

The effect of *F. nucleatum* in CRC was investigated by co-culturing the bacteria with DLD-1 and look for changes in proliferation and migration over time. Migration was monitored using a Wound Healing and Migration assay from Ibidi which is reported to be a reproducible assay owing to a defined 500 µm cell free gap with no leakage during cultivation. Figure 24 shows the percent wound closure at each time point for four biological replicates and Figure 25 shows images for one biological replicate taken with a phase contrast microscope. Compared with untreated cells, *F nucleatum* did not promote cell migration in DLD-1 after treatment at 24 hours (P = 0.415), 48 hours (P = 0.201) or 72 hours (P = 0.304). However, by observing the images in Figure 25 there was an indication that cells cultured with *F. nucleatum* were migrating faster than cells cultured without *F. nucleatum*. Additional information, including all images and calculations, are available in Supplementary section 5.

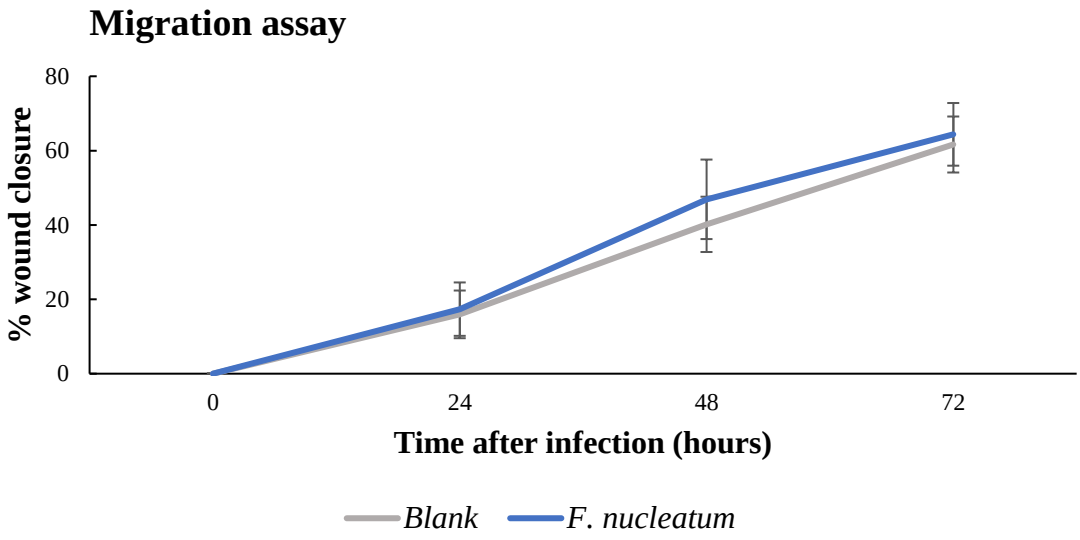


***Figure 24**. Percent wound closure of DLD-1 that is untreated (Blank) or treated with F. nucleatum at a MOI of 300. Results are the means for each time point ± standard deviations (SD) of four biological replicates.*

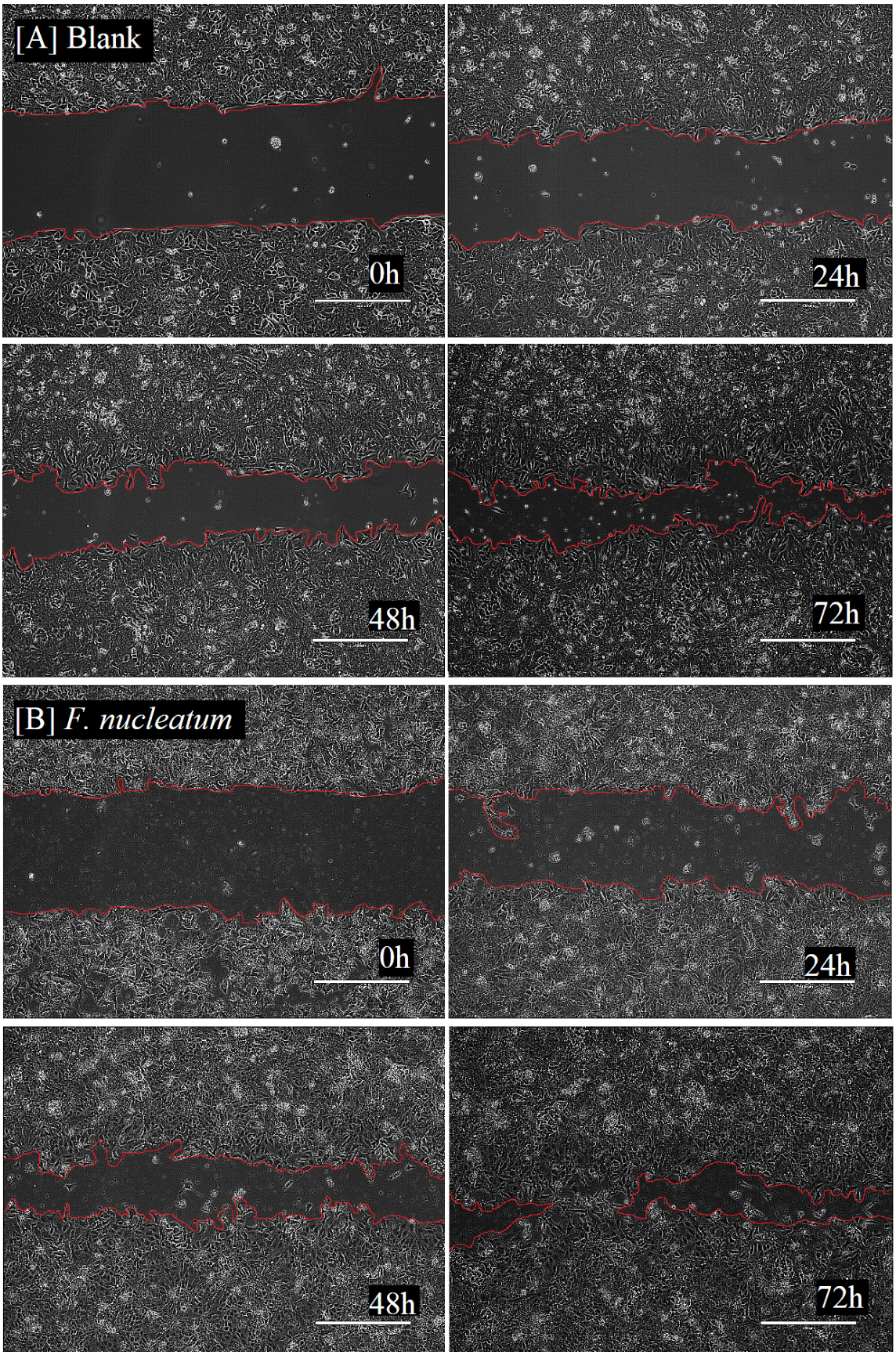


Figure 25. *Wound-healing and migration assay of DLD-1 that is untreated (Blank)* ***[A]*** *or infected with F. nucleatum* ***[B]*** *at a MOI of 300. The wound size was measured at 0, 24, 48 and 72 hours. DLD-1 cells were photographed at 10X magnification in a phase contrast microscope. Scale bar: 200 µm.*

Cell proliferation assay was monitored using the µ-Slide 8 Well Grid-500 from Ibidi. Figure 26 shows images for one technical replicate taken with a phase contrast microscope, and Figure 27 shows percent proliferation at each time point for four technical replicates. Compared with untreated cells, *F nucleatum* did not promote cell growth in DLD-1 after treatment at 24 hours (P = 0.277) or 48 hours (P = 0.071). However, since there is only one biological replicate with four technical replicates, we cannot tell the statistical significance of the results. Additional information, including all images and calculations, are available in Supplementary section 6.

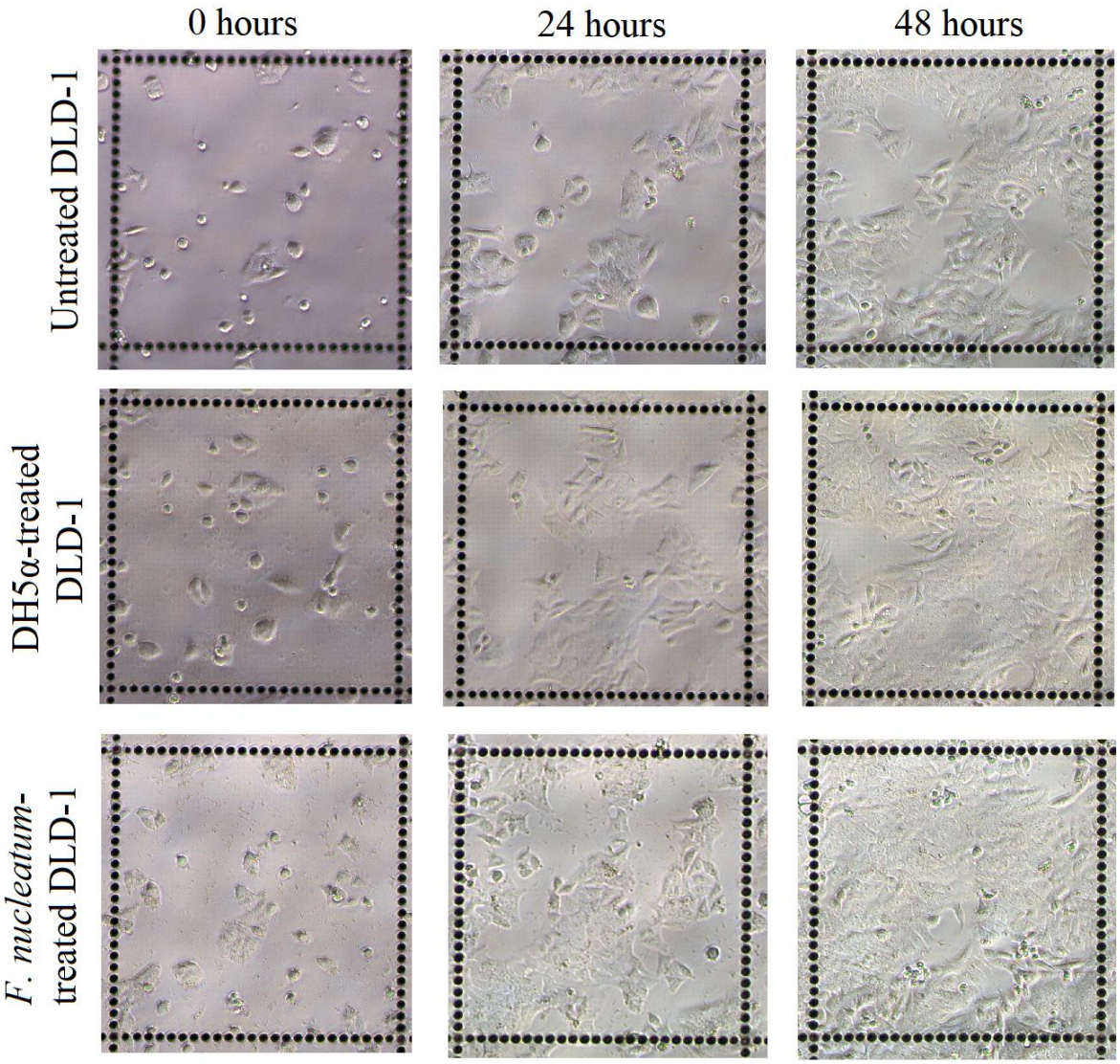


Figure 26. *Proliferation assay of untreated-, E. coli DH5α-treated- or F. nucleatum-treated DLD-1 cells at a MOI of 500. The cell-free area was measured at 0, 24 and 48 hours. DLD-1 cells were photographed at 20X magnification in a phase contrast microscope. Size of grid: 500 µm.*

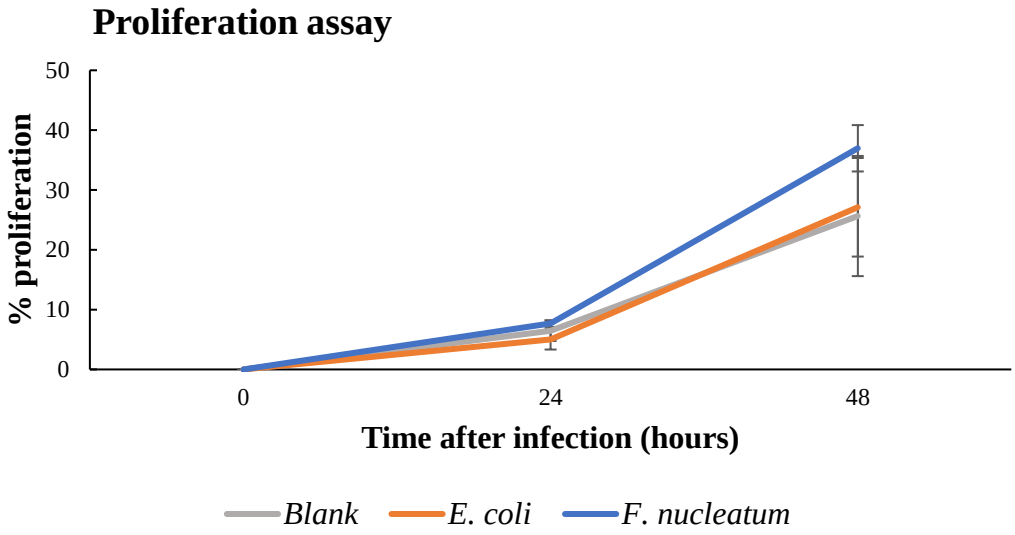


Figure 27. *Percent proliferation of untreated (Blank), E. coli DH5α-treated or F. nucleatum-treated DLD-1 at a MOI of 500. Results are the means for each time point ± standard deviations (SD) of four technical replicates.*

4.2.2. Intracellular localization of *F. nucleatum* in DLD-1 cells

It has been demonstrated that *F. nucleatum* adheres to and invades human epithelial cells, activates the β-catenin signalling pathway, induces oncogenic gene expression and promotes growth of CRC cells via the fusobacterial adhesin FadA[47]. An intracellular localization is considered advantageous for bacteria to evade acquired immunity, as well as antibiotic pressure, leading to intracellular persistence, proliferation and invasion of adjacent tissues[165]. Because the effect of *F. nucleatum* has persisted even though the bacteria were removed from the cell culture after 6 hours of incubation and then incubated with media containing antibiotics, we determined whether the bacterium was maintained intracellularly. *F. nucleatum* infection of DLD-1 cells was confirmed by confocal fluorescence microscopy, which showed intracellular *F. nucleatum* aggregates after 2 hours of incubation (Figure 28A and B). *F. nucleatum* was stained with CFSE (red) prior to the infection. Fixed DLD-1 cells were stained with Rhodamine Phalloidin F-actin staining (green) and DAPI staining solution (blue) for the actin filaments and DNA, respectively.

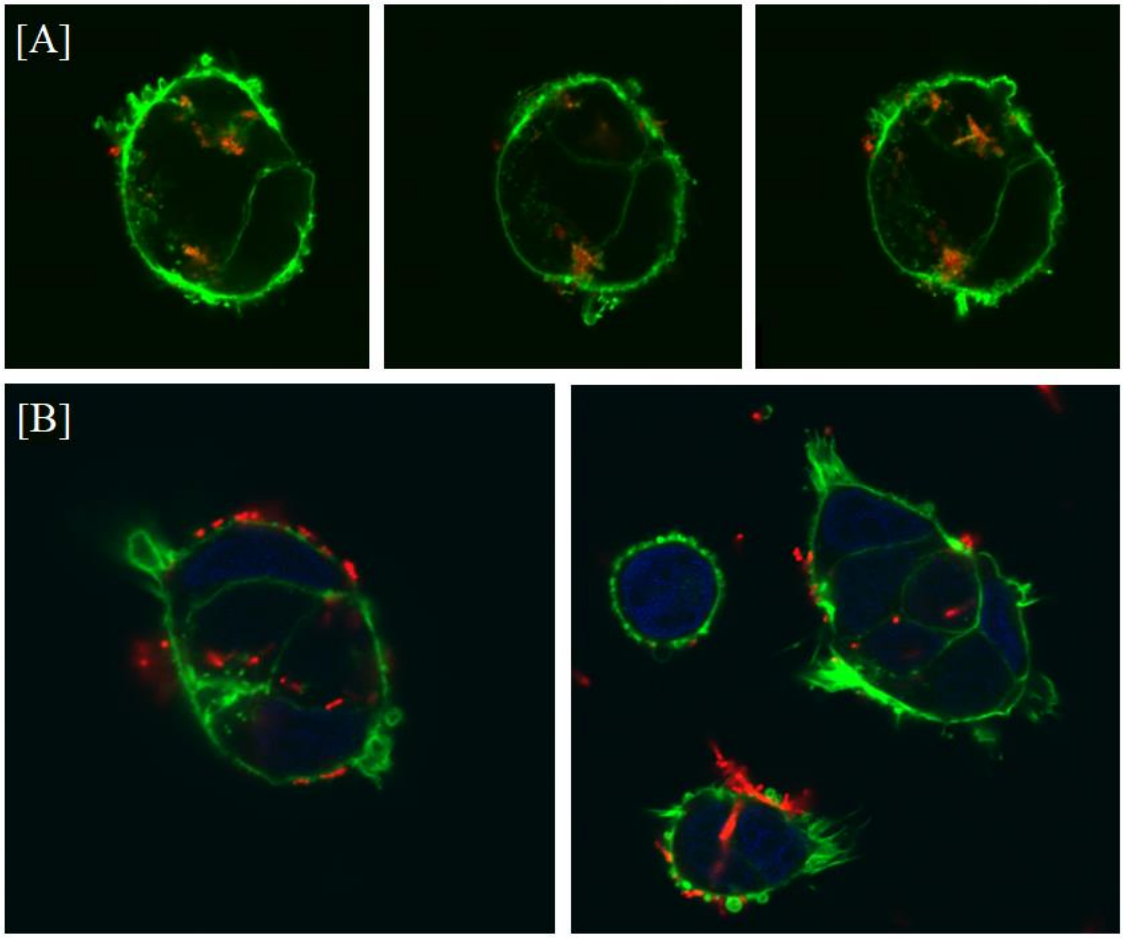


Figure 28. *Intracellular localization of F. nucleatum (500 MOI) in DLD-1 at 2 hours post-infection. DLD-1 cells were photographed at 63X magnification in an immunofluorescence confocal microscopy. Fixed DLD-1 was stained with Rhodamine Phalloidin F-Actin staining (green) to show actin filaments, and F. nucleatum was stained with CFSE staining solution (red) to show the localization of the bacteria.* ***[A]*** *Different optical sections through the middle of the cell shows the intracellular localization of the bacteria.* ***[B]*** *DNA was stained with DAPI staining solution (blue) to show the nucleus of the cells.*

Furthermore, by stacking an image of DLD-1 from different optical sections, a 3D structure was made (Figure 29). The section through the middle of the cell confirms the intracellular localization of the bacteria.

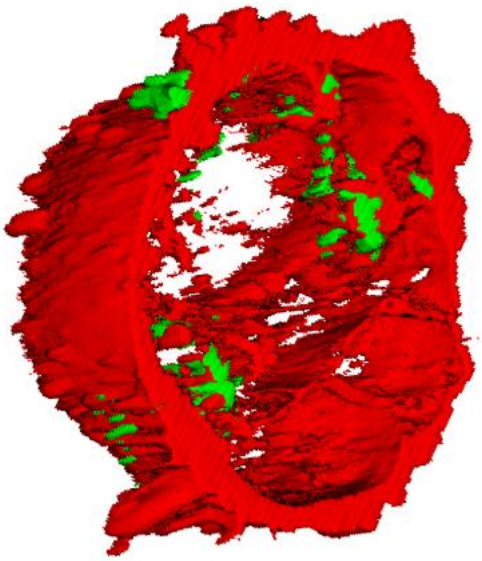


Figure 29. *3D-model of DLD-1 infected with F. nucleatum, reconstructed from a confocal z-stack. The actin filaments are displayed in red and F. nucleatum in green.*

4.2.3. Effect of *F. nucleatum* on *CSF2* and *CCL20* expression by DLD-1

In normal colonic mucosa, both *CCL20* and *CSF2* are weakly expressed. However, in response to an inflammatory stimulus, the expression of *CCL20* and *CSF2* are significantly increased. Both *CCL20* and *CSF2* play important roles in the recruitment of immune cells. Previous studies have shown that aberrant infiltration of immune cells and subsequent inflammation may induce tumour progression, invasion, and metastasis[113,166]. Here, we want to examine the level of *CCL20* and *CSF2* in DLD-1 after infection of *F. nucleatum*. DLD-1 cells were treated with *F. nucleatum* (500 MOI) for 0, 3, 6, 12 and 24 hours, and RNA was extracted for further detection by two-step RT-qPCR. The mRNA expression of *CCL20* was higher in *F. nucleatum*-treated DLD-1 cells at all time points compared to non-treated DLD-1 (Figure 30A). However, since there is only one biological replicate with three technical replicates, we cannot tell the statistical significance of the results. Furthermore, the release of CCL20 protein in cell-free supernatants at the indicated time points (excluding 3 hours) were investigated by doing a Quantikine ELISA assay. The Quantikine ELISA assay demonstrated an increase in CCL20 release at 6 hours following *F. nucleatum* infection, with a plateauing of CCL20 release 12 hours post-infection (Figure 30B). The release of CCL20 was, however, not significant at 6 hours ($P = 0.109$), 12 hours ($P = 0.163$) or 24 hours ($P = 0.163$) post-infection due to the high standard deviation. Additional information, including raw data and calculations, are available in Supplementary section 7.1 (qPCR) and Supplementary section 8.1 (Quantikine ELISA).

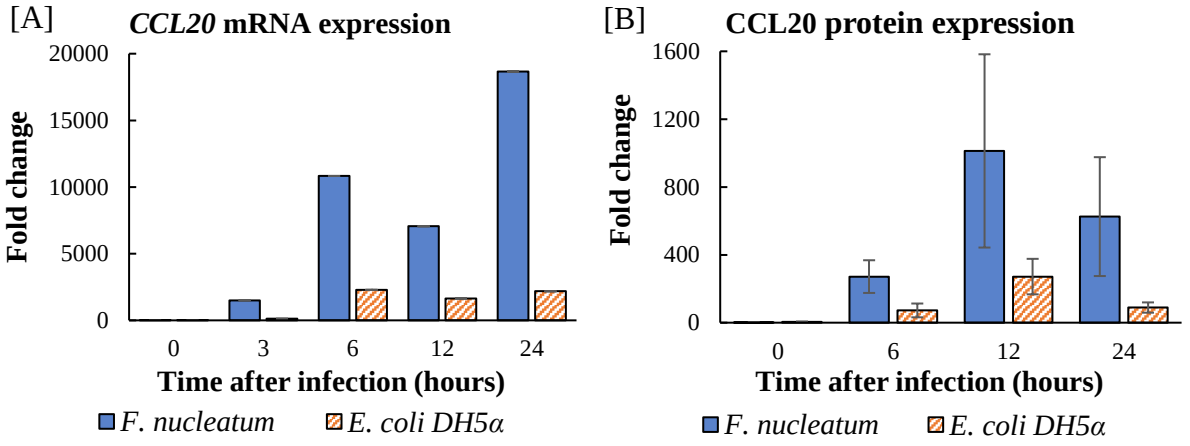


Figure 30. *Time course of CCL20 in F. nucleatum-treated- or E. coli DH5α-treated DLD-1.* ***[A]*** *The amount of CCL20 mRNA was measured by two-step RT-qPCR. CCL20 was normalized to the housekeeping gene ACTB and the untreated sample. Results are the means ± standard deviations (SD) of three technical replicates.* ***[B]*** *The amount of CCL20 protein in cell-free supernatants was measured by ELISA. CCL20 was normalized to the untreated cells. Results are the means ± standard deviations (SD) of two biological replicates.*

The mRNA expression of *CSF2* was higher in *F. nucleatum*-treated DLD-1 cells at all time points compared to non-treated DLD-1 (Figure 31). The fold increase in *CSF2* from *F. nucleatum*-treated cells compared to untreated cells, reached its maximum at 24 hours post-infection. However, as there is only one biological replicate with three technical replicates, we cannot tell the statistical significance of the results. Additional information, including raw data and calculations, are available in Supplementary section 7.2.

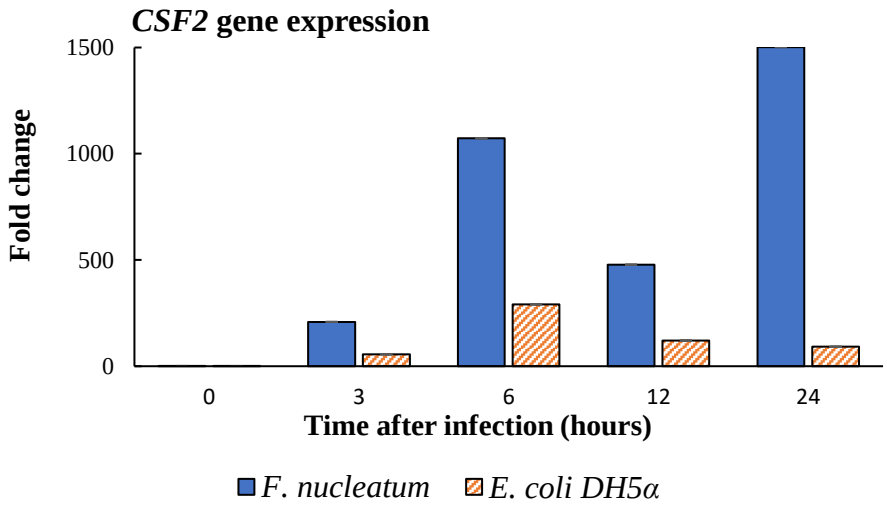


Figure 31. *Time course of F. nucleatum-induced expression of CSF2 mRNA in DLD-1. DLD-1 was infected with F. nucleatum or E. coli DH5α for the indicated time periods, and the amount of CCL20 mRNA was measured by two-step RT-qPCR. CSF2 was normalized to the housekeeping gene ACTB and the untreated sample. Results are the means ± standard deviations (SD) of three technical replicates.*

4.2.4. *F. nucleatum* induces release of CCL20 in a dose-dependent manner

To determine the minimum concentration of *F. nucleatum* to induce CCL20 release, DLD-1 was infected with increasing concentrations of *F. nucleatum* (0 MOI, 0.1 MOI, 1 MOI, 10 MOI and 100 MOI) for 6 hours, and the release of CCL20 protein in cell-free supernatants was investigated by a Quantikine ELISA assay. The release of CCL20 was significantly increased at both 10 MOI (P = 0.002) and 100 MOI (P = 0.045) treatment of *F. nucleatum* compared to 0 MOI treatment of *F. nucleatum*. However, apart from 100 MOI treatment, the release of CCL20 in *F. nucleatum*-treated cells was not stronger than the release of CCL20 from non-pathogenic *E. coli* DH5α (Figure 32). Additional information, including raw data and calculations, are available in Supplementary section 8.2.

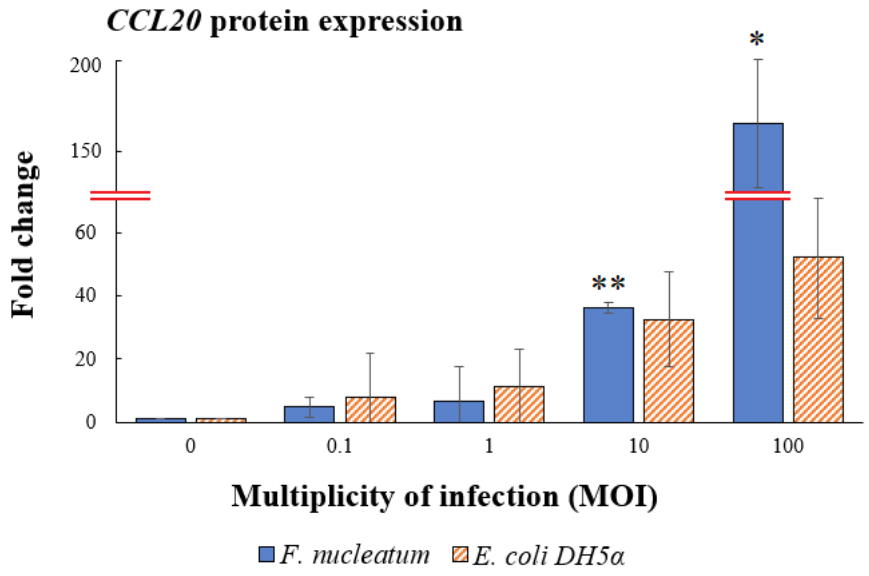


***Figure 32**. CCL20 release in DLD-1 after treatment with F. nucleatum and E. coli DH5α at different concentrations. DLD-1 was infected with F. nucleatum and E. coli DH5α at the indicated concentrations, and the amount of CCL20 in cell-free supernatants was measured by ELISA. CCL20 was normalized to the untreated cells (0 MOI). Results are the means ± standard deviations (SD) of two biological replicates; *p < 0.05, **p < 0.005*

4.3. EBV miRNA induces downregulation of *ELL2* and *MAT2B* in SW620

EBV miRNAs, including miR-BART10-3p, has been shown to target host mRNAs and genes involved in CRC development. In Subheading 4.1. we showed that *MAT2B* and *ELL2* were downregulated in SW620 cells after transfection with miR-BART10-3p. The altered genes have been studied in a variety of epithelial cells in which knockdown of *ELL2* has been shown to increase proliferation, migration, and invasion, whereas knockdown of *MAT2B* has been shown

to inhibit proliferation and migration. Here, the knockdown of *MAT2B* and *ELL2* in SW620-transfected cells were validated. SW620 was transfected with miR-BART10-3p or negative miRNA for 24 hours, and RNA was extracted for further detection by two-step RT-qPCR (Figure 33). *MAT2B* and *ELL2* mRNA levels were significantly reduced upon miR-BART10-3p transfection (p = 0.016 and p = 0.028, respectively).

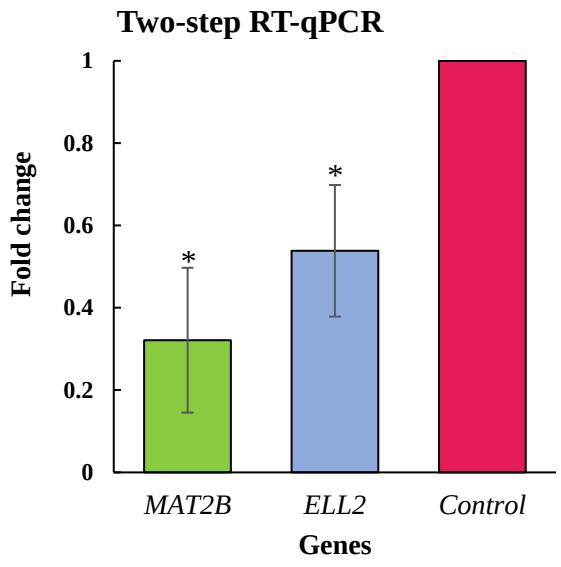


Figure 33. *Two-step RT-qPCR of MAT2B and ELL2 in miR-BART10-3p transfected SW620. SW620 was transfected with miR-BART10-3p or negative miRNA for 24 hours. The genes were normalized to the housekeeping gene ACTB. Results are the means ± standard deviations (SD) of three biological replicates; *p < 0.05. The significance indicates the differences relative to control.*

The genes were further validated using the Dual Luciferase Assay and the LightSwitch Luciferase Assay to determine that the altered gene expression was due to post-transcriptional regulations, in this matter the miR-BART10-3p. The Dual Luciferase Assay was used for the *MAT2B* 8mer target sequence inserted into a psiCHECK™-2 vector, whereas the LightSwitch Luciferace Assay was used for transfection-ready *MAT2B*- and *ELL2* 3'UTR Reporter GoClones. The reporter vector or a negative control was co-transfected with a miR-BART10-3p mimic or a negative miRNA mimic in SW620 cells for 48 hours, and the luciferase activity was measured (Figure 34A and B). *MAT2B* was significantly reduced upon miR-BART10-3p transfection in the LightSwitch Luciferase Assay (p = 0.004), as well as in the Dual Luciferase assay (P = 0.0005), indicating that miR-BART10-3p is a post-transcriptional regulator of *MAT2B*. Interestingly, *MAT2B* was stronger downregulated when using the LightSwitch Luciferase Assay compared to the Dual Luciferase Assay. The expression of *ELL2* was,

however, not significant (P = 0.102). Nevertheless, there was an indication that *ELL2* was downregulated in miR-BART10-3p transfected SW620 cells compared to SW620 cells transfected with negative miRNA mimic. Additional information, including raw data and calculations, are available in Supplementary section 9.1 (qPCR) and Supplementary section 9.2 (Luciferase assays).

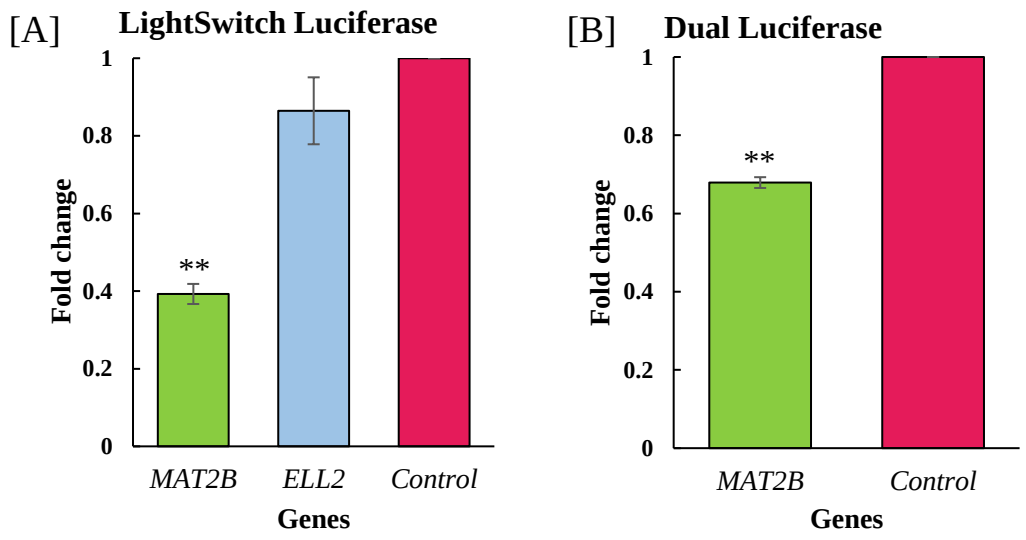


***Figure 34**: Luciferase assays of MAT2B and ELL2 in miR-BART10-3p transfected SW620. For the LightSwitch Luciferase Assay **[A]**, the MAT2B/ELL2 reporter vector or a Random 3'UTR control vector (R04_3UTR) was co-transfected with miR-BART10-3p or a negative miRNA mimic in SW620 for 48 hours, and the luciferase activity was measured. For the Dual Luciferase Assay **[B]**, the MAT2B reporter vector or an empty psiCHECK™-2 control vector was co-transfected with miR-BART10-3p or a negative miRNA in SW620 for 48 hours, and the luciferase activity was measured. Results are the means ± standard deviations (SD) of three biological replicates; **p < 0.005. The significance indicates the differences relative to control.*

4.4. Patient-derived colorectal cancer spheroids

A method to establish patient-derived CRC spheroids was developed. To establish the protocol, ten tissue samples (CRC-01-CRC-10) were retrieved for an initial phase in which the protocol was based on a method developed by Kondo et al.[92]. Several modifications were done in the initial phase. Then, seven tissue samples (CRC-11-CRC-17) were retrieved for an optimization phase in which the protocol was based on a combination between modifications from the initial phase and a modified protocol by Kondo et al., developed by Jeppesen et al.[160] to determine which cultivation condition was most suitable for the CTOSs. Success of sample was defined as ≥ 1 CTOS in the gel 2 weeks after cultivation.

4.4.1. Phase 1 – Initial work

Table 8 shows that most samples have been discarded after some time (8 out of 10) which indicates low success rate in the initial phase (<50%). In the first attempt to make CTOSs (CRC-01), digestion with Liberase DH was incomplete, resulting in few CTOSs in the filtrate. It was difficult to get enough CTOSs inside the gel droplet due to a low yield of CTOSs in the sample. The gel droplet was also very fragile and tended to disintegrate when adding the medium. Therefore, the gel was added in one layer instead of two layers by adding 130 µL gel directly to the plate and let it solidify for 15 minutes before the addition of CTOSs. This resulted in less fragile droplets and was therefore integrated in the protocol for the rest of the samples. In addition, different volumes of CTOSs were added to the gel droplet (10 µL, 20 µL, 30 µL) to observe the fragility of the gel. As the volume of 30 µL was observed to give a high yield and still maintain a stable gel, 30 µL was used for the additional samples. The droplets did however sometimes disintegrate during the incubation time, leading to the loss of CTOSs. In the sixth attempt (CRC-06), a gel composition for which the gel solidifies better was tested. Instead of using a gel ratio of 7:2:1 for Cellmatrix Type 1-A, DMEM and RB, the volume of RB was increased from 10% to 50%, giving the ratio 5:0:5. This resulted in a non-fragile 130 µL gel droplet. The CTOSs did however not tolerate the high amount of RB in the sample and subsequently died following three days in culture. In the seventh attempt (CRC-07), the CTOSs were added to the gel prior to adding the gel to each well, but the gel disintegrated, and no CTOSs were found. In the eighth and ninth attempt (CRC-08 and CRC-09), samples were treated with non-antibiotic medium, but the samples were infected within 24 hours, and had to be discarded. In the tenth attempt (CRC-10), after digestion with Liberase DH, mechanical disruption of CTOSs by using a syringe (10 mL) and a 1.1 needle was tested to digest more of the tissue and get more CTOSs in the filtrate. More tissue was filtered, resulting in a higher yield of CTOSs. Furthermore, CTOSs were centrifuged and resuspended in a smaller volume of serum-free stem cell medium (400 µL) to increase the concentration of CTOSs before adding it to the gel. This resulted in an increased number of CTOSs in the gel and was therefore integrated in the protocol.

Two tissue samples, CRC-01 and CRC-03, were declared established after two weeks of cultivation and both were successfully split and reseeded into new gels. Figure 35 shows CRC-01 five days after the first expansion and CRC-03 three days after expansion in a light microscope.

Table 8. *Observations from each sample in the initial phase. "Sample" is the patient ID; "Observations" is the observed results; "Status" is the current endpoint of the spheroid establishment. "Established" indicates that there are ≥ 1 CTOS in the gel 2 weeks after cultivation.*

Sample	Observations	Status
CRC-01	Low yield of CTOSs (< 10 CTOSs). Split and reseeded twice with high success.	Established
CRC-02	No CTOSs found after seeding to gel.	Discarded
CRC-03	Low yield of CTOSs (< 10 CTOSs). Split and reseeded once with high success.	Established
CRC-04	No CTOSs of appropriate size for expansion.	Discarded
CRC-05	Gel disrupted, CTOSs floated out in the medium and died.	Discarded
CRC-06	CTOSs died following three days in culture due to higher ratio of RB.	Discarded
CRC-07	No CTOSs found after seeding to gel.	Discarded
CRC-08	Infection, abundance of bacteria in medium.	Discarded
CRC-09	Infection, abundance of bacteria in medium.	Discarded
CRC-10	Gel disrupted, CTOSs flowed out in the medium and died.	Discarded

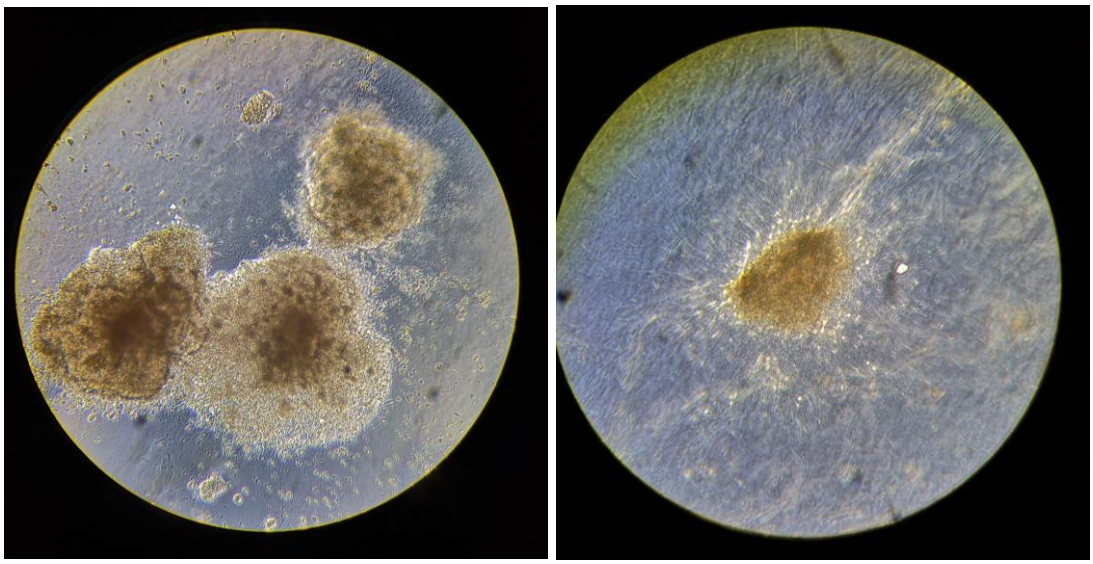


Figure 35. *Images of cancer-tissue originated spheroids (CTOSs) taken in a light microscope at 10X magnification. Left: CRC-01 five days after the first CTOS expansion showing three visible CTOSs in the field of view. Right: CRC-03 three days after expansion showing one visible CTOS in the field of view.*

4.4.2. Phase 2 - Optimization

In the optimization phase, one new digestion enzyme (Collagenase II) and three new gels (Cultrex, Geltrex and Matrigel) were tested in addition to the digestion enzyme (Liberase DH) and gel (Cellmatrix Type 1-A) used in the initial phase. In addition, the tissue sample was split using a digestion enzyme (Collagenase type II) instead of the manual expansion method using a needle. Digestion of tissue sample resulted in a higher yield of CTOSs for both Liberase DH and Collagenase II compared to the initial phase, i.e. there were approximately 0-10 CTOSs in each gel droplet in the initial phase and > 50 CTOSs in each gel droplet in the optimization phase. This indicates, regardless of the enzymes, that the magnetic stirrer, which was not accessible in the initial phase, improves the yield. However, Liberase DH seemed to be gentler with the sample, whereas Collagenase II seemed to give a higher yield of CTOSs (Figure 36A).

By observing the CTOSs in a phase contrast microscope, as shown in Figure 36B, both digestion enzymes partially digest the tissue samples into small fragments which spontaneously form spheroidal shapes. However, some tissue samples seem to be more digested than others, as for CRC-17 which is digested into small fragments and only show the characteristics of a spheroid after several days (Figure 36C). In contrast, CRC-14 shows a spheroidal shape already 2 days after cultivation. Furthermore, CRC-17 shows a dark core within the CTOSs.

As for the gels, most CTOSs have a reasonable round and regular shape, as seen in the example of CRC-14. However, some CTOSs, as seen in the example of CRC-17, have a non-spherical shape. There are observed ellipsoidal shapes, as seen in CRC-17 CM (Collagenase II, Matrigel), and there are observed irregular shapes, as seen in CRC-17 LG (Liberase DH, Geltrex). Furthermore, by observing the example of CRC-17, CTOSs digested in Collagenase II are surrounded by more single cells than CTOSs digested in Liberase DH, indicating cell detachment of the CTOSs. In addition, many single cells were observed in CTOSs seeded in Cultrex.

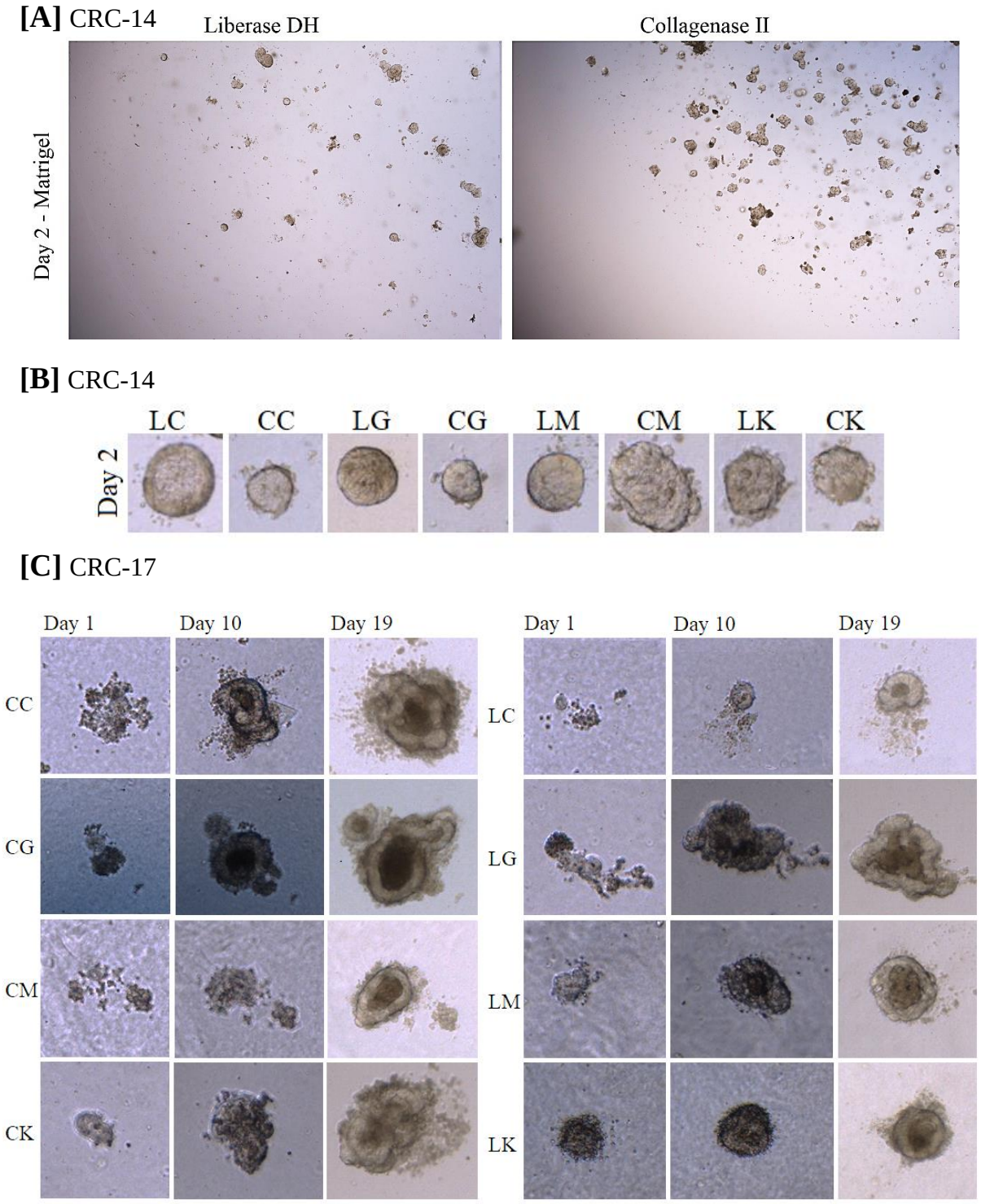


Figure 36. *Cancer-tissue originated spheroids (CTOSs) taken in a phase contrast microscope at 4X magnification.* ***[A]*** *CRC-14 digested with Liberase DH or Collagenase II and seeded in Matrigel. The images are taken 2 days after cultivation and shows one fourth of the gel.* ***[B]*** *CRC-14 digested with Liberase DH or Collagenase II and further seeded in Cultrex, Geltrex, Matrigel or Cellmatrix 1-A. Images are taken 2 days after cultivation.* ***[C]*** *CRC-17 digested with Liberase DH or Collagenase II and further seeded in Cultrex, Geltrex, Matrigel or Cellmatrix Type 1-A. Images are taken at day 1, 10 and 19 after cultivation. LC = Liberase DH, Cultrex; LG = Liberase DH, Geltrex; LM = Liberase DH, Matrigel; LK = Liberase DH, Cellmatrix Type 1-A; CC = Collagenase II, Cultrex; CG = Collagenase II, Geltrex; CM = Collagenase II, Matrigel; and CK = Collagenase II, Cellmatrix 1-A.*

As shown in Figure 37, samples digested with Liberase DH and seeded in Geltrex and Matrigel, as well as samples digested with Collagenase II and seeded in Geltrex, had the highest success rates with 10 out of 14 samples (7 tissue samples seeded in 2 technical replicates) with ≥ 1 CTOS in the gel 2 weeks after cultivation.

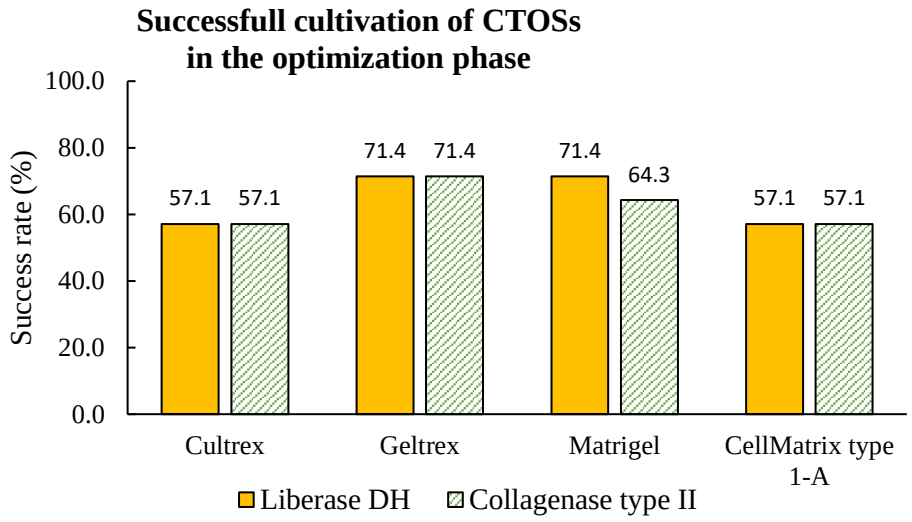


Figure 37. *Successful cultivation of cancer-tissue originated spheroids (CTOS) from CRC-11-CRC-17 digested with Liberase DH or Collagenase II and further seeded in Cultrex, Geltrex, Matrigel or Cellmatrix Type 1-A. Success of sample was defined as ≥ 1 CTOS in the gel 2 weeks after cultivation.*

Table 9 shows the samples for the optimization phase, which sample conditions that were discarded, if they have been successfully split and reseeded, and the current status of the samples. Supplementary Figure S10A-C shows examples of CRC-12, CRC-15, and CRC-16, which were discarded due to no spheroids in the gel, abundance of bacteria in the sample, and gel disruption, respectively. An established sample of CRC-14 is shown in Supplementary Figure S10D to show the contrast between established and discarded samples.

The expansion procedure was more convenient with enzyme digestion compared to manual expansion with a needle since all samples that were defined established (except CRC-17 which was frozen down before expansion) were successfully split and reseeded. However, the yield of CTOSs was still quite low with enzyme digestion, i.e. there were more than 50 CTOSs per gel droplet before enzyme digestion and less than 10 CTOSs per gel droplet after enzyme digestion. Furthermore, the freezing and thawing procedure was tested for CRC-11, CRC-13 and CRC-14 which maintained high viability (Supplementary Figure S10E), and CTOSs could safely be frozen down.

Table 9. *Observations from each sample in the optimization phase. "Sample" is the patient ID; "Observations" is the observed results; and "Current status" is the current endpoint of the spheroid establishment. The Table shows which samples has been split ans reseeded, as well as frozen and thawed. "Established" indicates that there are ≥ 1 CTOS in the gel 2 weeks after cultivation; High yield indicates > 50 CTOSs in the gel; "D"* = Discarded, no spheroids found after seeding the gel or gel disrupted; "LC" = Liberase DH, Cultrex; "LK" = Liberase DH, Cellmatrix Type 1-A; CC = Collagenase II, Cultrex; CM = Collagenase II, Matrigel; and CK = Collagenase II, Cellmatrix Type 1-A; Number behind sample condition (e.g. CM-1) indicates if one or two technical replicates has been discarded.

Sample	Observations	Split and reseeded	Frozen and thawed	Current status
CRC-11	High yield of CTOSs. All samples established.	Yes	Yes	Frozen at -80 °C < 10 CTOSs
CRC-12	No CTOSs found after seeding to gel. All samples discarded.	No	No	Discarded
CRC-13	High yield of CTOSs. D: CM-1, CK-1, LK-1.	Yes	Yes	Frozen at -80 °C < 10 CTOSs
CRC-14	High yield of CTOSs. All samples established.	Yes	Yes	Frozen at -80 °C < 10 CTOSs
CRC-15	Infected, abundance of bacteria in medium. All samples discarded.	No	No	Discarded
CRC-16	High yield of CTOSs. D: CC-2, CK-1, LC-2 and LK-1.	Yes	No	Frozen at -80 °C < 10 CTOSs
CRC-17	High yield of CTOSs. All samples established.	No	No	Frozen at -80 °C > 50 CTOSs

5. Discussion

Approximately 20% of human cancers are caused by infectious agents[28]. The role of viruses in human cancers has been well recognized, especially small DNA viruses such as polyomavirus in Merkel cell carcinoma, papillomavirus in cervical cancer and EBV in Burkitt's lymphoma, Hodgkin's lymphoma and nasopharyngeal carcinoma[28]. In addition, accumulating research evidence indicates a link between bacterial infection and cancer, particularly the association between *H. pylori* and GC[29]. Infectious agents have been acknowledged and listed as risk factors in a number of cancers, but despite the extensive amount of research, the association between certain bacterial or viral infections and the risk of CRC is still not clear. The present study aimed to increase the understanding of *F. nucleatum* and EBV miRNA and their potential roles in the development of CRC. Both *F. nucleatum*[43,44] and EBV miRNA[86] have previously been found to have an increased expression in CRC compared to healthy tissue. Several studies have indicated numerous functional traits of the two infectious agents in the carcinogenesis of CRC, but neither are acknowledged as potential risk factors nor potential biomarkers for detection of CRC.

The overall aim of the present study was to understand the role of viruses and bacteria in CRC and determine if and how they contribute to tumour development. We elucidated the effects of *F. nucleatum* on cell migration, cell proliferation and upregulation of the cancer-associated genes *CCL20* and *CSF2*, which were chosen based on previous RNA sequencing results in the lab. In addition, we elucidated the effect of EBV miRNA, specifically miR-BART10-3p, which is highly expressed in CRC tissue compared to healthy adjacent tissue. Two genes, *MAT2B* and *ELL2*, with predicted miRNA target sites for miR-BART10-3p were further investigated. Lastly, we established a method for making patient-derived CRC spheroids. The method gives the opportunity for future studies of *F. nucleatum* and miR-BART10-3p in a molecular and phenotypic landscape of an original tumour, and to study the native microbiota of the tumour.

5.1. *F. nucleatum* in the carcinogenesis of colorectal cancer

F. nucleatum infection was confirmed by confocal fluorescence microscopy, which showed intracellular *F. nucleatum* aggregation 2 hours post-infection. An intracellular localization is considered advantageous for bacteria to evade acquired immunity, as well as antibiotic pressure, leading to intracellular persistence, proliferation and invasion of adjacent tissues[165]. *F. nucleatum* has been reported to upregulate several pathways in the host cell which is consistent with the idea that the invasive bacterium adapt to an intracellular lifestyle and persist for some

time within the cell[167]. There is also evidence that, *in vivo, F. nucleatum* continue to replicate inside host cells, and mouse xenografts of primary colorectal adenocarcinomas has been found to retain viable *F. nucleatum* through several successive passages despite their obligate anaerobic physiology[168].

It has been reported that binding and cellular invasion of *F. nucleatum* to CRC cell line HCT116 selectively induces the secretion of the proinflammatory and metastatic cytokines IL-8 and CXCL1 which induces migration of the colon cancer cell line HCT116[53]. We wanted to see if we could promote migration in another CRC cell line, in particular DLD-1, but instead of incubating cells with conditioned and concentrated media obtained from a *F. nucleatum* infection of a desired cell line, as they did for the HCT116 cell line, we incubated the cells directly with *F. nucleatum* for 24 hours. Figure 24 shows that there was not a significant difference in migration between *F. nucleatum*-treated and untreated DLD-1 ($P > 0.05$). A reason could be that *F. nucleatum* does not directly induce migration of CRC cells, but indirectly through the secretion of CXCL1 and IL-8. *F. nucleatum* may not selectively induce the secretion of CXCL1 and IL-8 in DLD-1, but this hypothesis needs further investigation. Nevertheless, another study reported that the TPH-1 monocyte migration was markedly enhanced by the presence of *F. nucleatum* whereas the migration of the CRC cell line HCP1 was not[114]. These observations warrant further investigation to better understand the molecular mechanisms that regulate the interactions between *F. nucleatum* and their target cells.

As described in the introduction, *F. nucleatum* activates a number of pathways that are associated with increased proliferation, including the Wnt/β-catenin signalling pathway and the RAS/MAPK signalling pathway. Unexpectedly, *F. nucleatum* did not induce cell proliferation in DLD-1 ($P > 0.05$). The results are, however, only based on four technical replicates, and no statistical significance can be concluded. Images presented in Figure 26 shows a greater number of cells in *F. nucleatum*-treated cells than *E. coli* DH5α-treated- and untreated cells at 48 hours post-infection. In addition, from Figure 27, a slight increase in proliferation is observed at 48 hours post-infection, indicating that *F. nucleatum* may have a slight effect on cell proliferation. Although not significant, in all but one technical replicate the proliferation was higher in *F. nucleatum*-treated- compared to non-treated cells at 48 hours (Table S4). The p-value was 0.07 at 48 hours which is a considerable trend toward significance. The automated imaging software ImageJ only measure cell-free area and does not consider that cells may grow on top of each other instead of in a monolayer. Many factors may play a role in the measurement, and more experiments are needed to verify it. An alternative method for measuring cell proliferation could

be to measure the overall metabolic activity inside the cells by using a colorimetric dye that produces a distinct colour when added to proliferating cells[169]. An example of one such dye is MTT, which is metabolized from a yellow dye to a purple formazan dye when cells are proliferating[170]. Another method for measuring cell proliferation could be to measure specific markers within a cell that correlate with high cell proliferation[169]. The marker bromodeoxyuridine (BrdU) are used in the BrdU incorporation assays and is incorporated into the DNA of proliferating cells. By adding enzyme-linked antibodies against BrdU, a colour develops in proportion to the level of BrdU incorporated into the cells[171]. In both methods, the absorbance of the sample is read with a microplate reader to measure the metabolic activity or the antibody-bound markers, thus measuring the cell proliferation rate[170,171]. These methods are less laborious and time-consuming than the manual proliferation assay used in the present study. In addition, manual recognition of whether a cell is a single cell, a cluster of cells, cell debris or other particles can be challenging even for the trained eye.

Cells of the innate immune system regulate immune responses through the production of antimicrobial peptides, chemokines, and cytokines, such as CSF2 and CCL20[113,119]. In normal colonic mucosa, both CCL20 and CSF2 are lowly expressed. However, in response to an inflammatory stimulus, the expression of the antimicrobial cytokines is significantly increased[113,119]. As shown in Figure 6 and 7, *CCL20* and *CSF2* mRNA are highly expressed in the lung and the urinary bladder, organs which are constantly exposed to the external environment. To prevent pathogens from entering the respiratory- and urinary tract, the immune system is tightly regulated through cell-cell communication, and thus depending on signalling mediated by cytokines[172,173]. However, several studies have shown that inflammatory immune cells are essential players of cancer-related inflammation, in which they infiltrate tumours and engage in an extensive and dynamic crosstalk with cancer cells, contributing to tumour development[174].

In the present study, the ability of DLD-1 to express *CCL20* and *CSF2* mRNA in response to *F. nucleatum* was demonstrated. There was found a slight increase in *CCL20* (Figure 30A) and *CSF2* (Figure 31) at 3 hours post-infection suggesting that gene expression occurs at an early event post-infection. In addition, *CCL20* and *CSF2* mRNA reached its maximum fold increase at 24 hours post-infection, indicating that *F. nucleatum* upregulate the genes in a time-dependent manner. However, as only one biological replicate was performed, we cannot tell the statistical significance. Nevertheless, *CCL20* has previously been found to be significantly upregulated in synovial fibroblasts and keratinocytes by IL-17A, a cytokine which has a

significant positive correlation with *F. nucleatum* infection[51], for 8 to 24 hours and 24 to 48 hours, respectively[175,176]. In addition, *CCL20* was significantly upregulated in keratinocytes by TNF-α, another cytokine which has a significant positive correlation with *F. nucleatum* infection[51], for 6 to 48 hours[176]. This indicates that *CCL20* mRNA expression is time-dependent and has various time-course patters in different cell types and with different stimuli. Furthermore, in one study, *CSF2* was significantly upregulated in endothelial cells by the TLR2/6 agonist MALP-2 (macrophage-activating lipopeptide of 2 kDa), already at 1 hour after stimulation, and expression remained elevated for 24 hours[177]. In another study, stimulation with the proinflammatory cytokine IL-1β rapidly induced *CSF2* mRNA levels in A549 human alveolar carcinoma cells, which peaked at 2 hours before declining at 6 hours[178]. These studies indicate that *CSF2* expression occurs early after stimulation, but does, however, vary in time-course pattern in different cell lines and with different stimuli. This corresponds to the PCA plots in Figure S1, showing that CRC cell lines have different mRNA expression profiles when treated with *F. nucleatum*, *E. coli* DH5α and TSB.

Because of various levels post-transcriptional and post-translational regulations, the amount of mRNA does not always correlate with the amount of protein produced. Nevertheless, the protein release of CCL20 by DLD-1 suggested that this is an early event regulating the immune system in the early hours after infection. Figure 30B shows that *F. nucleatum* elicits rapid release of CCL20 from DLD-1 already at 6 hours post-infection with a maximum release of CCL20 at 12 hours post-infection. The results are, however, not significant ($P > 0.05$) due to only two biological replicates with high standard deviation. However, these findings are consistent with a previous study indicating that CCL20 is released at early hours in response to *F. nucleatum* infection of human oral epithelial cells (HOECs)[179], and the same time-course pattern of CCL20 release was reported in TNF-α stimulated keratinocytes[176], indicating that CCL20 release is expected at an early hour.

Figure 32 shows that DLD-1 releases CCL20 in small concentrations already at 0.1 MOI *F. nucleatum* infection, but this is, however, approximately the same amount of CCL20 that is released when treated with non-pathogenic *E. coli* DH5α. In contrast, at 100 MOI *F. nucleatum*, we can observe a great increase in CCL20 from DLD-1 cells compared with 100 MOI *E. coli* DH5α treated cells and untreated cells. The release of CCL20 was significantly upregulated at 10 MOI and 100 MOI *F. nucleatum* compared to untreated DLD-1, indicating that *F. nucleatum* stimulates CCL20 release in a dose-dependent manner. In addition, these findings are consistent

with a previous study indicating that the release of CCL20 is significantly increasing from 0 to 100 MOI in TNF-α stimulated keratinocytes[176] which reinforces this study.

5.2. EBV miRNA in the carcinogenesis of colorectal cancer

EBV miRNAs, most of which are encoded from the BART loci, have been shown to target host mRNAs and genes involved in CRC development[86]. For instance, miR-BART10-3p, which previously has been associated with NPC[87] and GC[88], has been found to be significantly elevated in CRC tissue[86]. Two cancer-associated genes, *MAT2B* and *ELL2*, were found to be downregulated in miR-BART10-3p transfected SW620 cells after RNA sequencing (Table 7), and the sequencing results were further validated by RT-qPCR and luciferase assays.

From RT-qPCR, we found that *MAT2B* and *ELL2* were significantly downregulated in miR-BART10-3p transfected SW620 cells with 65% and 50%, respectively. Furthermore, by doing luciferase assays, we found that miR-BART10-3p reduced expression of *MAT2B* by directly targeting the 3'UTR region of this gene. For the LightSwitch Luciferase Assay, we found that *MAT2B* was significantly downregulated by 60%, and *ELL2* was downregulated by 15%. The downregulation of *ELL2* was observed but was not significant. For the Dual Luciferase Assay, which we only did for *MAT2B* due to problems with the ligation of *ELL2* into the psiCHECK™-2 vector, we found that *MAT2B* was downregulated by approximately 30%. Interestingly, the Dual Luciferase Assay showed less downregulation for *MAT2B* than the LightSwitch Luciferase Assay. The Dual Luciferase construct only contain the target site and some flanking nucleotides on either side and therefore lack the sequence context around the target site. Studies have shown that the position of the miRNA target site can affect the efficiency of the site, and some influencing factors include target site accessibility, AU-richness, position relative to the start and end of the 3'UTR, and position relative to other miRNA target sites[180]. A construct with the whole 3'UTR will be influenced with many of these features, including the effects from other endogenous miRNAs, compared to the Dual Luciferase Assay that only contain the target site.

Downregulation of *MAT2B* has previously been shown to inhibit migration and activate apoptosis by inhibiting the AKT and ERK pathways in several cancers, including colon cancer[126-128], raising the possibility that the gene may be a potential oncogene. In hepatocellular carcinoma, recent studies have detected an increase in *MAT2B* levels[181,182]. In hepatocytes, the methyl donor SAM is the key regulator of proliferation, death and differentiation, and upregulation of *MAT2B* results in decreased SAM levels and increased hepatoma cell growth[182].

The berberine-induced miR-21-3p has been found to directly target the 3'UTR of *MAT2B* resulting in increased intracellular SAM levels and induced growth inhibition and apoptosis[182]. SAM has also been shown to suppress cell growth in the colon by reversing DNA hypomethylation of the oncogenic c-Myc and h-Ras, thus decreasing their expression[183]. This finding suggests that miR-BART10-3p may act as a tumour suppressor miRNA in CRC by targeting the 3'UTR of *MAT2B*. Apart from that, we discovered that *MAT2B* had a highly conserved target site among vertebrates using the TargetScan software (Table 6). The 8mer site is the most effective canonical site type and are expected to be the most responsive to miRNA. This is followed by the 7mer-m8 site, the 7mer-1a site and the 6mer site which were found in *ELL2*[164]. The 8mer site of *MAT2B* have been selectively maintained and similar pathways and biological processes may have been regulated through the same conserved miRNAs[184]. As presented in Figure 8, *MAT2B* is highly expressed in immune cells. *CCDC50* (coiled-coil domain-containing protein 50) has been found to be overexpressed in the B cell lymphoma cell line GRANTA-519, and knockdown of *CCDC50* resulted in *MAT2B* downregulation by approximately 40%, as well as 75% less proliferation[185]. These results suggest that there might be a link between *MAT2B* and increased cell proliferation in immune cell lymphomas. Similarly, EBV has been associated with a broad variety of lymphoproliferative lesions and malignant lymphomas of B-, T- and NK cells[186]. As both *MAT2B* and EBV are highly expressed in the immune cells, miR-BART10-3p may have been evolved to target *MAT2B* in the immune cells. In fact, there might me a possibility that EBV infect immune cells such as tumour-infiltrating T cells and macrophages and subsequently contribute to the immune response against CRC. However, further investigation is needed to understand the *in vivo* effects of miR-BART10-3p on *MAT2B*.

As described in the introduction, *ELL2* is found to be a susceptibility gene in MM by driving Ig secretory-specific mRNA production. It has been reported that *ELL2* is a target for the miR-30 family in the coordinating of plasma cell differentiation, and inhibition of the miR-30 family, which is aberrantly overexpressed in MM, enhances the differentiation of B cells to plasma cells[187]. These results indicate that *ELL2* may function as an oncogene in MM. On the other hand, downregulation of *ELL2* has been shown to increase proliferation, migration, and invasion in multiple prostate epithelial cells[135]. A study shows that *ELL2* knockdown enhances the expression of the anti-apoptotic factor Birc3 in prostate cells, suggesting that *ELL2* may play a role in driving prostate cancer proliferation[188]. In addition, *ELL2* and the prostate tumor suppressor ELL-associated factor 2 (*EAF2*) can functionally interact, and combined knockdown

of these two genes has been found to increase activation of the AKT pathway[189]. These findings are comprising the possibility that the gene may be a potential tumour suppressor in the prostate epithelial cells, but whether *ELL2* exhibit tumour suppressive properties in colorectal tissue is yet unknown.

5.3. Patient-derived colorectal cancer spheroids

In the initial phase, the success rate of established spheroid samples was 20%. In contrast, the success rate of established spheroid samples in the optimization phase was > 50%. The success rate had increased from 20% to 57.1% for tissue samples treated with Liberase DH and Cellmatrix Type 1-A, indicating that the magnetic stirrer, which was introduced in the optimization phase, applies a more forceful agitation than a stirring water bath, thus increasing the yield. Nevertheless, some tumour samples had a harder consistency than others, and the magnetic stirrer seemed to be more mechanically aggressive towards some samples than others. As observed in Figure 36C, CRC-17 was digested into small fragments, some of which had lost cell-cell contact, indicating that the tissue had been over-digested. Therefore, digestion time and rotation speed on the magnetic stirrer should be determined and monitored for each and every tumour sample. Interestingly, fractions of non-dissociated CTOS cells accumulated in the gel and formed spheroids after several days. Furthermore, cells located in the core of large CTOSs underwent apoptosis or necrosis as a result of less transport of oxygen and nutrients to the innermost cells, thus giving the spheroids a dark core[190].

As observed in Figure 36A, the CTOSs showed high shape and size variability, and there could be observed budding of one or more spheroids. A recent study found that budding of spheroids and cell detachment are more frequently detected in spheroids with non-spherical shape[190], which is important information when selecting the most appropriate spheroids for future studies. There has also been reported that different morphological differences such as size and sphericity affect reproducibility of the results[191]. Several methods, such as microfluidic devices, has been developed to reduce the heterogeneity of spheroid sizes[191]. However, these methods use cancer stem cells as the starting material, and not tumour fractions. Collection of equally sized and shaped CTOSs usually requires handpicking under the microscope or sieving through nylon meshes[192].

When comparing the efficiency of Liberase DH and Collagenase II in the present study, the CTOS yield was observed to be higher with Collagenase II, whereas Liberase DH was observed

to be gentler with the tissue sample resulting in less cell dissociation. In contrast to Collagenase II, Liberase DH is high in Dispase, which is a gentle enzyme for the separation of tissue that does not damage cell membranes[161]. Nevertheless, Collagenase is said to be a relatively gentle enzyme that does not require mechanical agitation, suggesting that a lower rotation speed could lead to less cell dissociation for this digestion enzyme[161]. Moreover, when factoring in required volume for each digestion run, Liberase DH costs six times as much as Collagenase II[193,194], making Collagenase II a much more favourable option.

As for the gels, Cellmatrix Type 1-A and Cultrex are both hard to work with; Cellmatrix Type 1-A is very sensitive to high temperatures and sometimes solidify before the addition to the plate; and Cultrex forms very thin gel droplets, resulting in CTOSs growing in a 2D monolayer. In addition to being laborious to work with, Cellmatrix Type 1-A and Cultrex are the gels with the lowest success rates for both digestion enzymes (Figure 37), indicating that the low yield is associated with the problems of seeding the gel. In contrast, Matrigel and Geltrex are easy to work with, and are also the gels with the highest success rates for both digestion enzymes (Figure 37). Both gels are soluble forms of secreted extracellular matrix proteins, purified from murine Engelbreth-Holm-Swarm tumor cells[195,196]. They are very similar, but Matrigel is, however, the most commonly used gel and has been around for longer; Matrigel has more than 12,000 numbers of citations in journal publications dating back to 1987, whereas Geltrex has 38 citations as of May 2020 dating back to 2010[197]. The costs are similar for both gels when factoring in sizes and required volume for each gel droplet, but Matrigel has to be bought in a larger volume than Geltrex[195,196].

All in all, we have developed a procedure that successfully generate viable spheroids. Due to limited access to the laboratory, future improvements are needed to reduce cell detachment, variable shapes and sizes in the gel, and to figure out how to improve the yield of spheroids after expansion. Furthermore, Collagenase II is the most favourable digestion enzyme due to the high yield and low cost but needs to be tested with lower rotation speed to determine whether it leads to gentler digestion. Subsequently, decisions must be made about which gel, Matrigel or Geltrex, should be excluded from the protocol.

The protocol established in the present study aimed to be used in future studies to look at the effect of *F. nucleatum* and miR-BART10-3p in a molecular and phenotypic landscape of an original tumour, and to study the native microbiota of the spheroids. A recent study demonstrated that tumour spheroids supported co-culture with *F. nucleatum* and consistently

remained viable for at least 48 hours, which most likely is due to the hypoxic core that is characteristic of large spheroids[198]. Another study identified 33 differentially expressed miRNA in cells growing in breast multicellular tumour spheroids compared to cells growing in a monolayer, of which the top ten dysregulated miRNAs were associated with breast cancer[199]. These discoveries provide new insight of how anaerobic bacteria, such as *F. nucleatum*, function in the tumor microenvironment, and how the ECM and cell-cell contact influences the biology of miRNA in CRC cells[200].

6. Conclusion

In the present study, the direct effect of *F. nucleatum* in CRC has been validated by co-culturing the bacteria with DLD-1 and look for changes in proliferation, migration and gene expression. Surprisingly, *F. nucleatum* was not found to promote cell migration or cell proliferation in DLD-1. However, by RT-qPCR and ELISA, *F. nucleatum* was observed to induce mRNA expression of *CCL20* and *CSF2*, which are important cytokines in the regulation of inflammation. In addition, *F. nucleatum* significantly induced CCL20 protein release in a dose-dependent manner. Furthermore, an intracellular localization of *F. nucleatum* was confirmed upon *F. nucleatum* co-culturing with CRC cell lines using confocal microscopy.

RNA seq of miR-BART10-3p-treated SW620 led to the identification of the cancer-related genes *MAT2B* and *ELL2*. *In vitro* RT-qPCR validation of *MAT2B* and *ELL2* confirmed that miR-BART10-30 significantly downregulates these genes, and the luciferase assays confirmed that *MAT2B* is a direct target for this miRNA. These results suggest that miR-BART10-3p may function as a tumor suppressor miRNA by directly downregulating *MAT2B*, a gene which has been shown to activate the AKT and ERK pathways in CRC, as well as an oncomiR by potentially downregulating *ELL2*, a gene with tumor suppressor functions in prostate epithelial cells. However, further investigations of the target genes are needed to understand the effect of miR-BART10-3p in the colorectal carcinogenesis.

Lastly, an *in vitro* procedure for establishing patient-derived CRC spheroids was developed for future studies of *F. nucleatum* and miR-BART10-3p in a molecular and phenotypic landscape of an original tumour, but further decisions are needed in order to improve the cellular viability and the cellular growth of the spheroids, and to reduce the heterogeneity of the spheroid sizes.

7. References

1. Rawla, P., Sunkara, T. & Barsouk, A. Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors. *Prz Gastroenterol* **14**, 89-103 (2019).
2. Cancer Registry of Norway. Cancer in Norway 2018. (Oslo: Cancer Registry of Norway, 2018).
3. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians* **68**, 394-424 (2018).
4. Chan, D.S. *et al.* Red and processed meat and colorectal cancer incidence: meta-analysis of prospective studies. *PLoS One* **6**, e20456 (2011).
5. Johnson, C.M. *et al.* Meta-analyses of colorectal cancer risk factors. *Cancer Causes Control* **24**, 1207-22 (2013).
6. Kuipers, E.J. *et al.* Colorectal cancer. *Nat Rev Dis Primers* **1**, 15065 (2015).
7. Braendegaard Winther, S., Baatrup, G., Pfeiffer, P. & Qvortrup, C. Trends in colorectal cancer in the elderly in Denmark, 1980-2012. *Acta Oncol* **55 Suppl 1**, 29-39 (2016).
8. Aran, V., Victorino, A.P., Thuler, L.C. & Ferreira, C.G. Colorectal Cancer: Epidemiology, Disease Mechanisms and Interventions to Reduce Onset and Mortality. *Clin Colorectal Cancer* **15**, 195-203 (2016).
9. The American Cancer Society medical and editorial content team. Colorectal Cancer Signs and Symptoms. (American Cancer Society, 2018).
10. The American Cancer Society medical and editorial content team. Colorectal Cancer Stages. (American Cancer Society, 2018).
11. ASCO: American Society of Clinical Oncology. Colorectal Cancer: Statistics (Cancer.net, 2019).
12. The American Cancer Society medical and editorial content team. Survival Rates for Colorectal Cancer. (American Cancer Society, 2020).
13. The American Cancer Society medical and editorial content team. Treatment of Colon Cancer, by Stage. (American Cancer Society, 2018).
14. McQuade, R.M., Stojanovska, V., Bornstein, J.C. & Nurgali, K. Colorectal Cancer Chemotherapy: The Evolution of Treatment and New Approaches. *Curr Med Chem* **24**, 1537-1557 (2017).
15. Baudino, T.A. Targeted Cancer Therapy: The Next Generation of Cancer Treatment. *Curr Drug Discov Technol* **12**, 3-20 (2015).
16. Mehta, S.R., Suhag, V., Semwal, M. & Sharma, N. Radiotherapy: Basic Concepts and Recent Advances. *Med J Armed Forces India* **66**, 158-62 (2010).
17. Worthley, D.L. & Leggett, B.A. Colorectal cancer: molecular features and clinical opportunities. *Clin Biochem Rev* **31**, 31-8 (2010).
18. Nguyen, H.T. & Duong, H.Q. The molecular characteristics of colorectal cancer: Implications for diagnosis and therapy. *Oncol Lett* **16**, 9-18 (2018).
19. Smith, L., Anti, S., White, N. & Saldanha, S. *Epigenetic Mechanisms in Cancer. Chapter 4 - The Genetics and Epigenetics of Colorectal Cancer Health Disparity,*, (Academic Press, 2018).
20. Santarpia, L., Lippman, S.M. & El-Naggar, A.K. Targeting the MAPK-RAS-RAF signaling pathway in cancer therapy. *Expert Opin Ther Targets* **16**, 103-19 (2012).
21. Khanal, N., Upadhyay, S. & Silberstein, P.T. Colorectal Carcinoma and Emerging Targeted Therapies. *Fed Pract* **32**, 27s-31s (2015).
22. Fearon, E.R. & Vogelstein, B. A genetic model for colorectal tumorigenesis. *Cell* **61**, 759-67 (1990).
23. Kohno, T. *et al.* Mutation and expression of the DCC gene in human lung cancer. *Neoplasia* **2**, 300-5 (2000).
24. Zilfou, J.T. & Lowe, S.W. Tumor suppressive functions of p53. *Cold Spring Harb Perspect Biol* **1**, a001883 (2009).

25. Vogelstein, B. *et al.* Genetic alterations during colorectal-tumor development. *N Engl J Med* **319**, 525-32 (1988).
26. Sameer, A.S. Colorectal cancer: molecular mutations and polymorphisms. *Front Oncol* **3**, 114 (2013).
27. Martinez, J.D., Parker, M.T., Fultz, K.E., Ignatenko, N.A. & Gerner, E.W. Molecular Biology of Cancer. in *Burger's Medicinal Chemistry and Drug Discovery* 1-50 (2003).
28. Howley, P.M. Gordon Wilson Lecture: Infectious Disease Causes of Cancer: Opportunities for Prevention and Treatment. *Trans Am Clin Climatol Assoc* **126**, 117-32 (2015).
29. National Cancer Institute. Infectious Agents. (National Cancer Institute, 2019).
30. Yim, E.K. & Park, J.S. The role of HPV E6 and E7 oncoproteins in HPV-associated cervical carcinogenesis. *Cancer Res Treat* **37**, 319-24 (2005).
31. Yassour, M. *et al.* Strain-Level Analysis of Mother-to-Child Bacterial Transmission during the First Few Months of Life. *Cell Host Microbe* **24**, 146-154.e4 (2018).
32. Nigro, G., Hanson, M., Fevre, C., Lecuit, M. & Sansonetti, P.J. Intestinal Organoids as a Novel Tool to Study Microbes-Epithelium Interactions. *Methods Mol Biol* **1576**, 183-194 (2019).
33. Antonic, V. *et al.* Significance of infectious agents in colorectal cancer development. *J Cancer* **4**, 227-40 (2013).
34. Rinninella, E. *et al.* What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases. *Microorganisms* **7**(2019).
35. Kelly, D., Yang, L. & Pei, Z. Gut Microbiota, Fusobacteria, and Colorectal Cancer. *Diseases* **6**(2018).
36. Sears, C.L. & Garrett, W.S. Microbes, microbiota, and colon cancer. *Cell Host Microbe* **15**, 317-28 (2014).
37. Garrett, W.S. Cancer and the microbiota. *Science* **348**, 80-6 (2015).
38. Shang, F.M. & Liu, H.L. Fusobacterium nucleatum and colorectal cancer: A review. *World J Gastrointest Oncol* **10**, 71-81 (2018).
39. Wu, S., Morin, P.J., Maouyo, D. & Sears, C.L. Bacteroides fragilis enterotoxin induces c-Myc expression and cellular proliferation. *Gastroenterology* **124**, 392-400 (2003).
40. Abdulamir, A.S., Hafidh, R.R. & Abu Bakar, F. The association of Streptococcus bovis/gallolyticus with colorectal tumors: the nature and the underlying mechanisms of its etiological role. *J Exp Clin Cancer Res* **30**, 11 (2011).
41. Grahn, N., Hmani-Aifa, M., Fransen, K., Soderkvist, P. & Monstein, H.J. Molecular identification of Helicobacter DNA present in human colorectal adenocarcinomas by 16S rDNA PCR amplification and pyrosequencing analysis. *J Med Microbiol* **54**, 1031-5 (2005).
42. Balamurugan, R., Rajendiran, E., George, S., Samuel, G.V. & Ramakrishna, B.S. Real-time polymerase chain reaction quantification of specific butyrate-producing bacteria, Desulfovibrio and Enterococcus faecalis in the feces of patients with colorectal cancer. *J Gastroenterol Hepatol* **23**, 1298-303 (2008).
43. Castellarin, M. *et al.* Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. *Genome Res* **22**, 299-306 (2012).
44. Ito, M. *et al.* Association of Fusobacterium nucleatum with clinical and molecular features in colorectal serrated pathway. *Int J Cancer* **137**, 1258-68 (2015).
45. Tahara, T. *et al.* Fusobacterium in colonic flora and molecular features of colorectal carcinoma. *Cancer Res* **74**, 1311-8 (2014).
46. Abed, J. *et al.* Fap2 Mediates Fusobacterium nucleatum Colorectal Adenocarcinoma Enrichment by Binding to Tumor-Expressed Gal-GalNAc. *Cell Host Microbe* **20**, 215-25 (2016).
47. Rubinstein, M.R. *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).

48. Gur, C. *et al.* Binding of the Fap2 protein of Fusobacterium nucleatum to human inhibitory receptor TIGIT protects tumors from immune cell attack. *Immunity* **42**, 344-355 (2015).
49. Yu, T. *et al.* Fusobacterium nucleatum Promotes Chemoresistance to Colorectal Cancer by Modulating Autophagy. *Cell* **170**, 548-563.e16 (2017).
50. Yang, Z. & Ji, G. Fusobacterium nucleatum-positive colorectal cancer. *Oncol Lett* **18**, 975-982 (2019).
51. McCoy, A.N. *et al.* Fusobacterium is associated with colorectal adenomas. *PLoS One* **8**, e53653 (2013).
52. Xia, Y., Shen, S. & Verma, I.M. NF-kappaB, an active player in human cancers. *Cancer Immunol Res* **2**, 823-30 (2014).
53. Casasanta, M.A. *et al.* Fusobacterium nucleatum host cell binding and invasion induces IL-8 and CXCL1 secretion that drives colorectal cancer cell migration. *bioRxiv*, 2020.01.15.907931 (2020).
54. Brennan, C.A. & Garrett, W.S. Fusobacterium nucleatum — symbiont, opportunist and oncobacterium. *Nature Reviews Microbiology* **17**, 156-166 (2019).
55. Kim, D. & Rossi, J. RNAi mechanisms and applications. *Biotechniques* **44**, 613-6 (2008).
56. Chery, J. RNA therapeutics: RNAi and antisense mechanisms and clinical applications. *Postdoc J* **4**, 35-50 (2016).
57. Poulsen, C., Vaucheret, H. & Brodersen, P. Lessons on RNA silencing mechanisms in plants from eukaryotic argonaute structures. *Plant Cell* **25**, 22-37 (2013).
58. Kim, V.N. Small RNAs just got bigger: Piwi-interacting RNAs (piRNAs) in mammalian testes. *Genes Dev* **20**, 1993-7 (2006).
59. Carthew, R.W. & Sontheimer, E.J. Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136**, 642-55 (2009).
60. Senti, K.A. & Brennecke, J. The piRNA pathway: a fly's perspective on the guardian of the genome. *Trends Genet* **26**, 499-509 (2010).
61. Winter, J., Jung, S., Keller, S., Gregory, R.I. & Diederichs, S. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat Cell Biol* **11**, 228-34 (2009).
62. Wilson, R.C. & Doudna, J.A. Molecular mechanisms of RNA interference. *Annu Rev Biophys* **42**, 217-39 (2013).
63. Dueck, A., Ziegler, C., Eichner, A., Berezikov, E. & Meister, G. microRNAs associated with the different human Argonaute proteins. *Nucleic Acids Res* **40**, 9850-62 (2012).
64. Macfarlane, L.A. & Murphy, P.R. MicroRNA: Biogenesis, Function and Role in Cancer. *Curr Genomics* **11**, 537-61 (2010).
65. Shivdasani, R.A. MicroRNAs: regulators of gene expression and cell differentiation. *Blood* **108**, 3646-53 (2006).
66. Zhang, B., Pan, X., Cobb, G.P. & Anderson, T.A. microRNAs as oncogenes and tumor suppressors. *Dev Biol* **302**, 1-12 (2007).
67. Shenouda, S.K. & Alahari, S.K. MicroRNA function in cancer: oncogene or a tumor suppressor? *Cancer Metastasis Rev* **28**, 369-78 (2009).
68. Chen, B. *et al.* Emerging microRNA biomarkers for colorectal cancer diagnosis and prognosis. *Open Biol* **9**, 180212 (2019).
69. Gatalica, Z., Vranic, S., Xiu, J., Swensen, J. & Reddy, S. High microsatellite instability (MSI-H) colorectal carcinoma: a brief review of predictive biomarkers in the era of personalized medicine. *Fam Cancer* **15**, 405-12 (2016).
70. Nosho, K. *et al.* Association of microRNA-31 with BRAF mutation, colorectal cancer survival and serrated pathway. *Carcinogenesis* **35**, 776-83 (2014).
71. Rohr, C. *et al.* High-throughput miRNA and mRNA sequencing of paired colorectal normal, tumor and metastasis tissues and bioinformatic modeling of miRNA-1 therapeutic applications. *PLoS One* **8**, e67461 (2013).

72. Nagel, R. *et al.* Regulation of the adenomatous polyposis coli gene by the miR-135 family in colorectal cancer. *Cancer Res* **68**, 5795-802 (2008).
73. Drusco, A. *et al.* MicroRNA profiles discriminate among colon cancer metastasis. *PLoS One* **9**, e96670 (2014).
74. Schiller, J.T. & Lowy, D.R. Vaccines to prevent infections by oncoviruses. *Annu Rev Microbiol* **64**, 23-41 (2010).
75. Iizasa, H., Nanbo, A., Nishikawa, J., Jinushi, M. & Yoshiyama, H. Epstein-Barr Virus (EBV)-associated gastric carcinoma. *Viruses* **4**, 3420-39 (2012).
76. Baumforth, K.R., Young, L.S., Flavell, K.J., Constandinou, C. & Murray, P.G. The Epstein-Barr virus and its association with human cancers. *Mol Pathol* **52**, 307-22 (1999).
77. Fields, B.N., Knipe, D.M. & Howley, P.M. *Fields virology*, (Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia, 2007).
78. Mui, U.N., Haley, C.T. & Tyring, S.K. Viral Oncology: Molecular Biology and Pathogenesis. *J Clin Med* **6**(2017).
79. Alfieri, C. *et al.* Epstein-Barr virus transmission from a blood donor to an organ transplant recipient with recovery of the same virus strain from the recipient's blood and oropharynx. *Blood* **87**, 812-7 (1996).
80. Stanfield, B.A. & Luftig, M.A. Recent advances in understanding Epstein-Barr virus. *F1000Res* **6**, 386 (2017).
81. Young, L.S. & Rickinson, A.B. Epstein-Barr virus: 40 years on. *Nat Rev Cancer* **4**, 757-68 (2004).
82. Hutt-Fletcher, L.M. The Long and Complicated Relationship between Epstein-Barr Virus and Epithelial Cells. *J Virol* **91**(2017).
83. Wang, M. *et al.* Epstein-Barr virus-encoded microRNAs as regulators in host immune responses. *Int J Biol Sci* **14**, 565-576 (2018).
84. Zhang, J. *et al.* The oncogenic role of Epstein-Barr virus-encoded microRNAs in Epstein-Barr virus-associated gastric carcinoma. *J Cell Mol Med* **22**, 38-45 (2018).
85. Tsao, S.W., Tsang, C.M., To, K.F. & Lo, K.W. The role of Epstein-Barr virus in epithelial malignancies. *J Pathol* **235**, 323-33 (2015).
86. Mjelle, R., Sjursen, W., Thommesen, L., Saetrom, P. & Hofsli, E. Small RNA expression from viruses, bacteria and human miRNAs in colon cancer tissue and its association with microsatellite instability and tumor location. *BMC Cancer* **19**, 161 (2019).
87. Yan, Q. *et al.* EBV-miR-BART10-3p facilitates epithelial-mesenchymal transition and promotes metastasis of nasopharyngeal carcinoma by targeting BTRC. *Oncotarget* **6**, 41766-82 (2015).
88. Min, K. & Lee, S.K. EBV miR-BART10-3p Promotes Cell Proliferation and Migration by Targeting DKK1. *Int J Biol Sci* **15**, 657-667 (2019).
89. Li, Y. & Kumacheva, E. Hydrogel microenvironments for cancer spheroid growth and drug screening. *Sci Adv* **4**, eaas8998 (2018).
90. Hofmann, C. *et al.* Cell-cell contacts prevent anoikis in primary human colonic epithelial cells. *Gastroenterology* **132**, 587-600 (2007).
91. Chiarugi, P. & Giannoni, E. Anoikis: a necessary death program for anchorage-dependent cells. *Biochem Pharmacol* **76**, 1352-64 (2008).
92. Kondo, J. *et al.* Retaining cell-cell contact enables preparation and culture of spheroids composed of pure primary cancer cells from colorectal cancer. *Proc Natl Acad Sci U S A* **108**, 6235-40 (2011).
93. Nath, S. & Devi, G.R. Three-dimensional culture systems in cancer research: Focus on tumor spheroid model. *Pharmacol Ther* **163**, 94-108 (2016).
94. Wang, C. *et al.* Three-dimensional in vitro cancer models: a short review. *Biofabrication* **6**, 022001 (2014).

95. Yip, D. & Cho, C.H. A multicellular 3D heterospheroid model of liver tumor and stromal cells in collagen gel for anti-cancer drug testing. *Biochem Biophys Res Commun* **433**, 327-32 (2013).
96. Ibidi. 3D Cell culture. (ed. 3D cell culture applications in comparison to the traditional 2D monolayer) (n.d.).
97. Gionet-Gonzales, M.A. & Leach, J.K. Engineering principles for guiding spheroid function in the regeneration of bone, cartilage, and skin. *Biomed Mater* **13**, 034109 (2018).
98. Xu, H. *et al.* Organoid technology and applications in cancer research. *J Hematol Oncol* **11**, 116 (2018).
99. Ooft, S.N. *et al.* Patient-derived organoids can predict response to chemotherapy in metastatic colorectal cancer patients. *Sci Transl Med* **11**(2019).
100. van de Wetering, M. *et al.* Prospective derivation of a living organoid biobank of colorectal cancer patients. *Cell* **161**, 933-45 (2015).
101. Sachs, N. *et al.* A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity. *Cell* **172**, 373-386.e10 (2018).
102. Boj, S.F. *et al.* Organoid models of human and mouse ductal pancreatic cancer. *Cell* **160**, 324-38 (2015).
103. Broutier, L. *et al.* Human primary liver cancer-derived organoid cultures for disease modeling and drug screening. *Nat Med* **23**, 1424-1435 (2017).
104. Gao, D. *et al.* Organoid cultures derived from patients with advanced prostate cancer. *Cell* **159**, 176-187 (2014).
105. Jee, J.H. *et al.* Development of Collagen-Based 3D Matrix for Gastrointestinal Tract-Derived Organoid Culture. *Stem Cells Int* **2019**, 8472712 (2019).
106. Almeqdadi, M., Mana, M.D., Roper, J. & Yilmaz, O.H. Gut organoids: mini-tissues in culture to study intestinal physiology and disease. *Am J Physiol Cell Physiol* **317**, C405-c419 (2019).
107. Williamson, I.A. *et al.* A High-Throughput Organoid Microinjection Platform to Study Gastrointestinal Microbiota and Luminal Physiology. *Cell Mol Gastroenterol Hepatol* **6**, 301-319 (2018).
108. Neerincx, M. *et al.* MiR expression profiles of paired primary colorectal cancer and metastases by next-generation sequencing. *Oncogenesis* **4**, e170 (2015).
109. Sun, G. *et al.* Signature miRNAs in colorectal cancers were revealed using a bias reduction small RNA deep sequencing protocol. *Oncotarget* **7**, 3857-72 (2016).
110. Rossi, D.L., Vicari, A.P., Franz-Bacon, K., McClanahan, T.K. & Zlotnik, A. Identification through bioinformatics of two new macrophage proinflammatory human chemokines: MIP-3alpha and MIP-3beta. *J Immunol* **158**, 1033-6 (1997).
111. Hieshima, K. *et al.* Molecular cloning of a novel human CC chemokine liver and activation-regulated chemokine (LARC) expressed in liver. Chemotactic activity for lymphocytes and gene localization on chromosome 2. *J Biol Chem* **272**, 5846-53 (1997).
112. Hromas, R. *et al.* Cloning and characterization of exodus, a novel beta-chemokine. *Blood* **89**, 3315-22 (1997).
113. Frick, V.O., Rubie, C., Keilholz, U. & Ghadjar, P. Chemokine/chemokine receptor pair CCL20/CCR6 in human colorectal malignancy: An overview. *World J Gastroenterol* **22**, 833-41 (2016).
114. Ye, X. *et al.* Fusobacterium Nucleatum Subspecies Animalis Influences Proinflammatory Cytokine Expression and Monocyte Activation in Human Colorectal Tumors. *Cancer Prev Res (Phila)* **10**, 398-409 (2017).
115. Liu, J. *et al.* CCR6 is a prognostic marker for overall survival in patients with colorectal cancer, and its overexpression enhances metastasis in vivo. *PLoS One* **9**, e101137 (2014).
116. Brand, S. *et al.* Cell differentiation dependent expressed CCR6 mediates ERK-1/2, SAPK/JNK, and Akt signaling resulting in proliferation and migration of colorectal cancer cells. *J Cell Biochem* **97**, 709-23 (2006).
117. The human protein atlas. CCL20, RNA expression overview, Consensus dataset. (n.d.).

118. Xu, Z. *et al.* Demethylation and Overexpression of CSF2 are Involved in Immune Response, Chemotherapy Resistance, and Poor Prognosis in Colorectal Cancer. *Onco Targets Ther* **12**, 11255-11269 (2019).
119. Hong, I.S. Stimulatory versus suppressive effects of GM-CSF on tumor progression in multiple cancer types. *Exp Mol Med* **48**, e242 (2016).
120. Urdinguio, R.G. *et al.* Immune-dependent and independent antitumor activity of GM-CSF aberrantly expressed by mouse and human colorectal tumors. *Cancer Res* **73**, 395-405 (2013).
121. Gutschalk, C.M. *et al.* GM-CSF enhances tumor invasion by elevated MMP-2, -9, and -26 expression. *Cancer Med* **2**, 117-29 (2013).
122. The human protein atlas. CSF2, RNA expression overview, Consensus dataset. (n.d.).
123. Yuan, Y. *et al.* MAT2B promotes proliferation and inhibits apoptosis in osteosarcoma by targeting epidermal growth factor receptor and proliferating cell nuclear antigen. *Int J Oncol* **54**, 2019-2029 (2019).
124. Lu, S.C. & Mato, J.M. S-adenosylmethionine in liver health, injury, and cancer. *Physiol Rev* **92**, 1515-42 (2012).
125. Yoshiyama, H., Imai, S., Shimizu, N. & Takada, K. Epstein-Barr virus infection of human gastric carcinoma cells: implication of the existence of a new virus receptor different from CD21. *J Virol* **71**, 5688-91 (1997).
126. Peng, H. *et al.* MAT2B-GIT1 interplay activates MEK1/ERK 1 and 2 to induce growth in human liver and colon cancer. *Hepatology* **57**, 2299-313 (2013).
127. Lei, Y. *et al.* Lentivirus-mediated downregulation of MAT2B inhibits cell proliferation and induces apoptosis in melanoma. *Int J Oncol* **49**, 981-90 (2016).
128. Xu, J., Wu, D., Wang, S. & Wang, Z. MAT2B expression correlates with poor prognosis in triple-negative breast cancer. *Cancer Manag Res* **11**, 5501-5511 (2019).
129. The human protein atlas. MAT2B, RNA expression overview, Consensus dataset. (n.d.).
130. Pascal, L.E. *et al.* Conditional deletion of ELL2 induces murine prostate intraepithelial neoplasia. *J Endocrinol* **235**, 123-136 (2017).
131. Luo, Z., Lin, C. & Shilatifard, A. The super elongation complex (SEC) family in transcriptional control. *Nat Rev Mol Cell Biol* **13**, 543-7 (2012).
132. Luo, Z. *et al.* The super elongation complex family of RNA polymerase II elongation factors: gene target specificity and transcriptional output. *Mol Cell Biol* **32**, 2608-17 (2012).
133. The human protein atlas. ELL2, RNA expression overview, Consensus dataset. (n.d.).
134. Ali, M. *et al.* The multiple myeloma risk allele at 5q15 lowers ELL2 expression and increases ribosomal gene expression. *Nat Commun* **9**, 1649 (2018).
135. Yang, T. *et al.* Regulation of ELL2 stability and polyubiquitination by EAF2 in prostate cancer cells. *Prostate* **78**, 1201-1212 (2018).
136. Sigma-Aldrich. Cell Viability and Proliferation Assays. (n.d.).
137. Ibidi. The Principle of a Wound Healing Assay and Migration Assay. (n.d.).
138. Ibidi. µ-Slide 8 Well Grid-500. (Ibidi, n.d.).
139. Ibidi. Culture-Insert 2 Well in µ-Dish 35 mm. (n.d.).
140. Ibidi. Creating the Gap: Different Approaches. (ed. Creating the Gap Using a Culture-Insert) (Ibidi, n.d.).
141. Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671-5 (2012).
142. Ibidi. Microscopy Techniques and Culture Surfaces: Find the Perfect Match. (n.d.).
143. Combs, C.A. Fluorescence microscopy: a concise guide to current imaging methods. *Curr Protoc Neurosci* **Chapter 2**, Unit2.1 (2010).
144. Ibidi. Microscopy Techniques and Culture Surfaces: Find the Perfect Match. (ed. Excitation and emission light pathways in a basic confocal microscope configuration) (n.d.).
145. Thermo Fisher Scientific. Basic Principles of RT-qPCR. (n.d.).

146. New England BioLabs. qPCR & RT-qPCR. (n.d.).
147. Livak, K.J. & Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-8 (2001).
148. R&D Systems. Quantikine®ELISA Human CCL20/MIP-3α Immunoassay. (R&D Systems,).
149. R&D Systems. Quantikine® ELISA Kit for Measuring High Molecular Weight Adiponectin. (ed. Assay Principle of the Quantikine Human HMW Adiponectin/Acrp30 ELISA Kit) (n.d.).
150. Ibidi. The Principle of Transfection. (n.d.).
151. Lodish H, B.A., Zipursky SL, et al. *Molecular Cell Biology. Section 7.1, DNA Cloning with Plasmid Vectors*, (W. H. Freeman, New York, 2000).
152. BioCat. How 3'UTR miRNA Target Clones Work. (n.d.).
153. SwitchGear Genomics. The LightSwitch™ 3´UTR Reporter GoClone® Collection. (n.d.).
154. Promega Corporation. psiCHECK™-1 and psiCHECK™-2 Vectors. (2020).
155. Active Motif. LightSwitch™ 3´UTR Reporter GoClone® Collection. (ed. 3'UTR Vectors) (n.d.).
156. Promega Corporation. Dual-Luciferase® Reporter Assay System. (2015).
157. SwitchGear Genomics. LightSwitch™ Luciferase Assay System Overview. (n.d.).
158. Stothard, P. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques* **28**, 1102, 1104 (2000).
159. SwitchGear Genomics. The LightSwitch™ 3´UTR Reporter GoClone® Collection. (n.d.).
160. Jeppesen, M. *et al.* Short-term spheroid culture of primary colorectal cancer cells as an in vitro model for personalizing cancer medicine. *PLoS One* **12**, e0183074 (2017).
161. Merck Life Science AS. Tissue Dissociation Guide: Collagenase, Dispase, and Liberase Enzyme Types. (Sigma-Aldrich, n.d.).
162. Berens, E.B., Holy, J.M., Riegel, A.T. & Wellstein, A. A Cancer Cell Spheroid Assay to Assess Invasion in a 3D Setting. *J Vis Exp* (2015).
163. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **57**, 289-300 (1995).
164. Agarwal, V., Bell, G.W., Nam, J.W. & Bartel, D.P. Predicting effective microRNA target sites in mammalian mRNAs. *Elife* **4**(2015).
165. Takeuchi, H., Furuta, N. & Amano, A. Cell entry and exit by periodontal pathogen via recycling pathway. *Commun Integr Biol* **4**, 587-9 (2011).
166. Man, Y.G. *et al.* Tumor-infiltrating immune cells promoting tumor invasion and metastasis: existing theories. *J Cancer* **4**, 84-95 (2013).
167. Cochrane, K., Robinson, A.V., Holt, R.A. & Allen-Vercoe, E. A survey of Fusobacterium nucleatum genes modulated by host cell infection. *Microb Genom* **6**(2020).
168. Bullman, S. *et al.* Analysis of Fusobacterium persistence and antibiotic response in colorectal cancer. *Science* **358**, 1443-1448 (2017).
169. Cell biolabs. Methods for Measuring Cell Proliferation. (n.d.).
170. Cell biolabs. MTT Cell Proliferation Assay. (n.d.).
171. Cell biolabs. BrdU Cell Proliferation ELISA Kit. (n.d.).
172. Hayes, B.W. & Abraham, S.N. Innate Immune Responses to Bladder Infection. *Microbiol Spectr* **4**(2016).
173. Branchett, W.J. & Lloyd, C.M. Regulatory cytokine function in the respiratory tract. *Mucosal Immunology* **12**, 589-600 (2019).
174. Grivennikov, S.I., Greten, F.R. & Karin, M. Immunity, inflammation, and cancer. *Cell* **140**, 883-99 (2010).
175. Hattori, T. *et al.* Gene Expression Profiling of IL-17A-Treated Synovial Fibroblasts from the Human Temporomandibular Joint. *Mediators Inflamm* **2015**, 436067 (2015).

176. Harper, E.G. *et al.* Th17 cytokines stimulate CCL20 expression in keratinocytes in vitro and in vivo: implications for psoriasis pathogenesis. *J Invest Dermatol* **129**, 2175-83 (2009).
177. Schuett, J. *et al.* NADPH oxidase NOX2 mediates TLR2/6-dependent release of GM-CSF from endothelial cells. *Faseb j* **31**, 2612-2624 (2017).
178. Bergmann, M., Barnes, P.J. & Newton, R. Molecular regulation of granulocyte macrophage colony-stimulating factor in human lung epithelial cells by interleukin (IL)-1beta, IL-4, and IL-13 involves both transcriptional and post-transcriptional mechanisms. *Am J Respir Cell Mol Biol* **22**, 582-9 (2000).
179. Ghosh, S.K., Gupta, S., Jiang, B. & Weinberg, A. Fusobacterium nucleatum and human beta-defensins modulate the release of antimicrobial chemokine CCL20/macrophage inflammatory protein 3alpha. *Infect Immun* **79**, 4578-87 (2011).
180. Bartel, D.P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-33 (2009).
181. Frau, M. *et al.* Role of transcriptional and posttranscriptional regulation of methionine adenosyltransferases in liver cancer progression. *Hepatology* **56**, 165-75 (2012).
182. Lo, T.F., Tsai, W.C. & Chen, S.T. MicroRNA-21-3p, a berberine-induced miRNA, directly down-regulates human methionine adenosyltransferases 2A and 2B and inhibits hepatoma cell growth. *PLoS One* **8**, e75628 (2013).
183. Luo, J., Li, Y.N., Wang, F., Zhang, W.M. & Geng, X. S-adenosylmethionine inhibits the growth of cancer cells by reversing the hypomethylation status of c-myc and H-ras in human gastric cancer and colon cancer. *Int J Biol Sci* **6**, 784-95 (2010).
184. Friedman, R.C., Farh, K.K., Burge, C.B. & Bartel, D.P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* **19**, 92-105 (2009).
185. Farfsing, A. *et al.* Gene knockdown studies revealed CCDC50 as a candidate gene in mantle cell lymphoma and chronic lymphocytic leukemia. *Leukemia* **23**, 2018-26 (2009).
186. Shannon-Lowe, C., Rickinson, A.B. & Bell, A.I. Epstein-Barr virus-associated lymphomas. *Philos Trans R Soc Lond B Biol Sci* **372**(2017).
187. Kassambara, A. *et al.* Global miRNA expression analysis identifies novel key regulators of plasma cell differentiation and malignant plasma cell. *Nucleic Acids Res* **45**, 5639-5652 (2017).
188. Wang, Z. *et al.* Anti-apoptotic factor Birc3 is up-regulated by ELL2 knockdown and stimulates proliferation in LNCaP cells. *Am J Clin Exp Urol* **7**, 223-231 (2019).
189. Pascal, L., Zhong, M., Qiong, S., Chen, W. & Wang, Z. ELL2 and EAF2 co-regulation of AKT in prostate cancer cells. in *Society for Endocrinology BES 2018* (Glascow, UK, 2018).
190. Zanoni, M. *et al.* 3D tumor spheroid models for in vitro therapeutic screening: a systematic approach to enhance the biological relevance of data obtained. *Sci Rep* **6**, 19103 (2016).
191. Lim, W. *et al.* Formation of size-controllable tumour spheroids using a microfluidic pillar array (muFPA) device. *Analyst* **143**, 5841-5848 (2018).
192. Kelm, J.M., Timmins, N.E., Brown, C.J., Fussenegger, M. & Nielsen, L.K. Method for generation of homogeneous multicellular tumor spheroids applicable to a wide variety of cell types. *Biotechnol Bioeng* **83**, 173-80 (2003).
193. Merck Life Science AS. Liberase™ DH Research Grade. (Sigma-Aldrich, n.d.).
194. Thermo Fisher Scientific. Collagenase, Type II, powder. (n.d.).
195. Corning. Corning® Matrigel® Matrix. (n.d.).
196. Gibco™. Geltrex™ LDEV-Free Reduced Growth Factor Basement Membrane Matrix. (ThermoFisher Scientific, n.d.).
197. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **46**, D8-d13 (2018).
198. Kasper, S.H. *et al.* Colorectal cancer-associated anaerobic bacteria proliferate in tumor spheroids and alter the microenvironment. *Scientific Reports* **10**, 5321 (2020).
199. Mandujano-Tinoco, E.A. *et al.* miRNA expression profile in multicellular breast cancer spheroids. *Biochim Biophys Acta Mol Cell Res* **1864**, 1642-1655 (2017).

200. Hyder, A. Effect of the pancreatic digestion with liberase versus collagenase on the yield, function and viability of neonatal rat pancreatic islets. *Cell Biol Int* **29**, 831-4 (2005).

Supplementary section

Table of contents

1. Supplementary section 1: Initial RNA-seq screening

1.1. Results for cell lines treated with *F. nucleatum*, *E. coli* DH5α or TSB

PCA plots (Figure S1) shows that the difference in mRNA expression is larger between different cell lines (DLD-1, LS411N, SW480 and SW620) compared to different treatments (*F. nucleatum*, *E. coli* DH5α or TSB). This indicates a large natural variation in gene expression between different colon cell lines, and that the cell lines respond to infection.

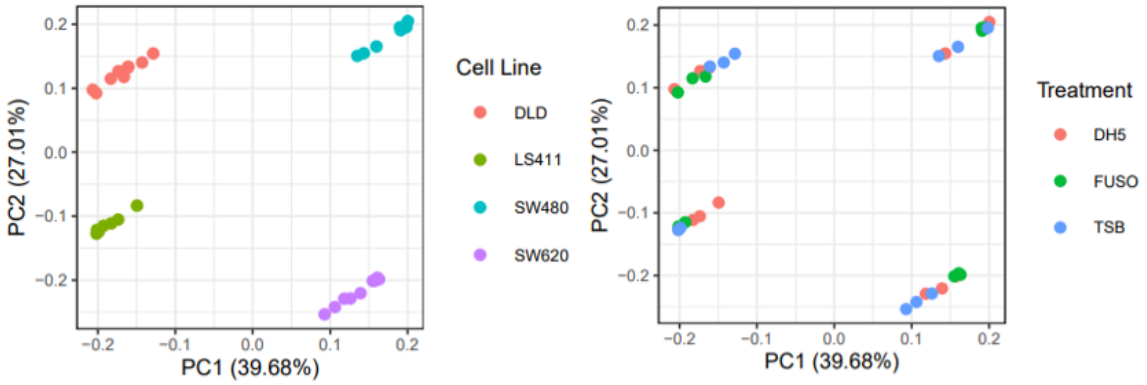


Figure S1. *PCA plots showing difference in mRNA between the cell lines DLD-1 (red), LS411N (green), SW480 (blue) and SW620 (purple) to the left, and the treatments F. nucleatum (green), E. coli DH5α (red) and TSB (blue) to the right. Difference in mRNA is bigger between the cell lines compared to the treatments.*

PCA plots for each cell line are presented in Figure S2 and show differentially expressed mRNA when treated with *F. nucleatum*, *E. coli* DH5α or TSB. The co-cultures indicate different responses for the different cell lines.

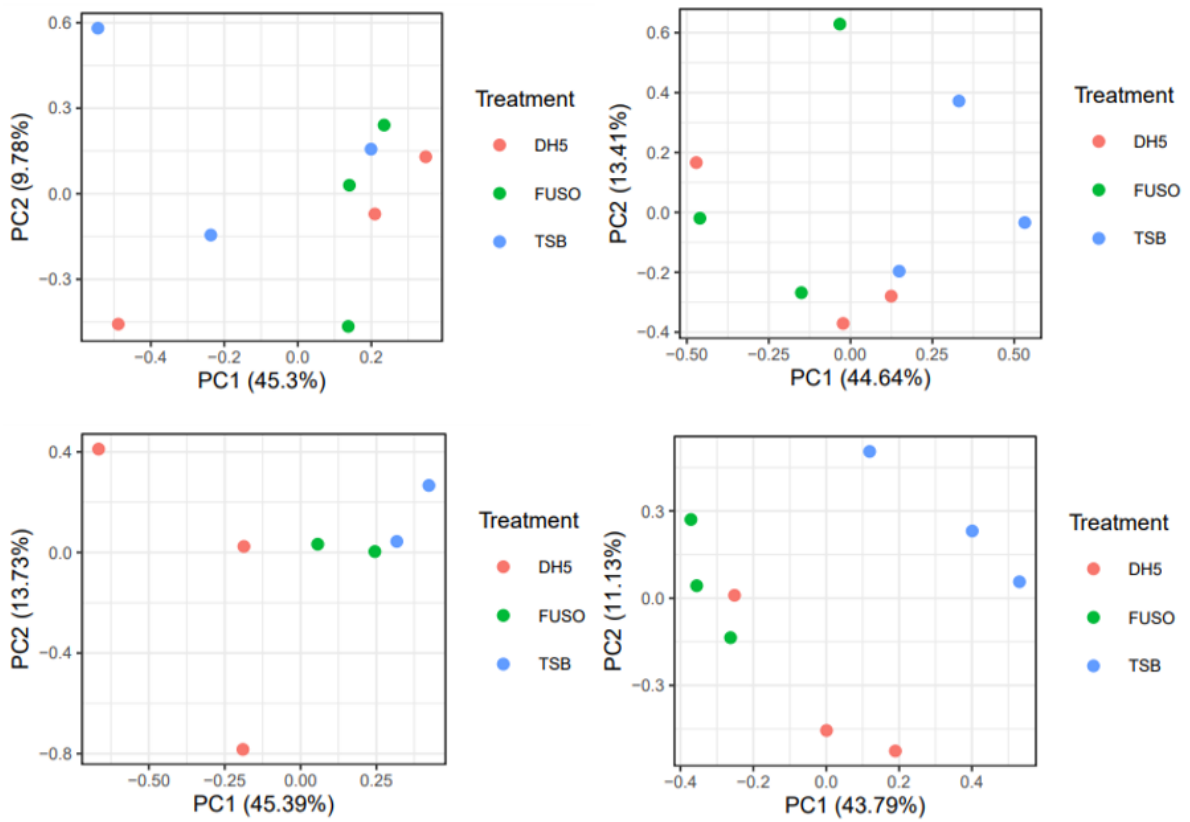


Figure S2. *PCA plots showing the difference in mRNA expression with different treatments (E. coli DH5α (red), F. nucleatum (green) or TSB (blue)) for each cell line. Upper left: SW420. Upper right: DLD-1. Lower left: LS411N. Lower right: SW620.*

Figure S3 shows a heatmap with the differentially expressed genes for SW620 in co-culture with *E. coli* DH5α, *F. nucleatum* or Tryptic soy broth (TSB). The clusters identify SW620 with different treatments (three parallels of each) whose mRNA expression levels are associated with similar levels of mRNA expression in the other treated cells. The colour of the heat map indicates the relative strength of differential expression of each particular mRNA (i.e., each pixel row) among each treatment, with red colour indicating higher expression and blue colour indicating lower expression of each mRNA. Not unexpectedly, SW620 cells with similar treatment clustered closely together, with *F. nucleatum*-treated cells far apart from TSB-treated cells. *E. coli* DH5α-treated cells showed similarities with the other treatments, but clusters closer together with *F. nucleatum*-treated cells, indicating that their gene expression profiles are more similar.

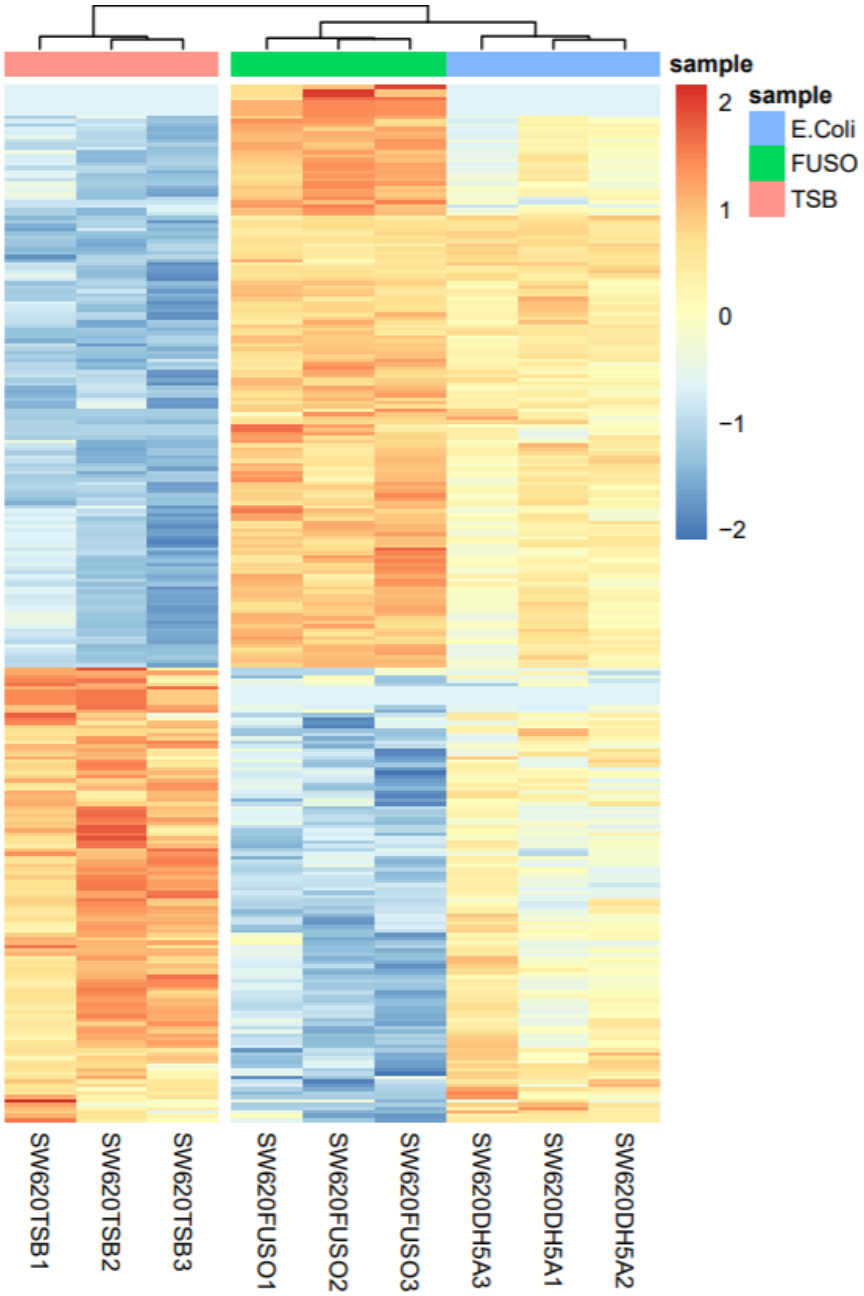


Figure S3. *A heatmap showing the differentially expressed genes for SW620 in co-culture with E. coli DH5α, F. nucleatum and TSB. The clusters identify SW620 cells with different treatments (in three parallels) whose mRNA expression levels are associated with similar levels of mRNA expression in the other treated cells. The colour of the heat map indicates the relative strength of differential expression of each particular mRNA (i.e., each pixel row) among each treatment, with red colour indicating higher expression and blue colour indicating lower expression of each mRNA. The values are z-score normalized log2-cpm expression.*

1.2. Results for cells transfected with miR-BART10-3p or negative RNA

PCA plot showing difference in mRNA between miR-BART10-3p- and negative miRNA-transfected cells (SW620 and LS411N) are presented in Figure S4. There are clear differences between the different groups, indicating that many genes are affected by the transfection. The plot shows that the variation is higher for the negative control samples compared to the miR-BART10-3p transfected cells.

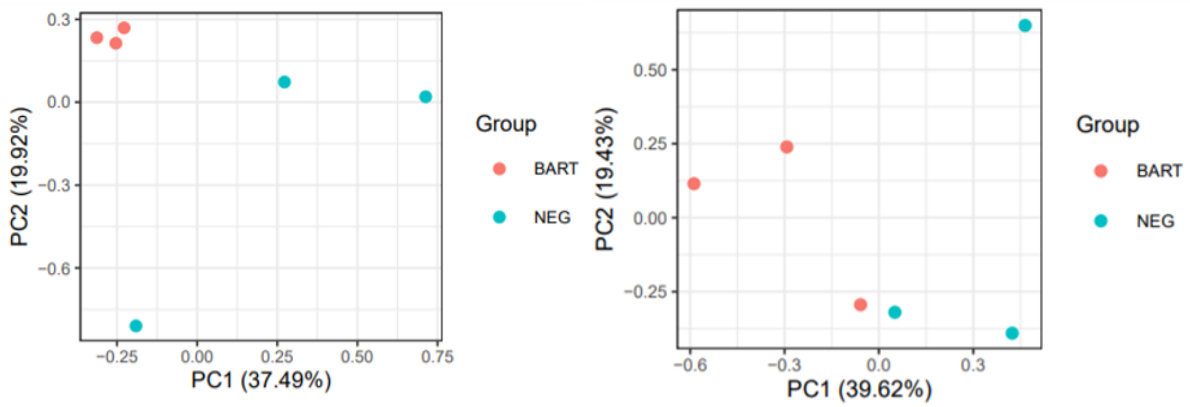


Figure S4. *PCA plots showing difference in mRNA expression in SW620 (left) and LS411N (right) transfected with miR-BART10-3p- (red) or negative miRNA (blue).*

Volcano plots showing the Benjamini-Hochberg adjusted p-value plotted against the fold change (log2) for differentially expressed mRNA in SW620 and LS411N transfected with miR-BART10-3p or a negative miRNA are presented in Figure S5. The red points indicate the significantly differentially expressed mRNAs. mRNA with a log fold change below zero are downregulated in miR-BART10-3p transfected cells compared to cells transfected with negative miRNA, whereas mRNA with a log fold change above zero are upregulated in miR-BART10-3p transfected cells compared to cells transfected with negative miRNA.

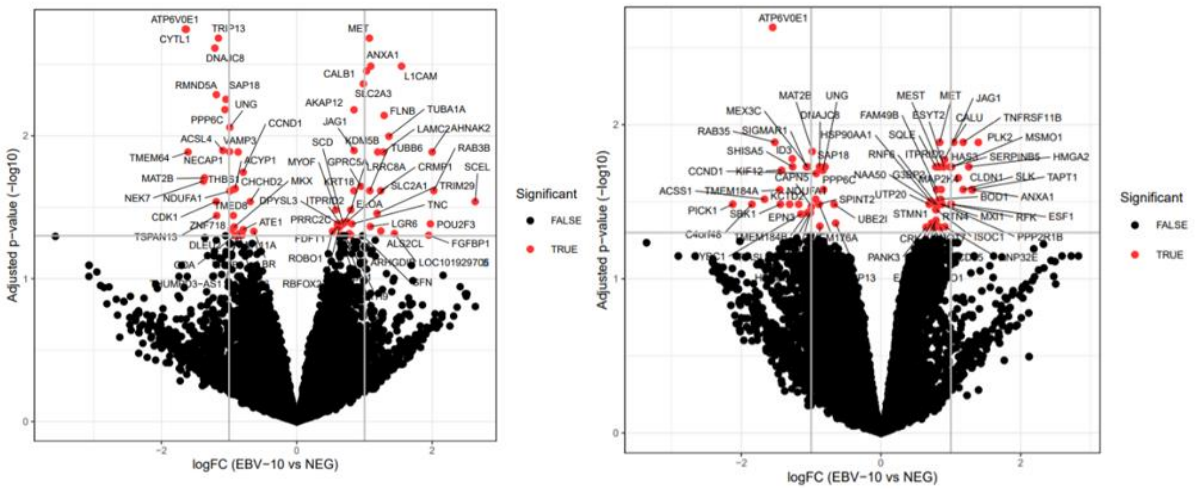


Figure S5. *Volcano plots showing the adjusted p-value plotted against the fold change (log2) for differentially expressed mRNA in SW620 (left) and LS411N (right) transfected with miR-BART10-3p or negative miRNA. Red points indicate the significantly differentially expressed mRNAs. mRNAs with a log fold change below zero are downregulated, whereas mRNAs with a log fold change above zero are upregulated. Grey horizontal line indicate significance (p = 0.05) and grey vertical line indicate fold change of -1 and 1.*

Bar plots showing the expression (count per million, cpm, log2) of *MAT2B* and *ELL2* in SW620 cells (Figure S6) and LS411N cells (Figure S7) transfected with miR-BART10-3p or a negative miRNA showing that *MAT2B* and *ELL2* are downregulated when treated with miR-BART10-3p.

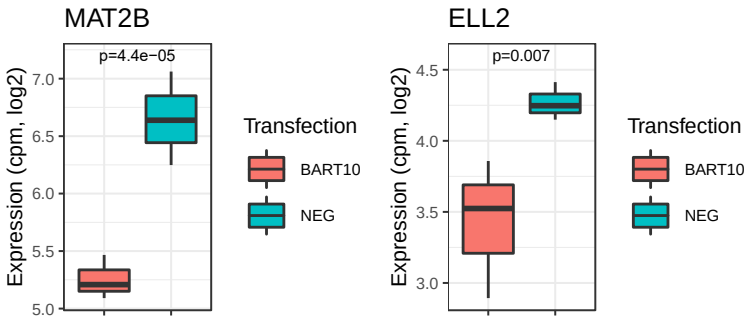


Figure S6. *Bar plot showing expression (cpm, log2) of MAT2B (left) and ELL2 (right) in SW620 cells transfected with either miR-BART10-3p (red) or negative miRNA (turquois). The p-values are non-adjusted p-values for the specific genes as calculated in limma in R.*

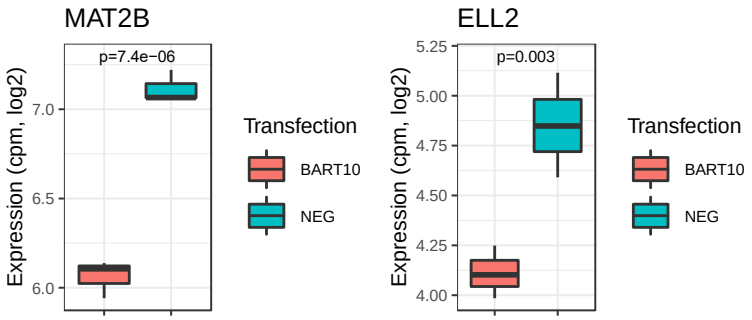


Figure S7. *Bar plot showing expression (cpm, log2) of MAT2B (left) and ELL2 (right) in LS411N cells transfected with either miR-BART10-3p (red) or negative miRNA (turquois). The p-values are non-adjusted p-values for the specific genes as calculated in limma in R.*

2. Supplementary section 2: Construction of vectors

2.1. Restriction Enzyme Cut Reaction with XhoI and NotI – Protocol

Introduction

Restriction Enzyme Cut Reaction with NotI and XhoI to prepare for ligation of *MAT2B* into the multiple cloning site of the psiCHECK™-2 vector

Materials

- NEBuffer 3.1
- NotI
- XhoI
- psiCHECK™-2
- dH_2O

Procedure

1. Thaw the frozen reagents except from the restriction enzymes.
2. Prepare an Eppendorf tube with the following reaction mixture:

Component	**Volume**
NEBuffer 3.1	5 µl
psiCHECK™-2	~ up to1 ug
NotI	1 µl
XhoI	1 µl
dH_2O	To a final volume of 50 µl
Final volume	**50 µl**

3. Incubate the tubes at 37 ºC overnight.
4. Check cut products by electrophoresis on a 0.8% agarose gel.

2.2. Gel electrophoresis – Protocol

Introduction

Agarose Gel electrophoresis of PCR products to determine the size and the purity. The gel will be stained in GelRed and visualized by the Gel Logic 200 Imaging System.

Materials

- Agarose
- TAE-buffer
- Horizon 58 Agarose Gel Electrophoresis Apparatus with voltage source
- GelRed® Nucleic Acid Gel Stain, 10 000X in water
- GelPilot DNA Loading Dye, 5x (QUIAGEN)
- GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific™)
- Gel Logic 200 Imaging System

Procedure

1. Make 0,8% agarose solution by mixing 0,8 g agarose with 100 ml 1X TAE-buffer in a glass bottle. Do not tighten the lid completely.
2. Melt the agarose in a microwave oven. Start with full power for 30-40 sec, then lower to 160 W and let boil approx. 1.5 min.
3. Let the solution equilibrate to 60 °C for 30 min in an incubator.
4. Mix the solution carefully and pour it into the electrophoresis cell with casting gates and a comb. Let the gel settle for 30 min.
5. Mix PCR-product 1:5 with GelPilot DNA Loading Dye.
6. Remove the casting gates and the comb carefully and add 1X TAE-buffer into the electrophoresis cell enough to cover the gel.
7. Apply GeneRuler 1 kb Plus DNA Ladder and the PCR-products carefully into the wells.
8. Run the gel at 100 V constant voltage until the dye line is approximately 75-80% of the way down the gel.
9. Transfer the gel to a GelRed® staining bath. Stain for 30 min.
10. Wash the gel briefly in RO-water and visualize by Molecular Imaging Software with Gel Logic 200 Imaging System.

2.3. Heat Shock Transformation – Protocol

Introduction

Transformation to identify and amplify psiCHECK^{TM}-2 vectors that contain the desired *MAT2B* 3´UTR insert. The psi-CHECK^{TM}-2 vector contains an Ampicillin (Amp)-resistance gene. Following transformation, bacteria (DH5α competent *E. coli*) are selected on an Amp-LB dish. Only bacteria that have taken up a vector are antibiotic-resistant and are able to form a colony on the Amp-LB dish.

Materials

- Ligation product (vector + insert)
- Control ligation product (only vector, no insert)
- DH5α competent *E. coli*
- SOC medium
- Amp-LB dishes
- Amp-LB medium

Procedure day 1 – Heat Shock Transformation

1. Set water bath to 42ºC.
2. Thaw 2 tubes of DH5α competent *E. coli* on ice (approximately 20-30 min).
3. Thaw 2 tubes of SOC medium and 2 Amp-LB dishes in a 37ºC incubator.
4. Add 1-10 µl (up to 100 ng) of the ligation product (vector + insert) to the one tube of DH5α competent *E. coli* and mark the tube (Insert).
5. Add 1-10 µl (up to 100 ng) of the control ligation product (only vector, no insert) to the other tube of DH5α competent *E. coli* and mark the tube (Control).
6. Incubate for 15-20 minutes on ice.
7. Incubate for 50 seconds in the 42ºC water bath.
8. Incubate for 2 minutes on ice to reduce damage to the DH5α competent *E. coli.*
9. Add 900 µl SOC medium to each of the tubes and incubate for 60-90 minutes at 37ºC in a shaking incubator.
10. Centrifuge at max speed for 1 minute and discard most of the supernatant.
11. Resuspend the pellet by pipetting up and down.
12. Seed the suspension evenly to an Amp-LB dish.
13. Incubate overnight at 37ºC with the plate lid face down.

Procedure day 2 – Pick colonies

1. Inspect the Amp-LB dishes. Ideally, many colonies should be present on the insert dish while few colonies should be present on the control dish. Decide how many colonies to pick from the insert-dish.
2. Prepare the appropriate number of culture tubes (if 10 colonies are to be picked, prepare 10 tubes).
3. Add 100 µl Amp to 100 ml LB medium. Distribute 3 ml of Amp-LB medium to 10 culture tubes. Store the remaining Amp-LB medium at 4ºC for up to 1 month.
4. Pick 10 colonies on the Amp-LB dish using a sterile tooth pick. Use a pair of tweezers when taking out sterile tooth picks from the container. Burn off the tweezers before and after each use. Put the correct tooth pick into the corresponding culture tube.
5. Incubate at 37ºC overnight.

Procedure day 3 – Preparations for DNA purification

1. Prepare 10 sterile Eppendorf tubes for each of the 10 culture tubes.
2. Pour 1,5 ml bacterial culture from the culture tubes into the Eppendorf tube.

   * Work sterile when handling bacterial cultures. Use a gas burner to sterilise the tip of the tubes before pouring. Burn off the tip of the tubes once again before putting the cap back on.
3. Centrifuge at max speed for 5 minutes and discard the supernatant.
4. Pour the rest of the bacterial culture (1,5 ml) to the same Eppendorf tube to increase the bacterial yield. Pipet up and down to resuspend the pellet.
5. The bacterial culture is now ready for DNA purification with the Wizard® Plus SV Minipreps DNA Purification System (Promega).

3. Supplementary section 3: Sequencing results *MAT2B*

3.1. FASTA sequence of *MAT2BB* inserted into a psiCHECK™-2 vector

Accurate cloning of the *MAT2B* insert into the psiCHECK™-2 vector was confirmed by detecting the reverse complement of MAT2B.1 (Table 3) within the psiCHECK™-2 sequence that was sequenced by GATC Biotech (Konstanz, Germany).

The reverse complement of MAT2B.1 is marked with yellow in the FASTA sequence, indicating a successful cloning of *MAT2B* into the psiCHECK™-2 vector.

>80GC99_91745231_91745231GAAGAATCATTTAGATCCTCACACAAAAAACCAACA
CACAGATGTAATGAAAATAAAGATATTTTATTGCGGCCAGCGGCCGCGTTCCAAT
ATACATAACATGTTGAAATATCTCAAGCCTCGAGCGATCGCCTAGAATTACTGCT
CGTTCTTCAGCACGCGCTCCACGAAGCTCTTGATGTACTTACCCATTTCATCTGGA
GCGTCCTCCTGGCTGAAGTGGAGGCCCTTCACCTTCACGAACTCGGTGTTAGGGA
ACTTCTTAGCTCCCTCGACAATAGCGTTGGAAAAGAACCCAGGGTCGGACTCGAT
GAACATCTTAGGCAGATCGTCGCTGGCCCGAAGGTAGGCGTTGTAGTTGCGGACA
ATCTGGACGACGTCGGGCTTGCCTCCCTTAACGAGAGGGATCTCGCGAGGCCAGG
AGAGGGTAGGCCGTCTAACCTCGCCCTTCTCCTTGAATGGCTCCAGGTAGGCAGC
GAACTCCTCAGGCTCCAGTTTCCGCATGATCTTGCTTGGGAGCATGGTCTCGACG
AAGAAGTTATTCTCAAGCACCATTTTCTCGCCCTCTTCGCTCTTGATCAGGGCGAT
ATCCTCCTCGATGTCAGGCCACTCGTCCCAGGACTCGATCACGTCCACGACACTC
TCAGCATGGACGATGGCCTTGATCTTGTCTTGGTGCTCGTAGGAGTAGTGAAAGG
CCAGACAAGCCCCCCAGTCGTGGCCCACAAAGATGATTTTCTTTGGAAGGTTCAG
CAGCTCGAACCAAGCGGTGAGGTACTTGTAGTGATCCAGGAGGCGATATGAGCC
ATTCCCGCTCTTGCCGGACTTACCCATTCCGATCAGATCAGGGATGATGCATCTA
GCCACGGGCTCGATGTGAGGCACGACGTGCCTCCACAGGTAGCTGGAGGCAGCG
TTACCATGCAGAAAAATCACGGCGTTCTCGGCGTGCTTCTCGGAATCATAGTAGT
TGATGAAGGAGTCCAGCACGTTCATTTGCTTGCAGCGAGCCCACCACTGAGGCCC
AGTGATCATGCGTTTGCGTTGCTCGGGGTCGTACACCTTGGAAGCCATGGTGGCT
AGCCTATAGTGAGTCGTATTAAGTACTCTAGCCTTAAGAGCTGTAATTGAACTGG
GGAGTGGA

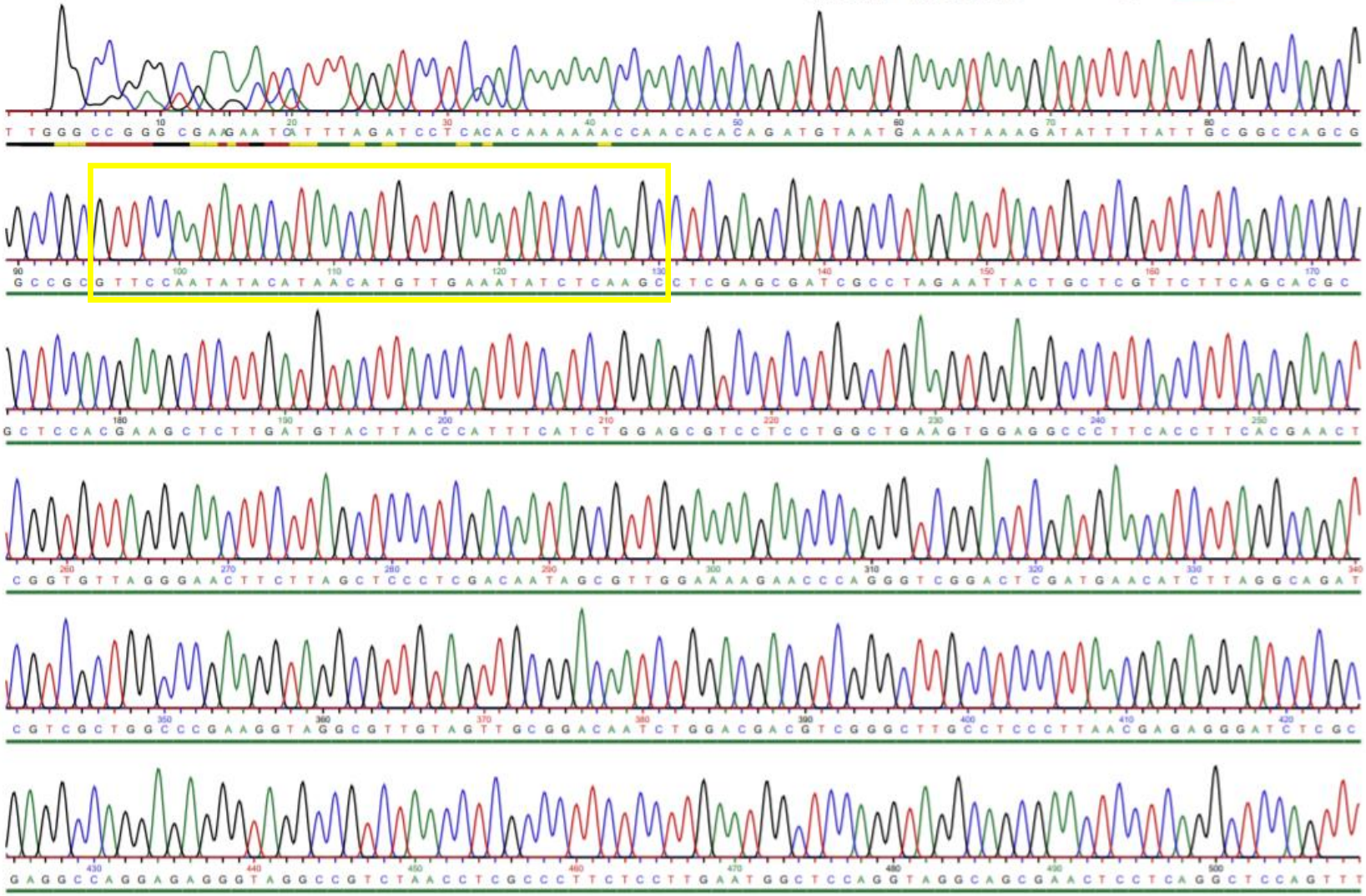
3.2. RNA sequencing chromatogram of *MAT2B* inserted into psiCHECK-2

RNA sequencing chromatogram showing successful cloning of the *MAT2B* insert into the psiCHECK™-2 vector are presented in Figure S8. The reverse complement of MAT2B.1 is marked with yellow.

Sequence: 91745231

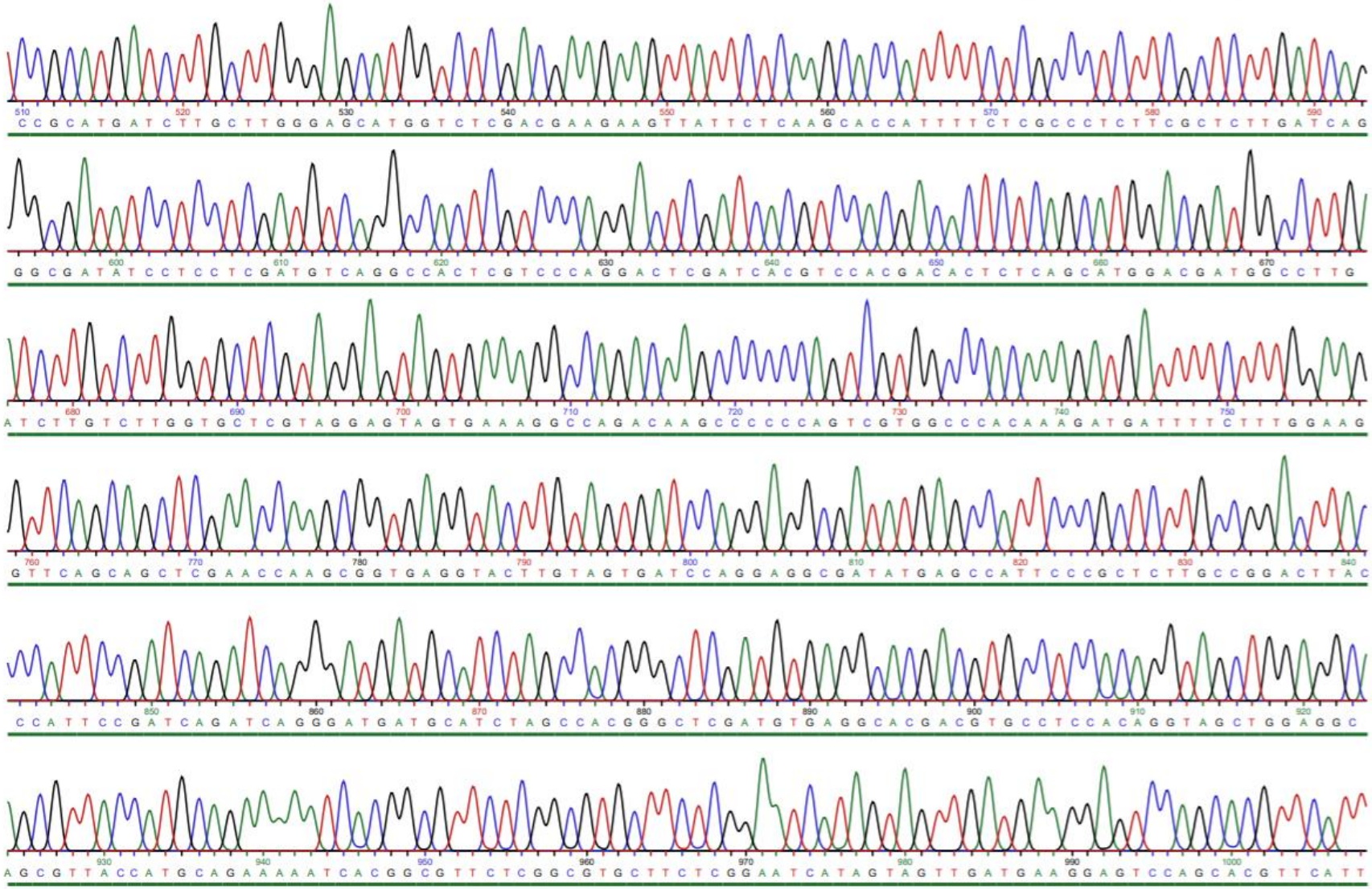
Samples:	15833	Quality: 0 - 9	
Bases:	1323	10 - 19	
Average spacing:	12.0	20 - 29	
Average quality >=	10: 75, 20: 68, 30: 1113	>= 30	

26.02.2020



Sequence: 91745231

Samples:	15833	Quality:	0 - 9
Bases:	1323		10 - 19
Average spacing:	12.0		20 - 29
Average quality >=	10: 75, 20: 66, 30: 1113		>= 30



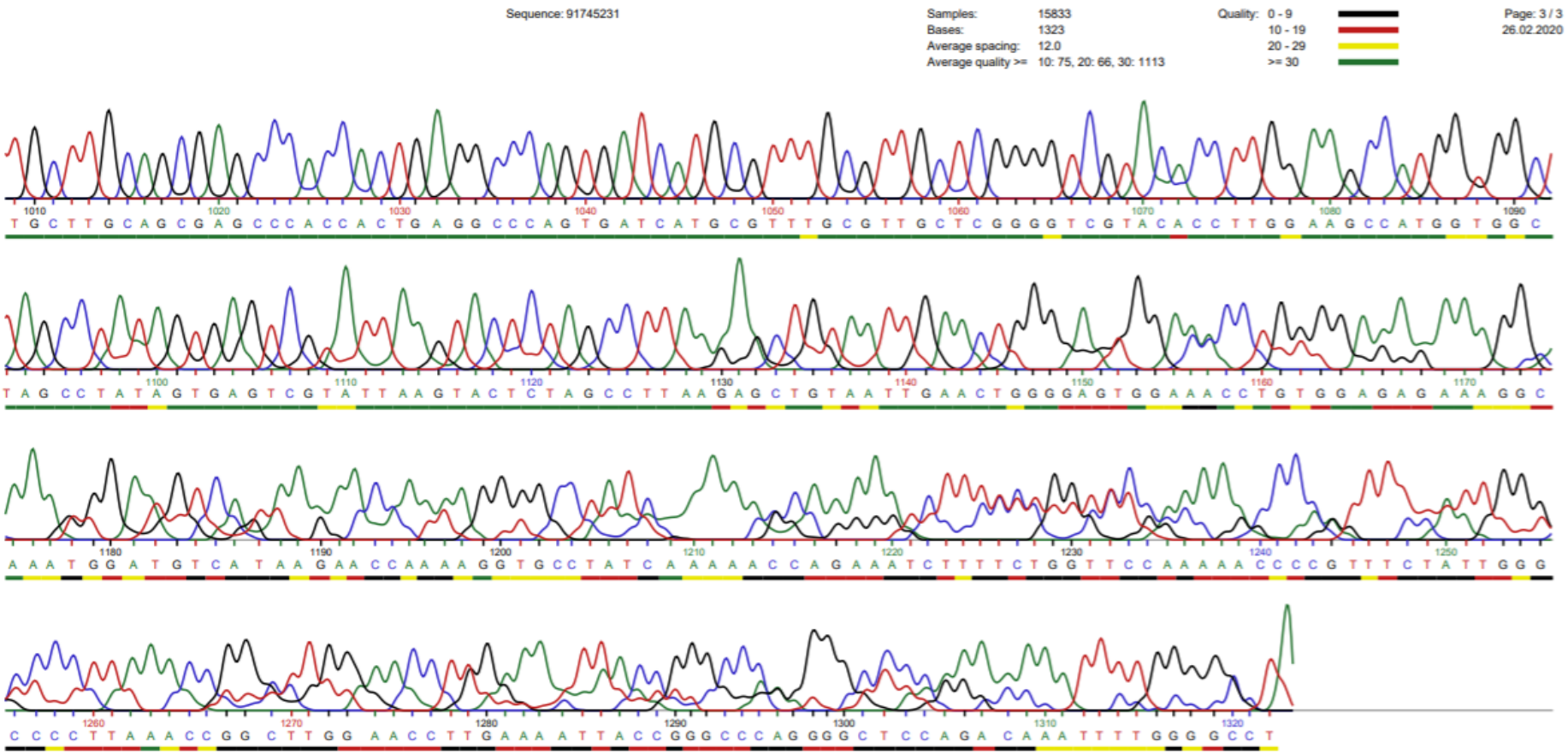


Figure S8. *RNA sequencing chromatogram (three pages) showing successful cloning of MAT2B into the psiCHECK 2 vector. The reverse complement of OLIGO1 is marked with yellow.*

4. Supplementary section 4: Patient-derived CRC spheroids

4.1. Applied reagent

Applied reagents with supplier are listed in Table S1.

Table S1. *Applied reagents and its ingredients with supplier*

Applied reagents	Ingredients and supplier
Supplemented DMEM	DMEM (D6429-500ML; Sigma Aldrich) 100 units/mL penicillin (15140-122; Gibco) 100 µg/mL streptomycin (15140-122; Gibco)
Supplemented DMEM + Liberase DH	Supplemented DMEM 0.28 units/mL Liberase DH (5401054001; Roche)
Supplemented DMEM + Collagenase type II	Supplemented DMEM 0.001 g/mL Collagenase type II (17101015; Gibco)
Hank's Balanced Salt Solution (HBSS)	HBSS, calcium, magnesium, no phenol red (14025-050; Gibco)
HBSS + BSA	HBSS, calcium, magnesium, no phenol red (14025-050; Gibco) 1% BSA
Phosphate-buffered saline (PBS)	
Fetal calf serum (FCS)	
Supplemented serum-free stem cell medium	DMEM/F12 + GlutamaxTM -I (10565-018; Gibco) StemPro® hESC Supplement (A10006-01; Gibco) BSA 1.8% (A10008-01; Gibco) 8 ng/mL bFGF (13256-029; Gibco) 0.1 mM 2-mercaptoethanol (21985-023; Thermo Fisher Scientific) 100 units/mL penicillin (15140-122; Gibco) 100 µg/mL streptomycin (15140-122; Gibco) 25 µg/mL amphotericin B (A2942-100ML; Sigma Aldrich)
Collagen solution	A, Cellmatrix type I-A (631-00651; Nitta Gelatin Inc.) B, DMEM (D6429-500ML; Sigma Aldrich) C, Reconstitution buffer 50 mM NaOH 260 mM NaHCO3 200 mM Hepes
DMEM + Collagenase Type IV	DMEM (11965-092; Gibco) 0.2 mg/ml Collagenase Type IV (LS004186; Worthington)
Cellmatrix Type 1-A	Cellmatrix Type 1-A (Collagen, Type I, 3 mg/mL, pH 3.0) (637-00653, Nitta Gelatin Inc.)
Cultrex	Cultrex RGF Basement Membrane Extract, Type 2 (3533-010-02, R&D Systems)
Geltrex	Geltrex™ LDEV-Free, hESC-Qualified, Reduced Growth Factor Basement Membrane Matrix (A1413302; Gibco)
Matrigel	Corning® Matrigel® Matrix, phenol red-free (734-0272, Corning)

4.2. Protocol for establishing patient-derived colorectal cancer spheroids

The CTOS preparation is described in Subheading 4.2.1. with the optimization steps written in cursive below the initial steps. For the CTOS expansion, two separate protocols are described for the initial phase (Subheading 4.2.2.) and the optimization phase (Subheading 4.2.3.). This protocol is a modified protocol developed by Evelina Folkesson.

4.2.1. CTOS preparation – Initial phase and optimization phase (in cursive)

1. Add 20 mL of supplemented DMEM in a 50-mL centrifuge tube and weight the tube.
2. Place the tissue sample in tube prepared in step 1 and put it on ice immediately after tumor resection or biopsy.
3. Store the specimen at 4 °C until ready to proceed. It is critical to start the following steps as soon as possible.
4. Weight the tube and subtract the weight from step 1 to get the weight of the sample.
5. Discard the storage medium and wash the samples with 20 mL HBSS by inverting the tube. Discard the HBSS wash solution.
6. Add 20 mL HBSS and transfer the medium and the samples to a 10-cm tissue culture dish.
7. Remove necrotic tissue using forceps or razor blades.

   Optimization procedure:

   - *Cut sample in two. Transfer the samples to two separate 50-mL tubes with 20 mL HBSS and label them "L" and "C". Weight the tubes before and after addition of the sample.*
8. Transfer the sample to 30 mL HBSS in a new 10-cm tissue culture dish.

   Optimization procedure:

   - *Transfer the medium and the sample to two separate 10-cm tissue culture dishes labelled "L" and "C".*
9. Mince the tissue with forceps or razor blades into small (1–2 mm) pieces.
10. Transfer the medium and the minced tissue to a 50-mL centrifuge tube.

    Optimization procedure:

    - *Transfer the medium and the minced tissue to two separate 50-mL centrifuge tubes labelled "L" and "C".*
11. Centrifuge at 200 × g at 4 °C for 5 min and discard the supernatant.

    Optimization procedure:

- *Centrifuge at 400 × g at 4 °C for 5 min and discard the supernatant.*

12. Wash the samples with 20 mL HBSS by inverting the tube.
13. Repeat steps 11-12.
14. Centrifuge at 1,000 rpm (200 × g) at 4 °C for 5 min and discard supernatant.
15. Resuspend the pellets in 20 mL supplemented DMEM + Liberase DH.

    Optimization procedure:

    - *Resuspend pellet in the "L" tube in 20 mL supplemented DMEM + Liberase DH*
    - *Resuspend pellet in the "C" tube in 20 mL supplemented DMEM + collagenase type II.*

16. Transfer the digestion mixture to a 100-mL sterile conical flask with a magnet bar.
17. Digest the samples for 2 h in a 37 °C water bath with constant stirring.

    Optimization procedure:

    - *Digest the "L" sample for 2 h in a 37 °C water bath with constant stirring.*
    - *Digest the "C" sample for 20 min in a 37 °C water bath with constant stirring.*

      **All steps below are done separately for "L" and "C". Mark all tubes and plates with "L" and "C" accordingly*

18. Transfer the digestion medium to a 50-mL centrifuge tube.
19. Centrifuge at 200 × g at 4 °C for 5 min and discard the supernatant.

    Optimization procedure:

    - *Centrifuge at 400 × g at 4 °C for 5 min and discard the supernatant.*

20. Wash the samples with 20 mL HBSS by inverting the tube.
21. Filter the samples with a stainless steel wire mesh (hole size 500 μm).

    Optimization procedure:

    - *Filter the samples with a 500 μm pluriStrainer filter.*

22. Filter the samples with a 40-μm cell strainer.
23. Dip the bottom of the cell strainer in 30 mL HBSS in a 10-cm tissue culture dish; swirl it gently to remove the debris, the single cells, and the cell clumps with diameters < 40 μm.
24. Transfer the cell strainer to a new 10-cm tissue culture dish containing 30 mL HBSS.
25. Collect the spheroids that remain in the cell strainer using a 1 mL pipette.
26. Centrifuge at 200 × g at 4 °C for 5 min and discard the supernatant.

    Optimization procedure:

    - *Centrifuge at 400 × g at 4 °C for 5 min and discard the supernatant.*

27. Wash the samples with 20 mL HBSS by inverting the tube.

28. Centrifuge at 200 × g at 4 °C for 5 min and discard the supernatant.

    Optimization procedure:

    - *Centrifuge at 400 × g at 4 °C for 5 min and discard the supernatant.*

      **From here, go straight to "CTOS Culture – optimization phase" in Subheading 4.2.2.*

29. Add 400 µL supplemented serum-free stem cell medium.
30. Transfer the spheroids and medium to a 6-cm non-treated dish.
31. Incubate in a 5% CO2-humidified chamber at 37 °C for 24 h.
32. View under a phase contrast microscope; CTOSs appear as bright, smooth spheres.

4.2.2. <u>CTOS culture and expansion – Initial phase</u>

1. Prepare a collagen solution by mixing the following: A, Cellmatrix type I-A (Nitta Gelatin); B, 5× DMEM (12100-038; Gibco); C, reconstitution buffer (50 mM NaOH, 260 mM NaHCO3, 200 mM Hepes). A:B:C = 7:2:1. After mixing A and B well, add C and again mix well. Keep all reagents on ice to prevent gel formation.
2. Pour 130 µL reconstituted collagen solution onto a 3.5-cm non-treated dish to create a gel base.
3. Allow the gel to solidify at 37 °C for 30 min.
4. Gently pick up 30 µL CTOS suspension using a pipette. Put the CTOSs into the upper gel layer, pipetting gently to disperse the CTOSs evenly.
5. Allow the gel to solidify at 37 °C for 15 min.
6. Add 3 mL supplemented serum-free stem cell medium.
7. Incubate in a 5% CO2-humidified chamber at 37 °C for 2–3 weeks until the diameters of the CTOSs are ~250 µm. Change the medium every 3. day.
8. Discard the medium
9. Digest the gels with 3 mL DMEM + Collagenase Type IV at 37 °C for 1 h.
10. Release the CTOSs from the digested gel by pipetting up and down.
11. Transfer the CTOS suspension to a 15-mL centrifuge tube.
12. Add 10 mL PBS and mix gently.
13. Centrifuge at 1,000 rpm (200 × g) for 2 min and discard the supernatant.
14. Add 5 mL HBSS containing 1% BSA and resuspend the CTOSs by pipetting; transfer them to a 6-cm non-treated culture dish.

15. Using microscopic observation, tear CTOSs using two sterile 23-gauge needles. One needle (held with the nondominant hand) holds the CTOS steady while a needle in the dominant hand tears the CTOS. Tear each CTOS into two to four pieces.
16. Add 3 mL serum-free stem cell medium to a new 3.5-cm non-treated dish.
17. Pick up the CTOS fragments with a pipette and transfer them to the dish prepared in step 16.
18. Incubate in a 5% CO2-humidified chamber at 37 °C overnight.
19. Use a phase contrast microscope to view the CTOSs, which appear as bright, smooth spheres.

4.2.3. <u>CTOS culture – Optimization phase</u>

1. Resuspend spheroids in 0.4 mL of HBSS.
2. Add a small volume (10 uL) to a glass slide and count the number of spheroids. Multiply the number of spheroids with 400 uL to get the number of spheroids in the sample.
3. Label 8 Eppendorf tubes: “LC”, “LG”, “LM”, “LK”, “CC”, “CG”, “CM”, “CK”. (“LC” = Liberase DH, Cultrex; “LG” = Liberase DH, Geltrex; “LM” = Liberase DH, Matrigel; “LK” = Liberase DH, Cellmatrix Type 1-A; “CC” = Collagenase type II, Cultrex; “CG” = Collagenase type II, Geltrex; “CM” = Collagenase type II, Matrigel; and “CK” = Collagenase type II, Cellmatrix Type 1-A).
4. Aliquot the 0.4 mL spheroid suspension to the four corresponding Eppendorf tubes (“LC”, “LG”, “LM”, “LK” for the “L” and “CC”, “CG”, “CM”, “CK” for the “C”).
5. Centrifuge at 200 × g at 4 °C for 5 min and discard the supernatant.
6. Put all the tubes on ice.
7. Resuspend spheroids in gel and medium according to the volumes below. Start by adding medium to all of the tubes, but then finish and seed the tubes one by one; i.e. after adding gel (and reconstitution buffer (RB)) immediately seed out 50 uL per well in two wells of a 24-well plate (layout below). Keep all reagent on ice.

Gel composition				
	LC, CC	LG, CG	LM, CM	LK, CK
Medium	33 uL	33 uL	50 uL	20 uL
Gel	67 uL	67 uL	50 uL	70 uL
RB	-	-	-	10 uL

Plate layout						
	1	2	3	4	5	6
A	LC	LC			CC	CC
B	LG	LG			CG	CG
C	LM	LM			CM	CM
D	LK	LK			CK	CK

8. Incubate the plate for 30 minutes at 37 °C.
9. Add 0.5 mL of supplemented serum-free stem cell medium to each well.
10. Incubate overnight at 37 °C.
11. Change medium following overnight incubation and then change the medium every third day.

4.2.4. <u>CTOS expansion, freezing and thawing – Optimization phase</u>

1. Prewarm a 24-well plate at 37°C.
2. Discard the medium.
3. Digest the gels with 0.5 mL DMEM + Collagenase Type IV at 37 °C for 1 hour.
4. Release the CTOSs from the digested gel by pipetting up and down.
5. Transfer the CTOS suspension to a 50-mL centrifuge tube.
6. Add double volume of PBS.
7. Centrifuge at 400 × g for 2 min and discard the supernatant.
8. Resuspend the pellet in 20 mL supplemented DMEM + collagenase type II.
9. Digest the sample for 20 min in a 37 °C water bath with constant stirring.
10. Repeat step 18-27 in the "CTOS preparation – Initial phase and optimization phase (in cursive)" in subheading 4.2.1. and follow the optimization steps written in cursive.
11. Centrifuge at 400 × g at 4 °C for 5 min.
12. Discard the HBSS wash solution and resuspend spheroids in 400 µL of HBSS.
13. Separate the sample into two separate tubes labelled "Split" and "Freeze" and continue with one or both of the following procedures.

<u>Expansion procedure:</u>

i) Centrifuge at 400 × g at 4 °C for 5 min and discard the HBSS wash solution.

ii) Resuspend in DMEM in a volume appropriate for the number of gels you want to seed out (20% of total volume).

iii) Add Cellmatrix Type 1-A; volume appropriate for the number of gels you want to seed out (70% of total volume).

iv) Add reconstitution buffer; volume appropriate for the number of gels you want to seed out (10% of total volume).

v) Add 50 µL per well to a prewarmed 24-well plate.

vi) Incubate at 37°C, then add 0.5 mL medium per well.

vii) Incubate overnight at 37°C.

viii) Change medium following overnight incubation and then change the medium every third day.

Freezing and thawing procedure:

i) Centrifuge at 400 × g at 4°C for 5 min and discard the HBSS wash solution

ii) Resuspend in the FCS volume appropriate for achieving a spheroid density of approximately 210 spheroids/mL.

iii) Add DMSO so that the DMSO concentration is 5% of the final volume.

iv) Immediately put the cryovial in an isopropanol box and store at -80 °C until use.

v) When sample is going to be thawed, place a 24-well culture plate in a 37 °C incubator to warm for at least 1 hour.

vi) Transfer the cryovial from -80 °C freezer and immediately place in a 37 °C water bath and thaw rapidly. Be careful not to submerge the neck of vial. This process should take less than 2 minutes.

vii) Transfer the contents of the cryovial drop-wise to a 15 mL conical tube containing 6 mL of serum-free stem cell medium.

viii) 5Wash the cryovial three times with serum-free stem cell medium (1 mL/wash) and transfer to the 15 mL tube – i.e. should be 10 mL in the tube afterwards.

ix) Centrifuge the conical tube at 400 x g at 4°C for 5 minutes.

x) Carefully aspirate the supernatant without disturbing the pellet while removing as much liquid as possible.

xi) Resuspend the pellet in 70 uL Cellmatrix Type 1-A, 20 uL DMEM and 10 uL reconstitution buffer in the respective order. Mix properly between each step.

xii) Quickly plate out two 50 uL gel drops in two of the wells in the 24-well plate. NB! The gel solidifies quickly once the reconstitution buffer is added.

xiii) Incubate at 37°C for 30 minutes and add 0.5 mL of serum-free stem cell medium.

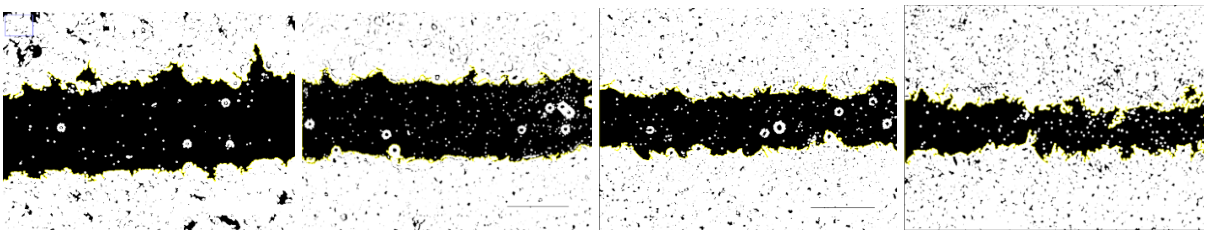
xiv) Incubate the at 37°C plate overnight – then change the medium every third day.

5. Supplementary section 5: Migration assay

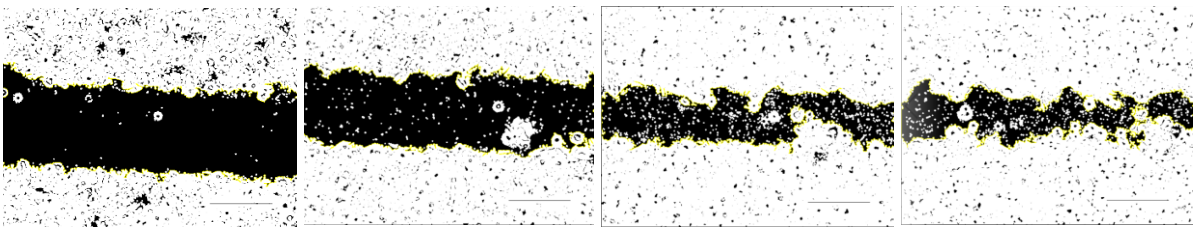
Images were taken in a phase contrast microscope at 10x and edited at ImageJ as follows; Image type was changed to 8-bit by choosing Image > Type > 8-bit; the bandpass filter was checked by choosing Process > FFT > Bandpass Filter; the image was changed to Black & White by choosing Image > Adjust > Treshold > B&W, and pixels in the image were highlighted so that cells turned white and background turned black; the filter was changed to a minimum radius of 7,0 pixels by choosing Process > Filters > Minimum; and then the wand (tracing) tool was used to mark the area to be measured. Edited images from four biological replicates are shown below.

Biological replicate 1

Untreated DLD-1: 0h, 24, 48h and 72h

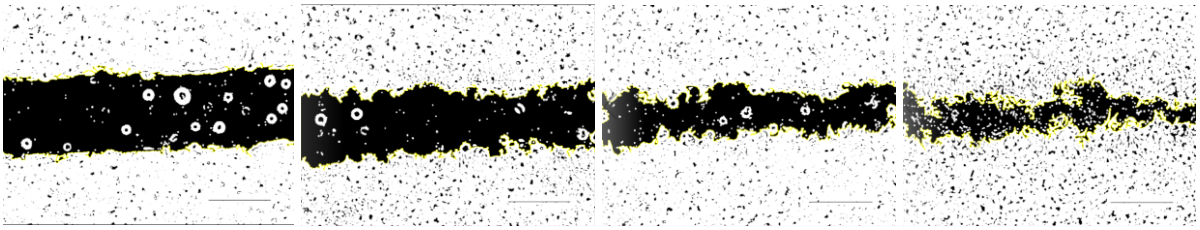


F. nucleatum treated DLD-1: 0h, 24h, 48h and 72h

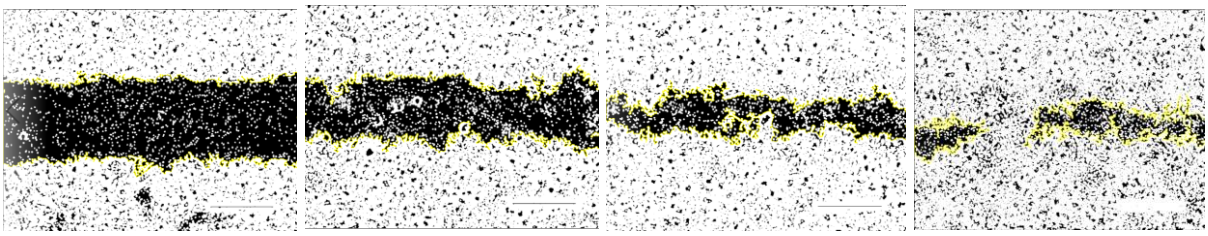


Biological replicate 2

Untreated DLD-1: 0h, 24, 48h and 72h

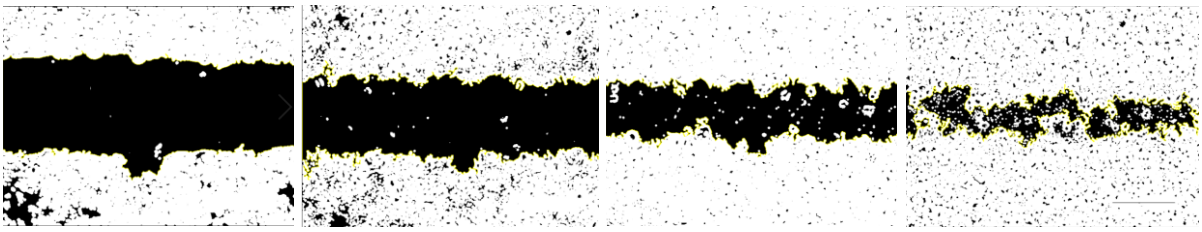


F. nucleatum treated DLD-1: 0h, 24h, 48h and 72h

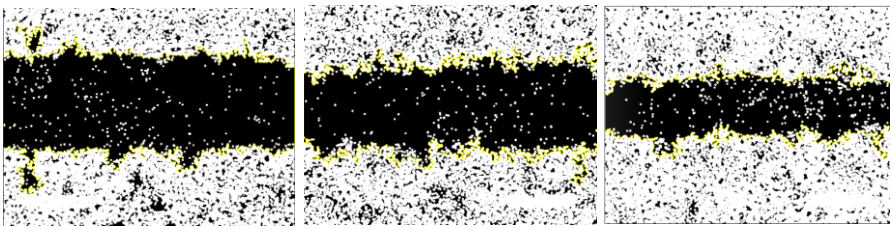


Biological replicate 3

Untreated DLD-1: 0h, 24, 48h and 72h

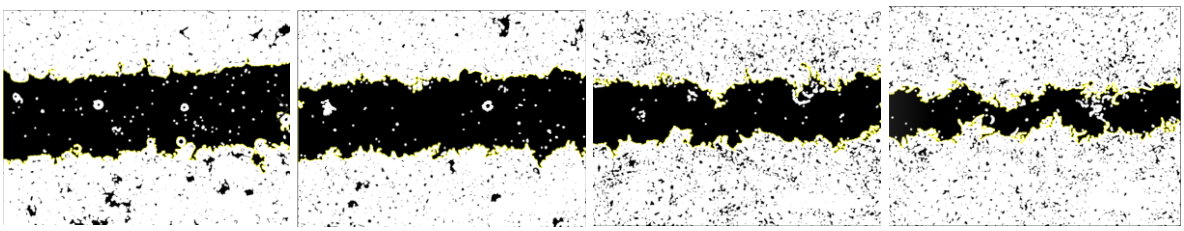


F. nucleatum treated DLD-1: 0h, 24h and 48h (Technical issue with image at 72h)

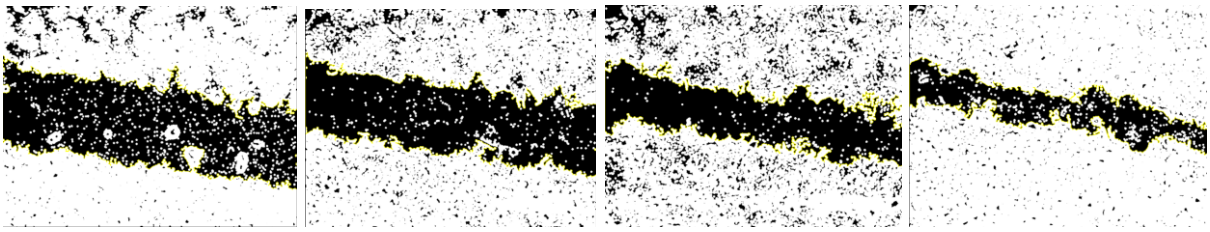


Biological replicate 4

Untreated DLD-1: 0h, 24, 48h and 72h



F. nucleatum treated DLD-1: 0h, 24h, 48h and 72h



The area was measured at ImageJ by choosing Analyze > Measure, and area size within the marked area was reported. Results were exported to excel, where percent wound closure of the area was calculated. Raw data, including fold change of the difference between each timepoint and start point and the percent wound closure are presented in Table S2.

Table S2. *Migration assay of untreated- and F. nucleatum treated DLD-1 at 0, 24, 48 and 72 hours. Table includes size of non-migrated area, fold change and percent wound closure.*

Biological replicate 1			
DLD Blank	**Area (pixels)**	**Fold change**	**% wound closure**
0h	1298970	1	0
24h	1081683	0.832723619	16.72763805
48h	784082	0.603618251	39.63817486
72h	569616	0.438513592	56.14864085
DLD F. nuc			
0h	1197559	1	0
24h	1004165	0.838509835	16.14901646
48h	592100	0.494422404	50.55775958
72h	496153	0.414303596	58.56964041
Biological replicate 2			
DLD blank	**Area (pixels)**	**Fold change**	**% wound closure**
0h	1080817	1	0
24h	866669	0.8018647	19.81352995
48h	582122	0.538594415	46.14055848
72h	426386	0.394503417	60.54965827
DLD F. nuc			
0h	1144670	1	0
24h	808029	0.705905632	29.40943678
48h	430137	0.375773804	62.42261962
72h	271297	0.23700892	76.29910804
Biological replicate 3			
DLD blank	**Area (pixels)**	**Fold change**	**% wound closure**
0h	1349710	1	0
24h	1064383	0.788601255	21.13987449
48h	724183	0.53654711	46.34528899
72h	374436	0.277419594	72.25804062
DLD F. nuc			
0h	1427057	1	0
24h	1246118	0.873208288	12.67917119
48h	849846	0.595523514	40.44764855
72h	Undetermined		
Biological replicate 4			
DLD Blank	**Area (pixels)**	**Fold change**	**% wound closure**
0h	1135218	1	0
24h	1056655	0.93079479	6.920520992
48h	817051	0.719730483	28.02695165
72h	543334	0.478616442	52.1383558
DLD F. nuc			
0h	1083777	1	0
24h	962352	0.887961269	11.20387312
48h	713113	0.657988682	34.20113178
72h	452194	0.417238971	58.27610293

The p-value was calculated in Microsoft Excel by doing a students' t-test for each time point. "T-test: Two-sample Assuming Unequal Variances" was chosen. As the *F. nucleatum*-treated cells are hypothesised to be greater than the untreated cells, the p-value is reported as one-tailed. As shown in Table S3, the p-value (P(T<=t) one-tail) was larger than 0.05 for all time points and a significant difference cannot be concluded.

Table S3. *T-test to determine if F. nucleatum treated DLD-1 are migrating faster than untreated DLD-1 at 24h, 48h and 72h post-infection. The table shows a T-test with two-sample assuming unequal variances. "Mean" is the average value for each experimental group; "Variance" is the statistical variance of the data for each experimental group; "Observations" is the number of samples in each experimental group; "Hypothesized Difference" is the selected hypothesized mean difference; "df" is the degree of freedom for the test; "t-Stat" is the t-statistic*; "*P(T<=t) one-tail" is the p-value when using a one-tailed analysis; "T-critical one-tail" is the t-statistic cut-off value when using the one-tailed analysis; "P(T<=t) two-tail" is the p-value when using a two-tailed analysis; "T-critical two-tail" is the t-statistic cut-off value when using the two-tailed analysis.*

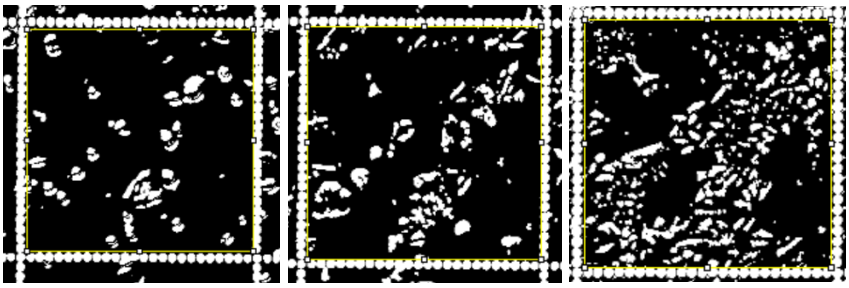
	24h Blank	*24h F. nuc*	*48h Blank*	*48h F. nuc*	*72h Blank*	*72h F .nuc*
Mean	15.95797514	17.36037439	40.0377435	46.90728988	60.27367388	64.38161713
Variance	61.69648063	68.8211546	73.81608558	152.4082685	75.63350372	106.5414833
Observations	4	4	4	4	4	3
Hypothesized Difference	0		0		0	
df	5		5		4	
t-Stat	-0.228188461		-0.913457567		-0.556847626	
P(T<=t) one-tail	0.414269078		0.201449968		0.303657957	
T-critical one-tail	2.015048373		2.015048373		2.131846786	
P(T<=t) two-tail	0.828538157		0.402899937		0.607315915	
T-critical two-tail	2.570581836		2.570581836		2.776445105	

6. Supplementary section 6: Proliferation assay

Images were taken in a phase contrast microscope at 20x and edited at ImageJ as follows; Image type was changed to 8-bit by choosing Image > Type > 8-bit; the bandpass filter was checked by choosing Process > FFT > Bandpass Filter; the image was changed to Black & White by choosing Image > Adjust > Treshold > B&W, and pixels in the image were highlighted so that cells turned white and background turned black; the filter was changed to a minimum radius of 2,0 pixels by choosing Process > Filters > Minimum; and then "rectangle" tool was used to mark the grid to be measured. Edited images from four technical replicates are shown below.

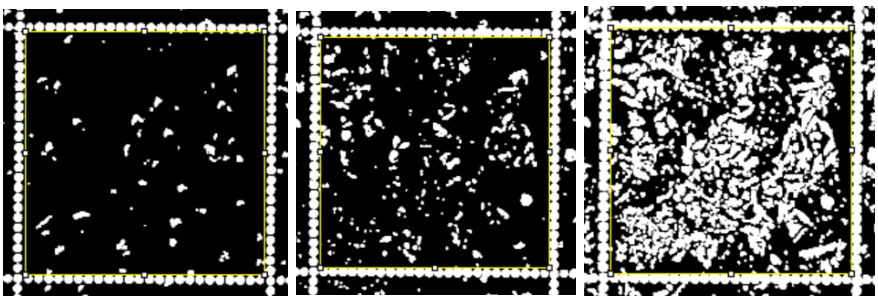
Technical replicate 1

Untreated DLD-1: 0h, 24h, 48h



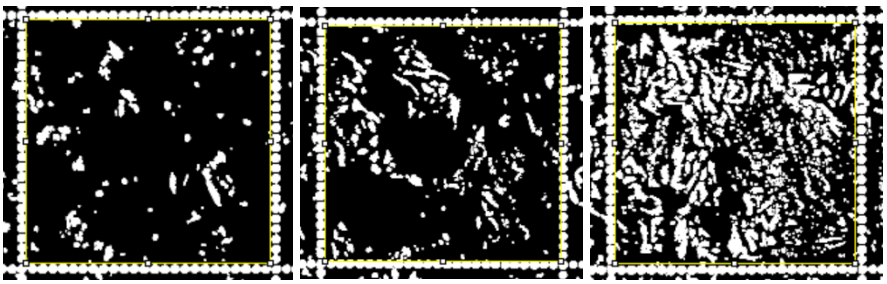
E. coli DH5α treated DLD-1: 0h, 24h, 48h and 72h

F. nucleatum treated DLD-1: 0h, 24h, 48h and 72h

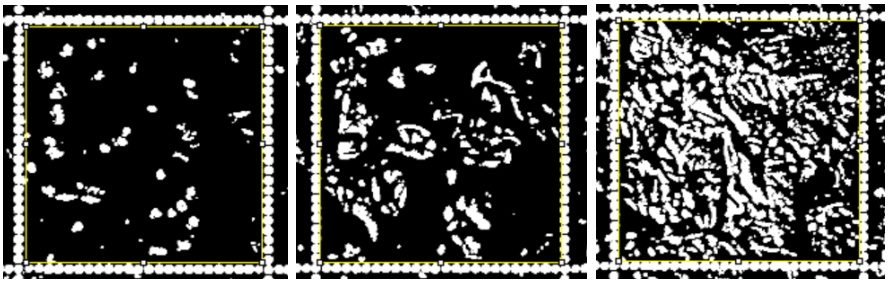


Technical replicate 2

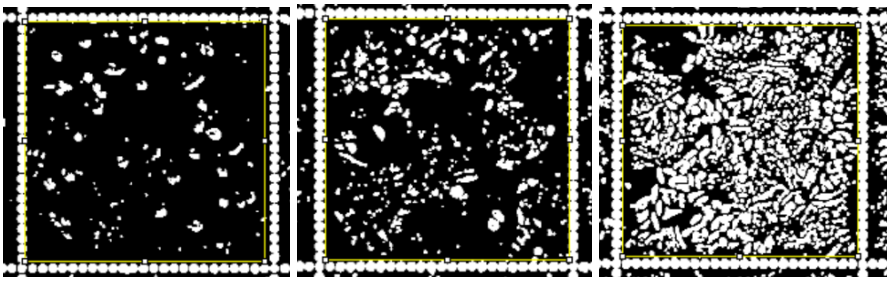
Untreated DLD-1: 0h, 24h, 48h



E. coli DH5α treated DLD-1: 0h, 24h, 48h and 72h

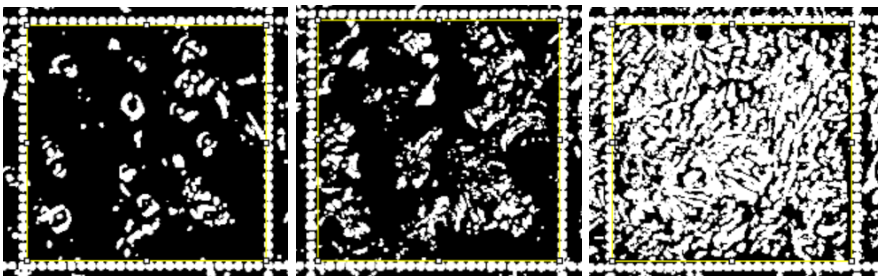


F. nucleatum treated DLD-1: 0h, 24h, 48h and 72h

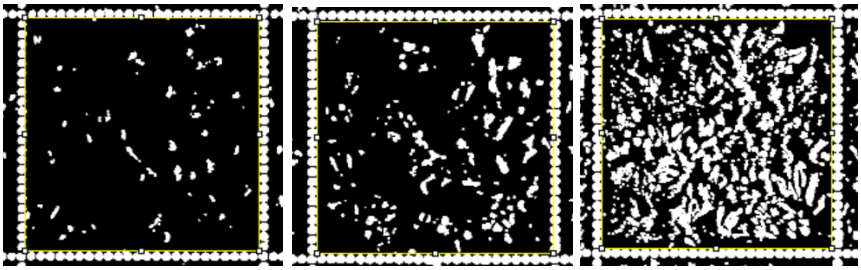


Technical replicate 3

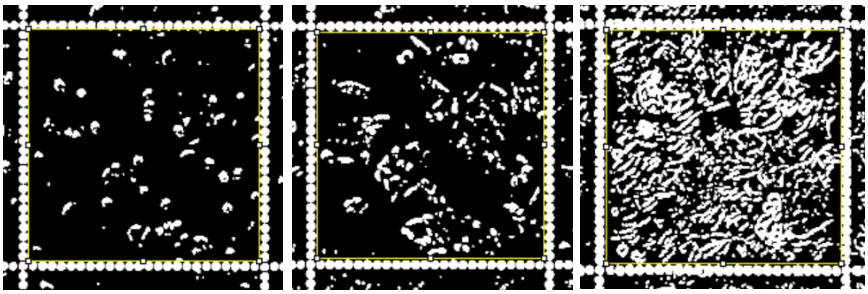
Untreated DLD-1: 0h, 24h, 48h



E. coli DH5α treated DLD-1: 0h, 24h, 48h and 72h

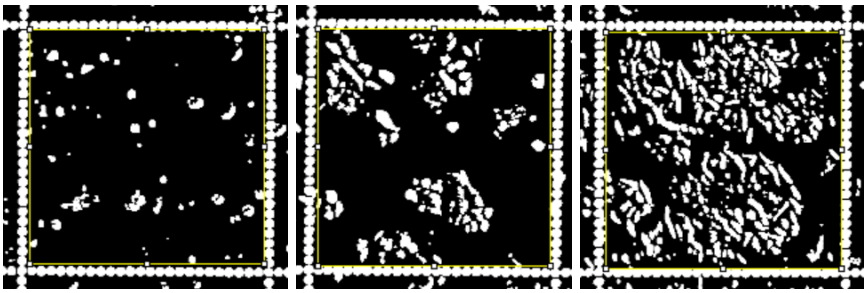


F. nucleatum treated DLD-1: 0h, 24h, 48h and 72h

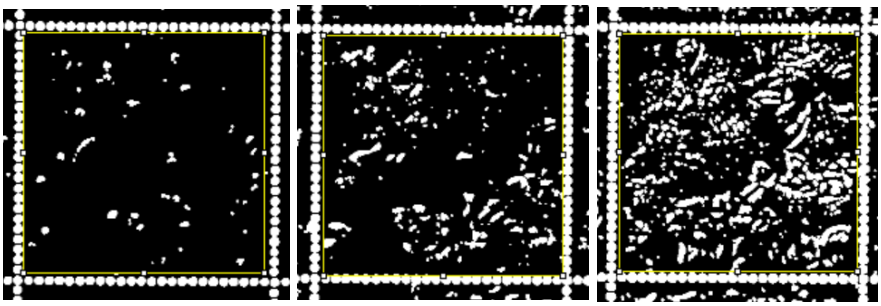


Technical replicate 4

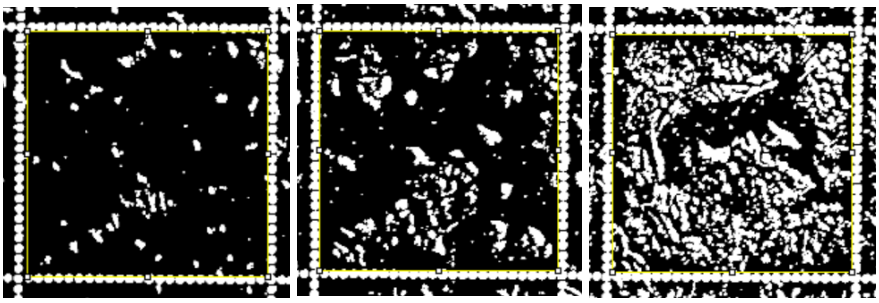
Untreated DLD-1: 0h, 24h, 48h



E. coli DH5α treated DLD-1: 0h, 24h, 48h and 72h



F. nucleatum treated DLD-1: 0h, 24h, 48h and 72h



The area was measured at ImageJ by choosing Analyze > Measure, and cell-free (black) area within the marked area was reported. Results were exported to excel, where percent wound closure of the area was calculated. Raw data, including fold change of the difference between each timepoint and start point and the percent proliferation are presented in Table S4.

Table S4. *Proliferation assay of untreated-, E. coli DH5α treated and F. nucleatum treated DLD-1 at 0, 24 and 48 hours. Table includes size of cell-free area, fold change and percent proliferation.*

	0h	**Area (pixels)**	**Fold change**	**% proliferation**
DLD-1 Untreated	Well 1	262118	1	0
	Well 2	263041	1	0
	Well 3	255430	1	0
	Well 4	267794	1	0
DLD-1 *E. coli*	Well 1	269558	1	0
	Well 2	257934	1	0
	Well 3	273207	1	0
	Well 4	265693	1	0
DLD-1 *F. nucleatum*	Well 1	270005	1	0
	Well 2	260587	1	0
	Well 3	267278	1	0
	Well 4	269147	1	0
	24h			
DLD-1 Untreated	Well 1	253192	0.965946635	3.405336528
	Well 2	254163	0.966248608	3.375139237
	Well 3	230296	0.901601221	9.839877853
	Well 4	243068	0.907667834	9.233216577
DLD-1 *E. coli*	Well 1	254857	0.945462572	5.453742794
	Well 2	243717	0.944881249	5.511875131
	Well 3	254334	0.930920511	6.907948918
	Well 4	259698	0.977436365	2.256363547
DLD-1 F. nucleatum	Well 1	251063	0.929845744	7.01542564
	Well 2	241760	0.927751576	7.224842375
	Well 3	245635	0.919024387	8.097561341
	Well 4	246786	0.916919007	8.308099291
	48h			
DLD-1 Untreated	Well 1	228683	0.872442946	12.75570545
	Well 2	186967	0.710790333	28.92096669
	Well 3	153410	0.600595075	39.9404925
	Well 4	211718	0.790600237	20.93997625
DLD-1 *E. coli*	Well 1	169641	0.629330237	37.06697631
	Well 2	182954	0.709305481	29.06945188
	Well 3	196350	0.718685832	28.13141684
	Well 4	228063	0.858370375	14.16296252
DLD-1 *F. nucleatum*	Well 1	163837	0.606792467	39.32075332
	Well 2	171798	0.659273103	34.07268974
	Well 3	154937	0.579684823	42.03151775
	Well 4	181841	0.675619643	32.43803572

The p-value was calculated in Microsoft Excel by doing a student's t-test for each time point. "T-test: Two-sample Assuming Unequal Variances" was chosen. As the *F. nucleatum*-treated cells are hypothesised to proliferate faster than the untreated cells, the p-value is reported as one-tailed. As shown in Table S5, the p-value (P(T<=t) one-tail) was larger than 0.05 for all time points and a significant difference cannot be concluded.

Table S5. *T-test to determine if F. nucleatum treated DLD-1 proliferate faster than untreated DLD-1 at 24h and 48h post-infection. The table shows a T-test with two-sample assuming unequal variances. "Mean" is the average value for each experimental group; "Variance" is the statistical variance of the data for each experimental group; "Observations" is the number of samples in each experimental group; "Hypothesized Difference" is the selected hypothesized mean difference; "df" is the degree of freedom for the test; "t-Stat" is the t-statistic; "P(T<=t) one-tail" is the p-value when using a one-tailed analysis; "T-critical one-tail" is the t-statistic cut-off value when using the one-tailed analysis; "P(T<=t) two-tail" is the p-value when using a two-tailed analysis; "T-critical two-tail" is the t-statistic cut-off value when using the two-tailed analysis.*

	24h Blank	***24h F. nuc***	***48h Blank***	***48h F. nuc***
Mean	6.463392549	7.661482162	25.6392852	36.9657491
Variance	12.65386443	0.405440698	134.454699	20.0260126
Observations	4	4	4	4
Hypothesized Differe	0		0	
df	3		4	
t-Stat	-0.663069824		-1.8225826	
P(T<=t) one-tail	0.277341424		0.07122456	
T-critical one-tail	2.353363435		2.13184679	
P(T<=t) two-tail	0.554682847		0.14244913	
T-critical two-tail	3.182446305		2.77644511	

7. Supplementary section 7: qPCR Time course assay

7.1. *CCL20* gene expression in *F. nucleatum*-treated DLD-1

Raw data and calculations for each technical replicate of *CCL20* gene expression in untreated-, *F. nucleatum*-treated-, and *E. coli* DH5α-treated DLD-1 cells are presented in Table S6. The table includes calculations of mean of each replicate, normalization to the housekeeping gene *ACTB* (Δsample), normalization to blank (ΔΔsample) and fold change.

Table S6. *Raw data (CT-values) for each technical replicate (Rep1-Rep3) of CCL20 expression in untreated-, F. nucleatum-treated-, and E. coli DH5α-treated DLD-1 cells. Calculations include mean of each replicate, normalization to the housekeeping gene ACTB (Δsample), normalization to blank (ΔΔsample) and fold change are calculated.*

Gene	DLD-1 treatment	Rep 1	Rep 2	Rep 3	Mean	Δsample	ΔΔsample	Fold change
CCL20	***F. nucleatum* 0h**	31.5647	31.8041	32.0695	31.8128	15.2487	0.1659	0.8914
CCL20	***F. nucleatum* 3h**	21.4831	21.6660	21.8480	21.6657	6.0244	-10.5524	1501.7665
CCL20	***F. nucleatum* 6h**	18.5494	18.7986	18.9102	18.7527	2.3262	-13.4023	10826.8254
CCL20	***F. nucleatum* 12h**	19.6004	19.7647	19.9422	19.7691	3.8481	-12.7849	7057.3231
CCL20	***F. nucleatum* 24h**	27.6237	27.7868	28.3226	27.9110	2.0451	-14.1882	18666.4429
CCL20	***E. coli* DH5α 0h**	31.5085	32.0099	31.9289	31.8158	14.9556	-0.1272	1.0922
CCL20	***E. coli* DH5α 3h**	24.9482	25.4249	25.5953	25.3228	9.6925	-6.8843	118.1351
CCL20	***E. coli* DH5α 6h**	20.9268	21.0370	21.2268	21.0635	4.5695	-11.1590	2286.6079
CCL20	***E. coli* DH5α 12h**	21.6946	21.9422	21.8911	21.8427	5.9472	-10.6859	1647.2858
CCL20	***E. coli* DH5α 24h**	24.9788	25.3300	25.0196	25.1095	5.1421	-11.0911	2181.5517
CCL20	**Untreated 0h**	30.6855	31.0876	31.4632	31.0788	15.0828		
CCL20	**Untreated 3h**	31.8137	31.9949	32.1139	31.9742	16.5768		
CCL20	**Untreated 6h**	31.7236	32.3202	32.4454	32.1631	15.7285		
CCL20	**Untreated 12h**	32.3967	32.5425	32.8121	32.5838	16.6330		
CCL20	**Untreated 24h**	32.7786	33.0806	33.2080	33.0224	16.2333		
ACTB	***F. nucleatum* 0h**	16.4235	16.5456	16.7230	16.5641			
ACTB	***F. nucleatum* 3h**	15.5162	15.5300	15.8778	15.6413			
ACTB	***F. nucleatum* 6h**	16.3465	16.3812	16.5521	16.4266			
ACTB	***F. nucleatum* 12h**	15.8094	15.9039	16.0497	15.9210			
ACTB	***F. nucleatum* 24h**	25.7558	25.8094	26.0326	25.8659			
ACTB	***E. coli* DH5α 0h**	16.8161	16.8266	16.9379	16.8602			
ACTB	***E. coli* DH5α 3h**	15.5803	15.6100	15.7005	15.6303			
ACTB	***E. coli* DH5α 6h**	16.4172	16.5007	16.5642	16.4940			
ACTB	***E. coli* DH5α 12h**	15.7831	15.7991	16.1043	15.8955			
ACTB	***E. coli* DH5α 24h**	19.8024	19.8032	20.2964	19.9673			
ACTB	**Untreated 0h**	15.8315	15.9895	16.1670	15.9960			
ACTB	**Untreated 3h**	15.2832	15.4226	15.4863	15.3974			
ACTB	**Untreated 6h**	16.3240	16.5767	16.4031	16.4346			
ACTB	**Untreated 12h**	15.7809	16.0303	16.0412	15.9508			
ACTB	**Untreated 24h**	16.6090	16.9171	16.8412	16.7891			

7.2. *CSF2* gene expression in *F. nucleatum*-treated DLD-1

Raw data for each technical replicate of *CSF2* gene expression in untreated-, *F. nucleatum*-treated-, and *E. coli* DH5α-treated DLD-1 cells are presented in Table S7. The table includes calculations of mean of each replicate, normalization to the housekeeping gene *ACTB* (Δsample), normalization to blank (ΔΔsample) and fold change.

Table S7. *Raw data (CT-values) for each technical replicate (Rep1-Rep3) of CSF2 expression in untreated-, F. nucleatum-treated-, and E. coli DH5α-treated DLD-1 cells. Calculations include mean of each replicate, normalization to the housekeeping gene ACTB (Δsample), normalization to blank (ΔΔsample) and fold change.*

Gene	DLD-1 treatment	Rep 1	Rep 2	Rep 3	Mean	Δsample	ΔΔsample	Fold change
CSF2	***F. nucleatum* 0h**	30.3262	30.4236	30.8246	30.5248	13.9607	-0.2347	1.1766
CSF2	***F. nucleatum* 3h**	23.9808	24.3415	24.2635	24.1953	8.5539	-7.7063	208.8408
CSF2	***F. nucleatum* 6h**	20.5065	20.8733	20.8642	20.7480	4.3214	-10.0668	1072.5102
CSF2	***F. nucleatum* 12h**	21.8937	22.0017	22.2704	22.0552	6.1342	-8.9006	477.9186
CSF2	***F. nucleatum* 24h**	28.7556	28.8705	28.8137	28.8132	2.9473	-10.5524	1501.7709
CSF2	***E. coli* DH5α 0h**	31.0136	31.3531	31.3769	31.2479	14.3877	0.1923	0.8752
CSF2	***E. coli* DH5α 3h**	25.9396	26.0921	26.2286	26.0868	10.4565	-5.8037	55.8590
CSF2	***E. coli* DH5α 6h**	22.4011	22.7313	22.9569	22.6964	6.2024	-8.1858	291.1908
CSF2	***E. coli* DH5α 12h**	23.9183	24.1053	24.0201	24.0146	8.1191	-6.9158	120.7419
CSF2	***E. coli* DH5α 24h**	26.6054	27.0542	27.1774	26.9456	6.9783	-6.5214	91.8650
CSF2	**Untreated 0h**	30.1726	30.2176	30.1838	30.1913	14.1954		
CSF2	**Untreated 3h**	31.3668	31.5787	32.0273	31.6576	16.2602		
CSF2	**Untreated 6h**	30.6700	30.8772	30.9212	30.8228	14.3882		
CSF2	**Untreated 12h**	30.8170	30.8570	31.2829	30.9856	15.0349		
CSF2	**Untreated 24h**	30.0817	30.2641	30.5209	30.2889	13.4998		
ACTB	***F. nucleatum* 0h**	16.4235	16.5456	16.7230	16.5641			
ACTB	***F. nucleatum* 3h**	15.5162	15.5300	15.8778	15.6413			
ACTB	***F. nucleatum* 6h**	16.3465	16.3812	16.5521	16.4266			
ACTB	***F. nucleatum* 12h**	15.8094	15.9039	16.0497	15.9210			
ACTB	***F. nucleatum* 24h**	25.7558	25.8094	26.0326	25.8659			
ACTB	***E. coli* DH5α 0h**	16.8161	16.8266	16.9379	16.8602			
ACTB	***E. coli* DH5α 3h**	15.5803	15.6100	15.7005	15.6303			
ACTB	***E. coli* DH5α 6h**	16.4172	16.5007	16.5642	16.4940			
ACTB	***E. coli* DH5α 12h**	15.7831	15.7991	16.1043	15.8955			
ACTB	***E. coli* DH5α 24h**	19.8024	19.8032	20.2964	19.9673			
ACTB	**Untreated 0h**	15.8315	15.9895	16.1670	15.9960			
ACTB	**Untreated 3h**	15.2832	15.4226	15.4863	15.3974			
ACTB	**Untreated 6h**	16.3240	16.5767	16.4031	16.4346			
ACTB	**Untreated 12h**	15.7809	16.0303	16.0412	15.9508			
ACTB	**Untreated 24h**	16.6090	16.9171	16.8412	16.7891			

8. Supplementary section 8: Quantikine ELISA assay

8.1. Time course of CCL20 release in *F. nucleatum*-treated DLD-1

Raw data and calculations of CCL20 release in untreated-, *F. nucleatum*-treated-, and *E. coli* DH5α-treated DLD-1 cells are presented in Table S8 (biological replicate 1) and Table S9 (biological replicate 2). Results were calculated using Microsoft Excel. Readings at 540 nm were subtracted from the readings at 450 nm. The mean of each duplicate was calculated, and the mean of the zero standard optical density was subtracted from each duplicate.

Table S8. *Biological replicate 1 for CCL20 release at different time points in untreated-, F. nucleatum-treated-, and E. coli DH5α-treated DLD-1 cells. Readings at 540 nm were subtracted from readings at 450 nm. The mean of each duplicate was calculated, and the mean of the zero standard optical density was subtracted from each duplicate. OD = Optical density.*

	450 nm Rep1 (OD)	540 nm Rep1 (OD)	450 nm Rep2 (OD)	540 nm Rep2 (OD)	450-540 nm Rep1 (OD)	450-540 nm Rep2 (OD)	Mean (OD)	Normalize to zero standard
Standard 0 pg/mL	0.077	0.033	0.078	0.032	0.044	0.046	0.045	0
Standard 7,8 pg/mL	0.083	0.031	0.089	0.031	0.052	0.058	0.055	0.01
Standard 15,6 pg/mL	0.092	0.034	0.094	0.035	0.058	0.059	0.0585	0.0135
Standard 31,3 pg/mL	0.113	0.035	0.104	0.035	0.078	0.069	0.0735	0.0285
Standard 62,5 pg/mL	0.144	0.035	0.164	0.04	0.109	0.124	0.1165	0.0715
Standard 125 pg/mL	0.233	0.037	0.241	0.053	0.196	0.188	0.192	0.147
Standard 250 pg/mL	0.391	0.039	0.405	0.038	0.352	0.367	0.3595	0.3145
Standard 500 pg/mL	0.689	0.066	0.6	0.042	0.623	0.558	0.5905	0.5455
***F. nucleatum* 0h**	0.08	0.033	0.069	0.032	0.047	0.037	0.042	-0.003
***F. nucleatum* 6h**	0.48	0.044	0.495	0.043	0.436	0.452	0.444	0.399
***F. nucleatum* 12h**	1.862	0.074	1.853	0.077	1.788	1.776	1.782	1.737
***F. nucleatum* 24h**	1.184	0.067	1.181	0.063	1.117	1.118	1.1175	1.0725
***E. coli* DH5α 0h**	0.083	0.031	0.078	0.033	0.052	0.045	0.0485	0.0035
***E. coli* DH5α 6h**	0.114	0.039	0.114	0.04	0.075	0.074	0.0745	0.0295
***E. coli* DH5α 12h**	0.266	0.04	0.266	0.041	0.226	0.225	0.2255	0.1805
***E. coli* DH5α 24h**	0.153	0.036	0.139	0.037	0.117	0.102	0.1095	0.0645
Untreated 0h	0.091	0.034	0.083	0.035	0.057	0.048	0.0525	0.0075
Untreated 6h	0.076	0.038	0.077	0.038	0.038	0.039	0.0385	-0.0065
Untreated 12h	0.079	0.035	0.075	0.035	0.044	0.04	0.042	-0.003
Untreated 24h	0.083	0.04	0.09	0.037	0.043	0.053	0.048	0.003

Table S9. *Biological replicate 2 for CCL20 release at different time points in untreated-, F. nucleatum-treated-, and E. coli DH5α-treated DLD-1 cells. Readings at 540 nm were subtracted from readings at 450 nm. The mean of each duplicate was calculated, and the mean of the zero standard optical density was subtracted from each duplicate. OD = Optical density.*

	450 nm Rep1 (OD)	540 nm Rep1 (OD)	450 nm Rep2 (OD)	540 nm Rep2 (OD)	450-540 nm Rep1 (OD)	450-540 nm Rep2 (OD)	Mean (OD)	Normalize to zero standard
Standard 0 pg/mL	0.089	0.04	0.08	0.037	0.049	0.043	0.046	0
Standard 7,8 pg/mL	0.104	0.037	0.113	0.037	0.067	0.076	0.0715	0.0255
Standard 15,6 pg/mL	0.113	0.039	0.118	0.036	0.074	0.082	0.078	0.032
Standard 31,3 pg/mL	0.173	0.038	0.168	0.038	0.135	0.13	0.1325	0.0865
Standard 62,5 pg/mL	0.25	0.042	0.238	0.039	0.208	0.199	0.2035	0.1575
Standard 125 pg/mL	0.486	0.042	0.47	0.042	0.444	0.428	0.436	0.39
Standard 250 pg/mL	0.863	0.048	0.854	0.048	0.815	0.806	0.8105	0.7645
Standard 500 pg/mL	1.601	0.059	1.576	0.058	1.542	1.518	1.53	1.484
***F. nucleatum* 0h**	0.08	0.032	0.072	0.035	0.048	0.037	0.0425	-0.0035
***F. nucleatum* 6h**	0.587	0.039	0.593	0.047	0.548	0.546	0.547	0.501
***F. nucleatum* 12h**	1.423	0.047	1.387	0.053	1.376	1.334	1.355	1.309
***F. nucleatum* 24h**	0.984	0.033	0.839	0.032	0.951	0.807	0.879	0.833
***E. coli* DH5α 0h**	0.105	0.036	0.07	0.034	0.069	0.036	0.0525	0.0065
***E. coli* DH5α 6h**	0.4	0.041	0.407	0.044	0.359	0.363	0.361	0.315
***E. coli* DH5α 12h**	1.475	0.031	1.492	0.034	1.444	1.458	1.451	1.405
***E. coli* DH5α 24h**	0.448	0.039	0.431	0.031	0.409	0.4	0.4045	0.3585
Untreated 0h	0.079	0.032	0.072	0.032	0.047	0.04	0.0435	-0.0025
Untreated 6h	0.069	0.031	0.063	0.032	0.038	0.031	0.0345	-0.0115
Untreated 12h	0.071	0.036	0.072	0.035	0.035	0.037	0.036	-0.01
Untreated 24h	0.072	0.038	0.073	0.036	0.034	0.037	0.0355	-0.0105

A standard curve was created by plotting the mean optical density for each standard against the concentration, and a best fit curve was made through the points on the graph by checking the linear trendline (Figure S9). Both standard curves showed a high squared correlation coefficient (R^2 greater than 0.99), indicating a high degree of accuracy.

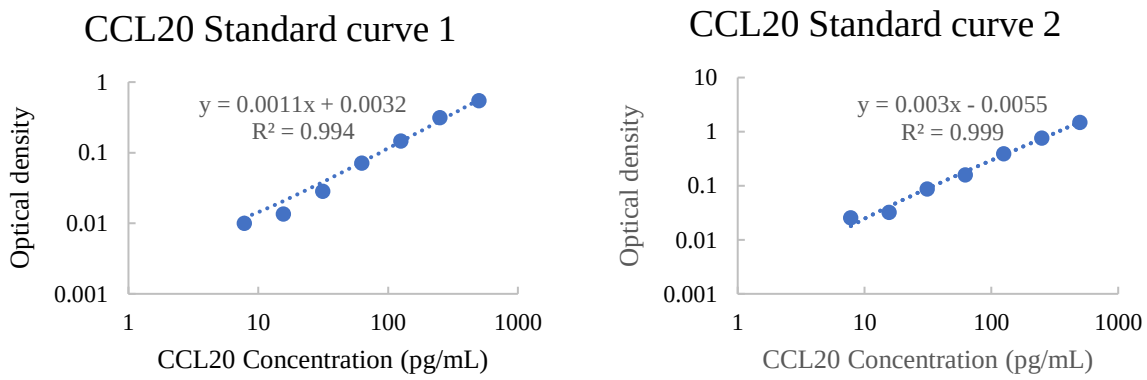


Figure S9. *Standard curve for biological replicate 1 (left) and biological replicate 2 (right) showing a linear relation between the mean optical density (OD) for each standard against the concentration in pg/mL (R^2 = 0,994 and R^2 = 0,999, respectively). y = ax + b, where y = optical density, x = concentration, a = slope, b = y-intercept.*

The equation of the trendline was further used to calculate the concentration of CCL20 in each biological replicate, and the mean of the biological replicates was calculated (Table S10).

Table S10. *Concentration of CCL20 in pg/mL in untreated-, F. nucleatum-treated- and E. coli DH5α-treated DLD-1 cells at different time points post-infection.*

Time (hours)	*F. nuc* p1 (pg/mL)	*E. coli* p1 (pg/mL)	Blank p1 (pg/mL)	*F. nuc* p2 (pg/mL)	*E. coli* p2 (pg/mL)	Blank p2 (pg/mL)	Mean *F. nuc* (pg/mL)	Mean *E. coli* (pg/mL)	Mean Blank (pg/mL)
0	0.6667	4.0000	1.0000	-4.0000	1.9091	5.5455	-1.6667	2.9545	3.2727
6	168.8333	106.8333	-2.0000	361.4545	25.5455	-7.1818	265.1439	66.1894	-4.5909
12	438.1667	372.1667	-1.5000	1577.8182	162.8182	-4.0000	1007.9924	267.4924	-2.7500
24	272.8333	118.0000	-1.6667	973.7273	57.3636	1.4545	623.2803	87.6818	-0.1061

The p-value was calculated in Microsoft Excel by doing a student's t-test: Two-sample Assuming Unequal Variances. As the *F. nucleatum*-treated cells are hypothesised to release more CCL20 than the untreated cells, the p-value is reported as one-tailed. As shown in Table S11, the p-value (P(T<=t) one-tail) was larger than 0.05 for all time points and a significant difference cannot be concluded.

Table S11. *T-test to determine if CCL20 release is significantly higher in F. nucleatum treated DLD-1 compared to untreated DLD-1. The table shows a T-test two-sample assuming unequal variances. "Mean" is the average value for each experimental group; "Variance" is the statistical variance of the data for each experimental group; "Observations" is the number of samples in each experimental group; "Hypothesized Difference" is the selected hypothesized mean difference; "df" is the degree of freedom for the test; "t-Stat" is the t-statistic; "P(T<=t) one-tail" is the p-value when using a one-tailed analysis; "T-critical one-tail" is the t-statistic cut-off value when using the one-tailed analysis; "P(T<=t) two-tail" is the p-value when using a two-tailed analysis; "T-critical two-tail" is the t-statistic cut-off value when using the two-tailed analysis.*

	6h F. nuc	*6h Blank*	*12h F. nuc*	*12h Blank*	*24h F. nuc*	*24h Blank*
Mean	265.14394	-4.590909	1007.99242	-2.75	623.280303	-0.106
Variance	18551.466	13.42562	649402.788	3.125	245626.157	4.871
Observations	2	2	2	2	2	2
Hypothesized Difference	0		0		0	
df	1		1		1	
t-Stat	2.7996636		1.77377028		1.77881459	
P(T<=t) one-tail	0.1092		0.16340576		0.16301934	
T-critical one-tail	6.3137515		6.31375151		6.31375151	
P(T<=t) two-tail	0.2184001		0.32681152		0.32603868	
T-critical two-tail	12.706205		12.7062047		12.7062047	

8.2. Concentration assay of CCL20 release in *F. nucleatum*-treated DLD-1

Raw data and calculations of CCL20 release in untreated-, *F. nucleatum*-treated-, and *E. coli* DH5α-treated DLD-1 cells are presented in Table S12 (biological replicate 1) and Table S13 (biological replicate 2). Results were calculated using Microsoft Excel. Readings at 540 nm were subtracted from the readings at 450 nm. The mean of each duplicate was calculated, and the mean of the zero standard optical density was subtracted from each duplicate.

Table S12. *Biological replicate 1. CCL20 release in different concentrations of F. nucleatum-treated- and E. coli DH5α-treated DLD-1 cells. Readings at 540 nm were subtracted from readings at 450 nm. The mean of each duplicate was calculated, and the mean of the zero standard optical density was subtracted from each duplicate. OD = Optical density.*

	450 nm Rep1 (OD)	540 nm Rep1 (OD)	450 nm Rep2 (OD)	540 nm Rep2 (OD)	450-540 nm Rep1 (OD)	450-540 nm Rep2 (OD)	Mean (OD)	Normalize to zero standard
***F. nucleatum* 0 MOI**	0.096	0.039	0.083	0.042	0.057	0.041	0.049	0.004
***F. nucleatum* 0,1 MOI**	0.087	0.033	0.078	0.035	0.054	0.043	0.0485	0.0035
***F. nucleatum* 1 MOI**	0.077	0.034	0.078	0.037	0.043	0.041	0.042	-0.003
***F. nucleatum* 10 MOI**	0.12	0.036	0.123	0.038	0.084	0.085	0.0845	0.0395
***F. nucleatum* 100 MOI**	0.731	0.045	0.771	0.053	0.686	0.718	0.702	0.657
***E. coli* DH5α 0 MOI**	0.083	0.034	0.078	0.038	0.049	0.04	0.0445	-0.0005
***E. coli* DH5α 0,1 MOI**	0.073	0.033	0.072	0.031	0.04	0.041	0.0405	-0.0045
***E. coli* DH5α 1 MOI**	0.077	0.033	0.082	0.034	0.044	0.048	0.046	0.001
***E. coli* DH5α 10 MOI**	0.096	0.034	0.104	0.034	0.062	0.07	0.066	0.021
***E. coli* DH5α 100 MOI**	0.175	0.035	0.17	0.034	0.14	0.136	0.138	0.093

Table S13. *Biological replicate 2. CCL20 release in different concentrations of F. nucleatum-treated- and E. coli DH5α-treated DLD-1 cells. Readings at 540 nm were subtracted from readings at 450 nm. The mean of each duplicate was calculated, and the mean of the zero standard optical density was subtracted from each duplicate. OD = Optical density.*

	450 nm Rep1 (OD)	540 nm Rep1 (OD)	450 nm Rep2 (OD)	540 nm Rep2 (OD)	450-540 nm Rep1 (OD)	450-540 nm Rep2 (OD)	Mean (OD)	Normalize to zero standard
***F. nucleatum* 0 MOI**	0.096	0.039	0.083	0.042	0.057	0.041	0.049	0.004
***F. nucleatum* 0,1 MOI**	0.088	0.033	0.089	0.033	0.055	0.056	0.0555	0.0105
***F. nucleatum* 1 MOI**	0.103	0.04	0.104	0.035	0.063	0.069	0.066	0.021
***F. nucleatum* 10 MOI**	0.123	0.036	0.125	0.035	0.087	0.09	0.0885	0.0435
***F. nucleatum* 100 MOI**	0.269	0.037	0.263	0.035	0.232	0.228	0.23	0.185
***E. coli* DH5α 0 MOI**	0.083	0.034	0.078	0.038	0.049	0.04	0.0445	-0.0005
***E. coli* DH5α 0,1 MOI**	0.107	0.034	0.1	0.032	0.073	0.068	0.0705	0.0255
***E. coli* DH5α 1 MOI**	0.109	0.038	0.108	0.035	0.071	0.073	0.072	0.027
***E. coli* DH5α 10 MOI**	0.137	0.036	0.133	0.036	0.101	0.097	0.099	0.054
***E. coli* DH5α 100 MOI**	0.17	0.036	0.153	0.036	0.134	0.117	0.1255	0.0805

The equation of the trendline in standard curve 1 and standard curve 2 (Figure S9) was used to calculate the concentration of CCL20 in biological replicate 1 and biological replicate 2, respectively. Concentrations of the respective gene are shown in Table S14.

Table S14. *Concentration of CCL20 in pg/mL in different concentrations of F. nucleatum-treated- and E. coli DH5α-treated DLD-1 cells. "p1" indicates biological replicate 1 and "p2" indicates biological replicate 2.*

MOI	*F. nuc* p1 (pg/mL)	*E. coli* p1 (pg/mL)	*F. nuc* p2 (pg/mL)	*E. coli* p2 (pg/mL)	Mean *F. nuc* (pg/mL)	Mean *E. coli* (pg/mL)
0	0.6	0.6	-1.6	-1.6	-0.5	-0.5
0.1	0.2	-7.1	6.5	20.2	3.4	6.5
1	-5.7	-2.1	16.1	21.5	5.2	9.7
10	32.9	16.1	36.5	46.1	34.7	31.1
100	220.5	31.5	165.6	70.2	193.1	50.8

The p-value was calculated in Microsoft Excel by doing a student's t-test: Two-sample Assuming Unequal Variances. As the *F. nucleatum*-treated cells were hypothesised to release more CCL20 than the untreated cells, the p-value was reported as one-tailed. As shown in Table S15, the p-value (P(T<=t) one-tail) was less than 0.05 for 10 MOI (P = 0.002) and 100 MOI (P = 0.045) *F. nucleatum* compared with 0 MOI *F. nucleatum* which means that the null hypothesis that there's no difference between the means were rejected and a significant difference does exist.

Table S15. *T-test to determine if CCL20 release is significantly higher in F. nucleatum treated DLD-1 compared to untreated DLD-1. The table shows a T-test two-sample assuming unequal variances. Significant differences (P < 0,05) are marked with a star (*). "Mean" is the average value for each experimental group; "Variance" is the statistical variance of the data for each experimental group; "Observations" is the number of samples in each experimental group; "Hypothesized Difference" is the selected hypothesized mean difference; "df" is the degree of freedom for the test; "t-Stat" is the t-statistic; "P(T<=t) one-tail" is the p-value when using a one-tailed analysis; "T-critical one-tail" is the t-statistic cut-off value when using the one-tailed analysis; "P(T<=t) two-tail" is the p-value when using a two-tailed analysis; "T-critical two-tail" is the t-statistic cut-off value when using the two-tailed analysis.*

	F. culeatum 0 MOI	*F. culeatum 0,1 MOI*	*F. culeatum 0 MOI*	*F. culeatum 1 MOI*	*F. culeatum 0 MOI*	*F. culeatum 10 MOI*	*F. culeatum 0 MOI*	*F. culeatum 100 MOI*
Mean	-0.5	3.35	-0.5	5.2	-0.5	34.7	-0.5	193.05
Variance	2.42	19.845	2.42	237.62	2.42	6.48	2.42	1507.005
Observations	2	2	2	2	2	2	2	2
Hypothesized Differe	0		0		0		0	
df	1		1		2		1	
t-Stat	-1.153889915		-0.52029307		-16.6864003		-7.0453473	
P(T<=t) one-tail	0.227296154		0.347290845		0.00178613		0.04488037	
T-critical one-tail	6.313751515		6.313751515		2.91998558		6.31375151	
P(T<=t) two-tail	0.454592309		0.694581691		0.00357226		0.08976073	
T-critical two-tail	12.70620474		12.70620474		4.30265273		12.7062047	

9. Supplementary section 9: EBV-miRNA transfection

9.1. qPCR results of SW620 transfected with miR-BART10-3p

Raw data of *MAT2B* and *ELL2* in miR-BART10-3p-transfected SW620 cells are presented in Table S16. The table includes CT values of 6 technical replicates for *MAT2B, ELL2* and *ACTB* transfected with miR-BART10-3p (EBV) or negative miRNA (Negative) in three biological replicates (Rep1-Rep3).

Table S16. *qPCR results of MAT2B and ELL2 gene expression in SW620 cells transfected with miR-BART10-3p (EBV) or negative miRNA (Negative). Table includes 6 technical replicates (MAT2B 1-6) and 3 biological replicates (Rep-1-Rep3). ACTB was used as housekeeping gene.*

EBV	CT Rep1	CT Rep2	CT Rep3	Negative	CT Rep1	CT Rep2	CT Rep3
MAT2B 1	23.3452	22.62575	24.9553	**MAT2B 1**	20.6477	20.3940067	23.9680843
MAT2B 2	22.8449	22.37274	24.942	**MAT2B 2**	20.5760	20.327404	23.8854141
MAT2B 3	23.4726	22.17135	24.694	**MAT2B 3**	19.8681	20.426302	23.6661453
MAT2B 4	22.8047	21.99561	24.6594	**MAT2B 4**	19.8509	20.1876488	23.6497574
MAT2B 5	22.8186	21.55392	24.6349	**MAT2B 5**	19.6323	20.0105438	24.0144291
MAT2B 6	22.8073	21.24316	24.5397	**MAT2B 6**	19.6751	20.1523724	23.9919643
ELL2 1	24.1803	25.86661	30.1697	**ELL2 1**	23.4515	24.9424152	26.9779854
ELL2 2	24.0702	25.5661	27.4466	**ELL2 2**	22.9492	24.6895542	26.9392853
ELL2 3	24.5121	24.95811	27.4291	**ELL2 3**	23.2366	24.5800114	26.3974094
ELL2 4	23.9595	25.37561	26.9599	**ELL2 4**	22.7903	24.6518707	26.7063904
ELL2 5	23.9966	25.17935	27.0568	**ELL2 5**	23.1546	24.5849953	26.9479313
ELL2 6	23.7535	25.17214	26.9749	**ELL2 6**	22.5613	24.6700249	26.9120407
ACTB 1	10.5764	17.9287	19.0944	**ACTB 1**	10.4266	16.9698181	19.173975
ACTB 2	9.6505	18.21277	18.9006	**ACTB 2**	10.0617	17.8907948	18.9276905
ACTB 3	9.8389	18.21183	18.9798	**ACTB 3**	10.1161	17.5389271	18.634037
ACTB 4	9.7231	17.58966	18.5732	**ACTB 4**	10.1489	17.2087631	18.5461388
ACTB 5	9.2601	17.41597	18.7605	**ACTB 5**	5.1642	17.7721405	19.0499954
ACTB 6	9.3035	17.65985	18.591	**ACTB 6**	5.1600	18.2437115	18.9715633

Further calculations of technical replicates in Table S16 are presented in Table S17. Calculations include mean of each replicate, normalization to the housekeeping gene *ACTB* (Δsample), normalization to the negative miRNA (ΔΔsample), and the fold change.

The p-value was calculated in Microsoft Excel by doing a student's t-test: Two-sample Assuming Unequal Variances. As the miR-BART10-3p-transfected cells were hypothesised to decrease expression of *MAT2B* and *ELL2* compared to cells transfected with negative miRNA, the p-value was reported as one-tailed. As shown in Table S18, the p-value (P(T<=t) one-tail) was less than 0.05 for *MAT2B* (P = 0.016) and *ELL2* (P = 0.028) compared to control, which

means that the null hypothesis that there's no difference between the means were rejected and a significant difference does exist.

Table S17. *Calculations of each biological replicate of MAT2B and ELL2 in SW620 transfected with miR-BART10-3p (EBV) or negative miRNA (Neg). Table includes mean of each technical replicate, normalization to the housekeeping gene ACTB (Δsample), normalization to the negative miRNA (ΔΔsample), and the fold change.*

	Genes	Mean EBV	Mean Neg	ΔEBV	ΔNeg	ΔΔEBV	Fold change
Rep1	**MAT2B**	23.0156	20.0417	13.2902	9.8534	3.4368	0.0923
	ELL2	24.0787	23.0239	14.3533	12.8356	1.5177	0.3492
	ACTB	9.7254	10.1883				
Rep2	**MAT2B**	21.9938	20.2497	4.1573	2.6457	1.5116	0.3507
	ELL2	25.3530	24.6865	7.5165	7.0825	0.4341	0.7402
	ACTB	17.8365	17.6040				
Rep3	**MAT2B**	24.7376	23.8626	5.9209	4.9787	0.9422	0.5204
	ELL2	27.6728	26.8135	8.8562	7.9296	0.9266	0.5261
	ACTB	18.8166	18.8839				

Table S18. *T-test to determine if MAT2B and ELL2 expression are significantly decreased after miR-BART10-3p-transfection of SW620 compared to negative miRNA. The table shows a T-test two-sample assuming unequal variances. "Mean" is the average value for each experimental group; "Variance" is the statistical variance of the data for each experimental group; "Observations" is the number of samples in each experimental group; "Hypothesized Difference" is the selected hypothesized mean difference; "df" is the degree of freedom for the test; "t-Stat" is the t-statistic; "P(T<=t) one-tail" is the p-value when using a one-tailed analysis; "T-critical one-tail" is the t-statistic cut-off value when using the one-tailed analysis; "P(T<=t) two-tail" is the p-value when using a two-tailed analysis; "T-critical two-tail" is the t-statistic cut-off value when using the two-tailed analysis.*

	MAT2B	***Control***	***ELL2***	***Control***
Mean	0.3211672	1	0.5385	1
Variance	0.0464694	0	0.03832	0
Observations	3	2	3	2
Hypothesized Difference	0		0	
df	2		2	
t-Stat	-5.454308		-4.0833	
P(T<=t) one-tail	0.0160044		0.02753	
T-critical one-tail	2.9199856		2.91999	
P(T<=t) two-tail	0.0320089		0.05507	
T-critical two-tail	4.3026527		4.30265	

9.2. Luciferase results of SW620 co-transfected with vector and miRNA

Raw data from the Dual Luciferase Assay of *MAT2B* expression in SW620 cells co-transfected with miR-BART10-3p and 3'UTR reporter vector are presented in Table S19. The table includes luciferase signal (*Renilla* and Firefly activity) of four technical replicates for *MATB* 3'UTR Reporter vector transfected with miR-BART10-3p (EBV) or negative miRNA (Neg) in three biological replicates (Rep1-Rep3). The control is an empty psiCHECK^{TM}-2 vector.

Table S19. *Renilla and Firefly luciferase activity from Dual Luciferase assay of MAT2B after being co-transfected with miR-BART10-3p (EBV) or negative miRNA (Neg) in SW620 cells. The table includes 3 biological replicates (Rep1-Rep3) with 4 technical replicates.*

		EBV1	Neg1	EBV2	Neg2	EBV3	Neg3	EBV4	Neg4
Rep1	**Blank**	5085	6748	6614	7535	7340	8346	8122	8549
Renilla	**MAT2B**	66759	97557	68891	105668	76248	118624	90349	125670
	Control	25357	25588	33678	30691	37954	27727	46010	34325
Rep1	**Blank**	4020	4527	4902	5018	4978	5364	5403	5254
Firefly	**MAT2B**	39648	39515	41632	41745	43653	43000	46874	46288
	psiCHECK	12307	12437	16088	13919	16296	12537	18443	14124
Rep2	**Blank**	7486	10233	10291	11469	11528	12409	11862	12252
Renilla	**MAT2B**	87301	124830	108057	140437	121805	150513	135282	171462
	Control	38976	9569	42921	11060	45589	12118	44074	12222
Rep2	**Blank**	5238	6255	7114	6869	7502	6964	7219	6733
Firefly	**MAT2B**	46113	47657	59512	51260	60305	51828	62864	53128
	Control	12307	15890	7238	17242	8349	18527	8336	17132
Rep3	**Blank**	58	68	77	99	96	108	64	6
Renilla	**MAT2B**	939	1916	1419	1988	1535	2414	1545	478
	Control	1266	1804	1946	2297	2326	2157	763	2877
Rep3	**Blank**	0	0	0	0	0	0	0	0
Firefly	**MAT2B**	356	396	410	466	456	531	474	87
	Control	286	459	524	604	516	547	165	574

Raw data from the LightSwitch Luciferase Assay of *MAT2B* gene expression in SW620 cells co-transfected with miR-BART10-3p and 3'UTR reporter vector are presented in Table S20. The table includes luciferase signal of four technical replicates for *MATB* 3'UTR Reporter vector transfected with miR-BART10-3p (EBV) or negative miRNA (Neg) in three biological replicates (Rep1-Rep3). The control is a Random 3'UTR control vector (R04_3UTR).

Table S20. *Luciferase signal from LightSwitch Luciferase assay of MAT2B after being co-transfected with miR-BART10-3p (EBV) or negative miRNA (Neg) in SW620 cells. The table includes 3 biological replicates (Rep1-Rep3) with 4 technical replicates.*

		EBV1	Neg1	EBV2	Neg2	EBV3	Neg3	EBV4	Neg4
Rep1	**Blank**	489	518	543	489	478	522	481	501
	MAT2B	1701	7797	2003	7190	1896	7515	1965	7818
	Control	1131	1500	1055	1426	1282	1809	802	993
Rep2	**Blank**	1321	1438	1765	1809	1549	1822	1291	1482
	MAT2B	26966	67789	30350	37143	24239	55432	13785	34330
	Control	5294	6222	10238	5085	5105	7010	1612	5743
Rep3	**Blank**	2052	2051	1823	2212	2262	2316	2135	2240
	MAT2B	12679	32717	15561	31458	12772	36978	14101	34997
	Control	3863	4200	4429	4329	4465	4336	5499	4759

Raw data from the LightSwitch Luciferase Assay of *ELL2* expression in SW620 cells co-transfected with miR-BART10-3p and 3'UTR reporter vector are presented in Table S21. The table includes luciferase signal of four technical replicates for *ELL2* 3'UTR Reporter vector transfected with miR-BART10-3p (EBV) or negative miRNA (Neg) in three biological replicates (Rep1-Rep3). The control is a Random 3'UTR control vector (R04_3UTR).

Table S21. *Luciferase signal from LightSwitch Luciferase assay of ELL2 after being co-transfected with miR-BART10-3p (EBV) or negative miRNA (Neg) in SW620 cells. The table includes 3 biological replicates (Rep1-Rep3) with 4 technical replicates.*

		EBV1	Neg1	EBV2	Neg2	EBV3	Neg3	EBV4	Neg4
Rep1	**Blank**	4039	5349	2453	5415	1976	4862	1278	4987
	ELL2	642	719	524	701	547	697	536	571
	Control	1739	1836	2635	2379	2484	2384	2945	2109
Rep2	**Blank**	1796	5609	4386	5425	1323	5274	1473	4708
	ELL2	525	634	685	649	489	526	691	575
	Control	2951	2472	2223	2155	1995	2158	2211	2551
Rep3	**Blank**	1893	5434	2344	5192	1546	4826	4628	4863
	ELL2	546	551	577	631	627	661	570	777
	Control	2195	2516	2510	2891	1957	1685	2701	2013

The degree of knockdown was calculated by first subtracting the mean blank from each sample. For the Dual Luciferase Assay, the *Renilla* luciferase activity was divided by the firefly luciferase activity. For both assays, the ratio of miR-BART10-3p divided by negative miRNA was calculated and normalized with the control. Results are shown in Table S22.

Table S22. *Fold change and standard deviation of MAT2B and ELL2 gene expression in SW620 co-transfected with miR-BART10-3p compared to negative miRNA.*

Assay	Genes	Mean Fold change	St. deviation
LightSwitch Luciferase	*ELL2*	0.864544674	0.025869693
	MAT2B	0.392634579	0.086226058
Dual Luciferase	*MAT2B*	0.679043438	0.013777863

The p-value was calculated in Microsoft Excel by doing a T-test: Two-sample Assuming Unequal Variances. As the miR-BART10-3p-transfected cells were hypothesised to decrease expression of *MAT2B* and *ELL2* compared to cells transfected with negative miRNA, the p-value was reported as one-tailed. As shown in Table S23, the p-value (P(T<=t) one-tail) was less than 0.05 for *MAT2B* compared to control in the LightSwitch Luciferase Assay (P = 0.004), as well as in the Dual Luciferase Assay (P = 0.0005), which means that the null hypothesis that there's no difference between the means were rejected and a significant difference does exist. The p-value for *ELL2* was, however, greater than 0.05 (P = 0.1026), and the downregulation was not significant.

Table S23. *T-test to determine if MAT2B and ELL2 expression are significantly decreased after miR-BART10-3p-transfection of SW620 compared to negative miRNA. The table shows a T-test two-sample assuming unequal variances. "Mean" is the average value for each experimental group; "Variance" is the statistical variance of the data for each experimental group; "Observations" is the number of samples in each experimental group; "Hypothesized Difference" is the selected hypothesized mean difference; "df" is the degree of freedom for the test; "t-Stat" is the t-statistic; "P(T<=t) one-tail" is the p-value when using a one-tailed analysis; "T-critical one-tail" is the t-statistic cut-off value when using the one-tailed analysis; "P(T<=t) two-tail" is the p-value when using a two-tailed analysis; "T-critical two-tail" is the t-statistic cut-off value when using the two-tailed analysis.*

	LightSwitch Luciferase Assay				Dual Luciferase Assay	
	MAT2B	*Control*	*ELL2*	*Control*	*MAT2B*	*Control*
Mean	0.392634579	1	0.86454467	1	0.67904344	1
Variance	0.009814808	0	0.01601793	0	0.00028474	0
Observations	3	2	3	2	3	2
Hypothesized Difference	0		0		0	
df	2		2		2	
t-Stat	-10.61866194		-1.8537598		-32.944232	
P(T<=t) one-tail	0.004376223		0.1024729		0.00046006	
T-critical one-tail	2.91998558		2.91998558		2.91998558	
P(T<=t) two-tail	0.008752445		0.20494579		0.00092011	
T-critical two-tail	4.30265273		4.30265273		4.30265273	

10. Supplementary section 10: Images from optimization phase

Figure S10A-C shows samples that were discarded due to no visible spheroids in the gel (CRC-12), a gel that disrupted (CRC-16), and an abundance of bacteria in the sample (CRC-15). Figure S10A shows an example of an established sample (CRC-14). Figure S10E shows successfully established spheroids of CRC-11, CRC-13 and CRC-14 4 days after the thawing procedure.

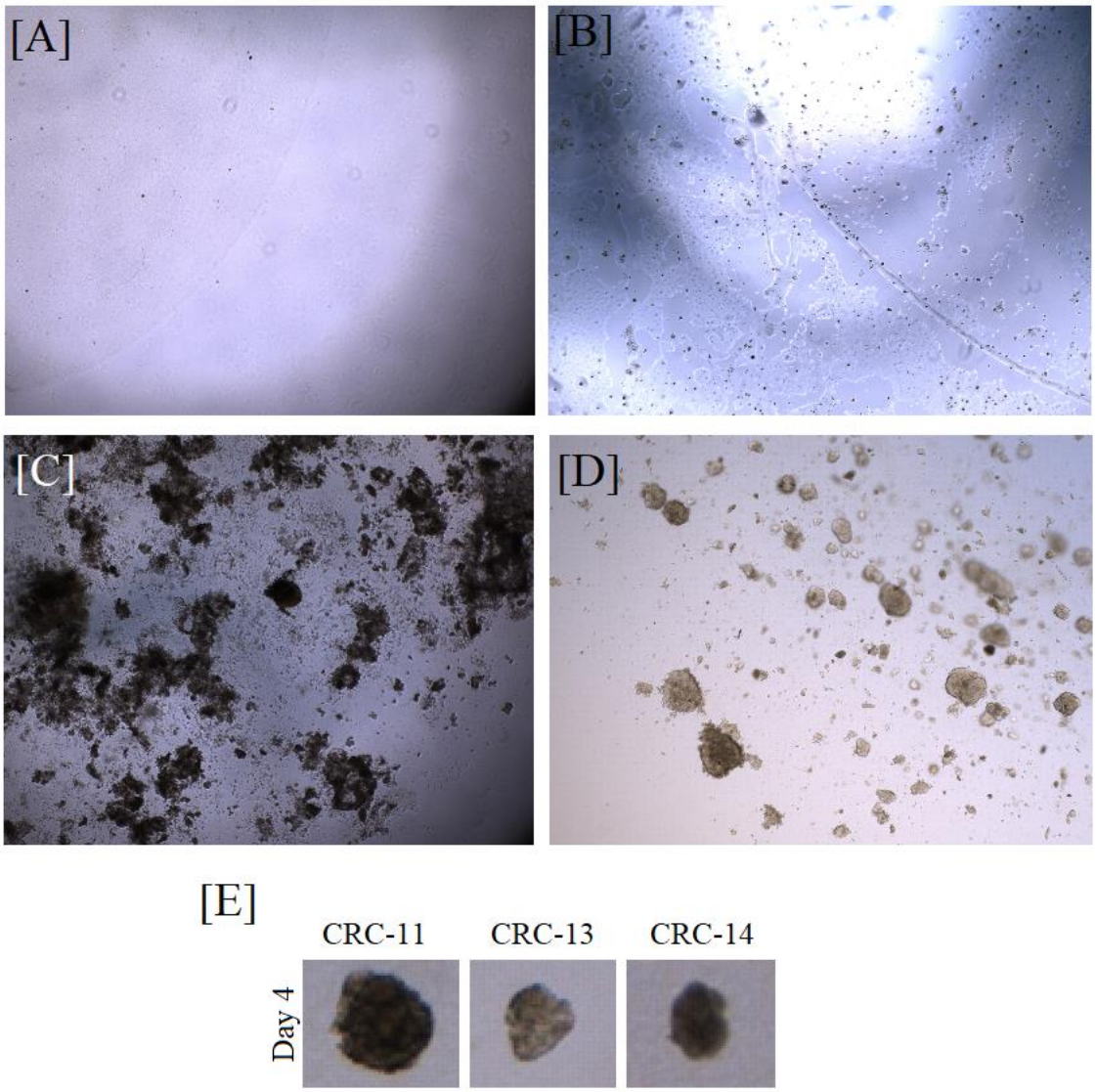


Figure S10. *Cancer-tissue originated spheroids (CTOSs) taken in a phase contrast microscope at 4X magnification. Samples were discarded in the optimization phase due to* ***[A]*** *no visible spheroids (CRC-12),* ***[B]*** *gel disruption (CRC-16), and* ***[C]*** *abundance of bacteria in the sample (CRC-15).* ***[D]*** *Established sample (CRC-14) with a high yield of spheroids in the gel.* ***[E]*** *CRC-11, CRC-13 and CRC-14 4 days after the thawing procedure.*

Vibeke Sognnes

F. nucleatum and miR-BART10-3p in colorectal cancer

NTNU

Norwegian University of
Science and Technology