Investigating the disease-promoting abilities of gut bacteria in colorectal tumour cells

James Lawrence Robson

A thesis submitted in partial fulfilment of the requirements of the University of the West of England, Bristol, for the degree of Doctor of Philosophy (PhD)

Faculty of Health & Applied Sciences

University of the West of England, Bristol

May 2021

Author's Declaration

This thesis is submitted in fulfilment of the requirements of the PhD Project,
and except where duly acknowledged or referenced it is entirely my own work. It has
not been submitted, either in whole or in part, for any other award at University of the
West of England (UWE) or elsewhere.

Signed: James Robson   **Date** 01/06/2021

Abstract

Colorectal cancer (CRC) is the third most common malignancy worldwide and the second leading cause of cancer-related death. The global burden of CRC is rising, particularly in younger populations. This increase has been attributed to lifestyle factors, chiefly diet, and their effects on the colonic microbiome. Several species of bacteria have been identified in high abundance from the gastrointestinal tract of CRC patients, and have been shown to exert direct and indirect pro-tumorigenic effects both *in vivo* and *in vitro*. However, what is unclear is whether these bacteria promote CRC development and progression across the entire adenoma-carcinoma sequence, or whether the interactions of bacteria with benign tumours differ from those with malignant ones.

Three CRC-associated species of bacteria, *Bacteroides fragilis*, *Enterococcus faecalis* and *Fusobacterium nucleatum*, in addition to the probiotic species *Escherichia coli* Nissle 1917, were investigated using the gentamicin protection assay to determine how they interact with the benign colorectal adenoma tumour cell line RG/C2, and the malignant adenocarcinoma line HCT116. All bacterial species were found to attach to and invade both cell lines. However, attachment was found to be higher in RG/C2 cells, in contrast to invasion which was higher in HCT116 cells. All species could also persist within tumour cells, and this bacterial infection promoted significant increases in cell yield whilst reducing apoptosis. There was evidence that *B. fragilis* and *F. nucleatum* may promote a Warburg-like metabolic effect in HCT116 cells, by upregulating glycolysis. In both wound healing and transwell filter migration assays, *E. coli* Nissle and *F. nucleatum* were able to promote both individual and collective cell migration in HCT116 cells; however, no bacterial species influenced migration in benign RG/C2 cells. Furthermore, *F. nucleatum* also promoted HCT116 invasion, suggesting that this species may be a driver of late-stage disease.

This thesis demonstrates that in a 2D model, there are differential bacterial interactions with colorectal tumour cells representing different CRC stages. It highlights that specific species are able to promote migration and invasion in malignant cells *in vitro*, but cannot confer this phenotype to benign cells. Despite this, all species investigated may promote tumour progression by increasing cell yield through the inhibition of apoptosis, possibly facilitating the accumulation of further genetic damage. In conclusion, this research demonstrates that colonisation of tumours by these species may contribute to poorer patient prognosis by accelerating CRC progression, and that these effects should be investigated in a more complex model of CRC. This thesis highlights the need to incorporate our knowledge of microbial contribution to CRC development and progression into future treatment and diagnostic strategies.

<u>Acknowledgements</u>

Table of Contents

List of Figures

x

List of Tables

List of abbreviations

ACF	Aberrant crypt foci
APC	Adenomatous polyposis coli
APS	Ammonium persulphate
ATP	Adenosine triphosphate
BFT	*Bacteroides fragilis* toxin
BHI	Brain heart infusion
BMP	Bone morphogenic protein
CFSE	Carboxyfluorescein
cfu	Colony forming units
CIMP	CpG island hypermethylation
CIS	Chromosomal instability
CMS	Consensus molecular subtype
CRC	Colorectal cancer
CRT	Catenin-related transcription
CSC	Colonic stem cell
DAPI	4′,6-diamidino-2-phenylindole
DMSO	Dimethyl sulphoxide
ECAR	Extracellular acidification rate
ECM	Extracellular matrix
EcN	*Escherichia coli* Nissle 1917
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EMT	Epithelial mesenchymal transition
EPEC	Enteropathogenic *Escherichia coli*
ETBF	Enterotoxigenic *Bacteroides fragilis*
FAP	Familial adenomatous polyposis
FBS	Foetal bovine serum
FCCP	Carbonyl cyanide-p-trifluoromethoxyphenylhydrazone
FIT	Faecal immunochemical test
FOBT	Faecal occult blood test
GALT	Gut-associated lymphoid tissue
GC/MS	Gas chromatography mass spectrometry

HCA	Hierarchical cluster analysis
HDACi	Histone deacetylase inhibitor
HIF-1α	Hypoxia-inducible factor 1 alpha
HNPCC	Hereditary non-polyposis colorectal cancer
IBD	Inflammatory bowel disease
IgA	Immunoglobulin A
LC/MS	Liquid chromatography mass spectrometry
Lgr5	Leucine-rich-repeat containing G-protein coupled receptor 5
LPS	Lipopolysaccharide
MHCII	Major histocompatibility complex class II
Min	Multiple intestinal neoplasia
MOI	Multiplicity of infection
NADPH	Reduced Nicotinamide adenine dinucleotide phosphate
OCR	Oxygen consumption rate
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCA	Principal component analysis
PCR	Polymerase chain reaction
PDK1	Pyruvate dehydrogenase kinase 1
PDX	Patient-derived xenograft
pks	Polyketide synthesis
PRR	Pattern recognition receptor
PVDF	Polyvinylidene difluoride
REG4	Regenerating family member 4
ROS	Reactive oxygen species
SASP	Senescence-associated secretory phenotype
SCFA	Short-chain fatty acid
SCNA	Somatic copy number variations
SDS	Sodium dodecyl sulphate
SFB	Segmented filamentous bacteria
SIFT-MS	Selected ion flow tube mass spectrometry
SOC	Super Optimal broth with Catabolite repression
SPME	Solid-phase microextraction
TEMED	Tetramethylethylenediamine

xiii

TME	Tumour microenvironment
VEGF	Vascular endothelial growth factor
UWE	University of the West of England

Chapter 1: Introduction

1.1 Structure and function of the human colon

The colon, or large intestine, is the distal portion of the bowel, and can be anatomically divided into four sections: the ascending colon, the transverse colon, the descending colon and the sigmoid colon, which is continuous with the rectum (Ellis and Mahadevan, 2014). The colon performs two primary functions: (1) the recovery of water and electrolytes from its luminal contents and (2) it facilitates the formation of stool, which is transported along the length of the colon to the rectum where it is excreted. Movement of colonic material is facilitated by peristalsis, coordinated contractions of the circular and longitudinal layers of the colonic muscle wall (Spencer *et al.*, 2016). The majority of colonic functions are performed by the highly specialised epithelial layer. For example, colonocytes actively absorb Na^+ ions from the lumen using glucose-dependent transporters; this maintains an osmotic gradient between the lumen and the epithelium, allowing continuous water absorption (Fordtran and Dietschy, 1966).

The colonic mucosa is composed of the epithelium and underlying lamina propria. It is the most luminal part of the bowel wall and is in direct contact with the luminal contents. Beneath this is the submucosa, followed by the muscularis propria and the adventitia (Levine and Haggitt, 1989). The epithelial layer is composed of polarised absorptive colonocytes and mucus-secreting goblet cells, which form a single-cell thick columnar epithelium. Cellular turnover at the mucosal surface is high, with the entire epithelial lining being renewed approximately every 4-5 days (Barker *et al.*, 2007). Unlike the small intestine, where the epithelium is arranged into millions of finger-like villi, the colon epithelial surface is mostly flat, with the exception of the

colonic crypts. These crypts are glandular invaginations of the epithelium, and are essential for cell renewal and colonic homeostasis (Clevers, 2013).

1.2 Colonic crypts are the drivers of colonic homeostasis

Cells at the colonic surface undergo apoptosis and are continuously sloughed off into the lumen (Blander, 2016). This prevents the accumulation of damaged cells, and the colonocytes are then replaced by new cells which emerge from the approximately 10 million crypts that are distributed along the length of the colon (Stamp *et al.*, 2018). Cells produced at the base of the crypt migrate upwards towards the colonic surface, in a conveyor belt-like fashion. To facilitate this constant cell renewal, each colonic crypt harbours a resident stem cell population (Barker, 2014). These stem cells undergo asymmetrical cell division to give rise to transit-amplifying progenitor cells, which migrate up the crypt, rapidly dividing, and undergo differentiation into several lineages before reaching the mucosal surface (Barker, 2014). Colonocytes are the primary cell lineage produced, in addition to mucus-secreting goblet cells, endocrine cells, microfold cells and tuft cells which also comprise the epithelium.

The colonic stem cells (CSC) are key to colon homeostasis and are found at the base of the crypt between the +1 and +3 positions (Figure 1.1). The maintenance of the stem cell niche is reliant upon concentration gradients of numerous signalling molecules, including Wnt and bone morphogenic protein (BMP), along the length of the crypt (Reya and Clevers, 2005). Thus, CSCs can be identified by the presence of the Wnt target, the leucine-rich-repeat containing G-protein coupled receptor 5 (Lgr5; Barker *et al.*, 2007). Wnt ligands are produced and secreted by specialised mesenchymal cells, which reside in the crypt base and surrounding tissue. Deep crypt secretory cells, characterised by the expression of regenerating family member 4 (*REG4*), are interspersed between CSCs and express Wnt ligands, in addition to

epidermal growth factor, Notch ligands and antimicrobial peptides (Sasaki *et al.*, 2016). Wnt ligands from the surrounding stroma are also essential, with GLI-expressing mesenchymal cells also required for the maintenance of the stem cell niche (Degirmenci *et al.*, 2018).

Wnt signalling regulates numerous genes, influencing differentiation, stemness and proliferation, having been extensively reviewed by Clevers and Nusse (2012). Altered Wnt signalling dramatically impacts colonic homeostasis, with subsequent damaging effects on gut health. In mouse models, inhibition of downstream Wnt effector proteins, including T-cell factor 4, leads to a complete loss of the CSC population; this prevents crypt formation and severely compromises the mucosa (Korinek *et al.*, 1998). Where colonic stem cells are lost from the mature mucosa through injury, the differentiated epithelium is able to recover this population and restore crypt function, as reviewed by de Sousa and de Sauvage, 2019. Remarkably, several cell lineages have been demonstrated to display plasticity and, after ablation of the stem cell population, express Lgr5 and migrate towards the crypt base. These include transit-amplifying progenitor cells and secretory cells. Aberrant increases in Wnt signalling within the stem cell population are also relevant to colonic disease, and are a driver for many forms of cancer, with 93% of colon cancers displaying mutations which lead to upregulated Wnt signalling (Muzny *et al.*, 2012).

The integrity of the stem cell population is therefore vital to colonic homeostasis; however, the mucosa faces numerous challenges. These include large fluctuations in pH, mechanical abrasion from the luminal contents and virulence from ingested pathogens. In addition, the mucosa is in constant contact with the gut microbiome, the community of bacteria, fungi, archaea, viruses and protozoa which colonise the gastrointestinal tract (Peterson *et al.*, 2009).
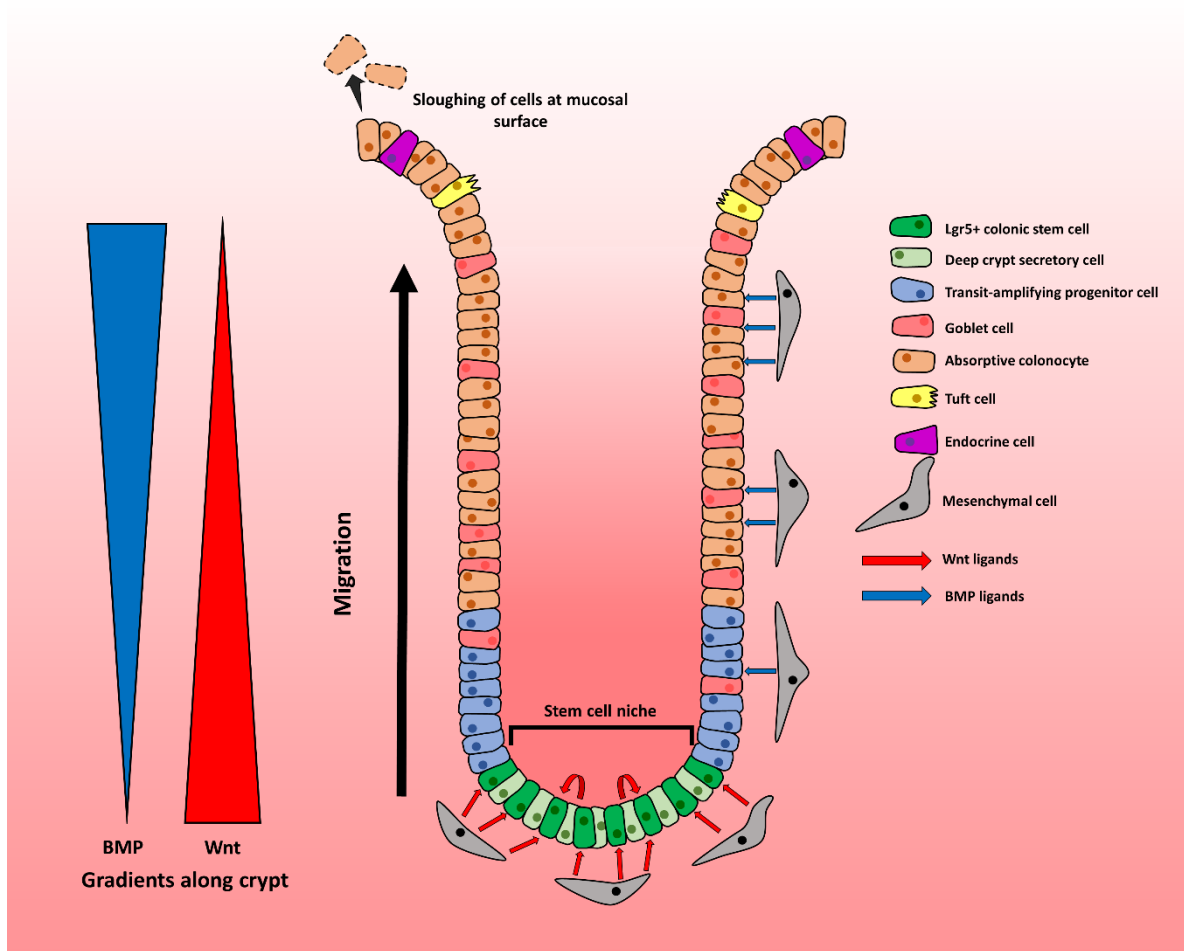
Figure 1.1: The structure and regulation of the colonic crypt. The stem cells, identified by the presence of Lgr5, at the base of the crypt are maintained through signalling by surrounding crypt cells and mesenchymal cells in the lamina propria. They asymmetrically divide into transit amplifying progenitor cells which migrate upwards and differentiate into one of the several cell lineages present in the colon. Increasing gradients of BMP and decreasing gradients of Wnt promote differentiation. At the mucosal surface, colonocytes undergo apoptosis and are sloughed into the lumen (Barker, 2014).

1.3 The Gut Microbiome

Microorganisms colonise numerous body sites, including the skin, oral cavity, vagina, and gastrointestinal tract. The gut microbiome refers to the consortia of microorganisms found within the gastrointestinal tract, which can form commensal and symbiotic relationships with the host (Hooper and Gordon, 2001). The colon harbours the largest bacterial load of the gastrointestinal tract, and more bacteria than all other body sites combined. Its approximately 10^{14} bacteria dwarf the combined bacterial population of the rest of the body, which is estimated to be $>10^{12}$ (Sender, Fuchs and Milo, 2016). Therefore, maintaining homeostasis in the colon is dependent on physical, chemical, and immune-regulatory processes which control interactions between the microbiome and the host.

1.3.1 The biogeography of the colonic microbiome

The colon provides a wide range of ingested and host-derived substrates for microorganisms, which foster the development of highly specialised niches. Many bacterial constituents of the gut microbiome are present only transiently within the colonic lumen; however, the colonic mucosa provides a habitat for many species to form permanent associations with the host (Figure 1.2; Derrien and van Hylckama Vlieg, 2015). Bacterial distribution along the length of the colon is influenced by gradients of pH and nutrient availability, whereas increasing oxygen concentration and the abundance of immune effectors and other secretions at the mucosal surface control bacterial colonisation across the mucosa-lumen axis (Donaldson, Lee and Mazmanian, 2015)
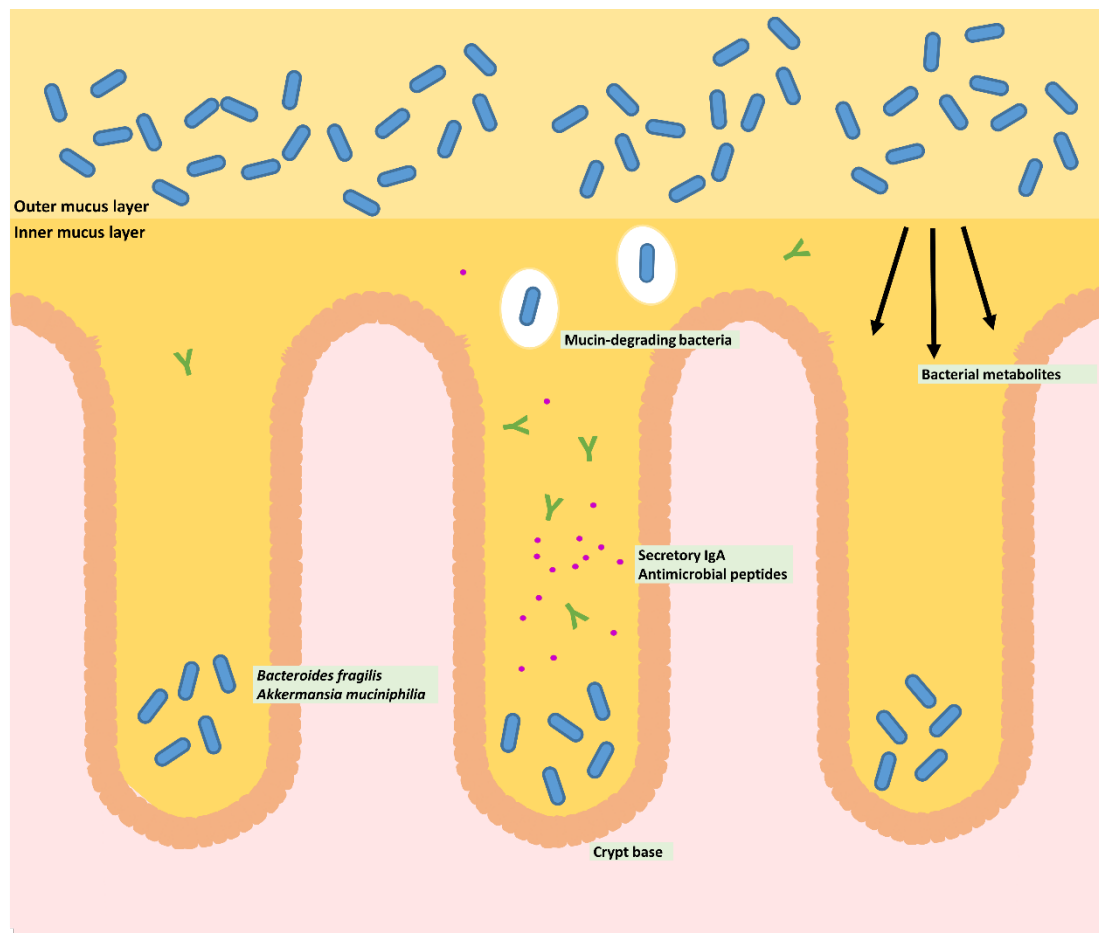
Figure 1.2: Spatial organisation of the colonic microbiota. The vast majority of bacteria (blue) are present in the outer mucus layer, which has a greater pore size and is loosely attached to the mucosa. Mucin-degrading species are able to penetrate the inner mucus layer and benefit from its abundant incorporated nutrients. In addition to increasing mucus density, a gradient of increasing oxygen concentration, antimicrobial peptides, and IgA controls bacterial growth in the inner mucus layer. A select few species, including *B. fragilis* and *A. muciniphilia* are able to access and colonise the colonic crypts (Donaldson, Lee and Mazmanian, 2015).

Maintaining proper spatial organisation of the colonic microbiota is essential for homeostasis. Goblet cells within the colonic crypts secrete mucus, comprising of heavily glycosylated high-molecular weight mucin proteins (Pelaseyed *et al.*, 2014). This mucus both protects the epithelium from bacteria whilst also providing a substrate for bacterial colonisation and metabolism (Rodríguez-Piñeiro *et al.*, 2013). In the

colon, the major mucus component is the gel-forming mucin protein Muc2 (Audie *et al.*, 1993). Mucin proteins are heavily glycosylated and form a net-like mesh, which has a relatively small pore size of 20-200 nm at the mucosal surface and acts to physically exclude bacteria (Johansson *et al.*, 2008). As the mucus travels towards the lumen, action by host and microbial proteases increases this pore size, exposing the abundant mucin-bound carbohydrates (e.g. O-linked oligosaccharides) and proteins present in the mucus and provides a potential growth substrate for microorganisms (Kaoutari *et al.*, 2013). This constant efflux of mucin from the mucosal surface protects the underlying epithelium, and specifically the intestinal crypt epithelium, from direct contact with bacteria (Round *et al.*, 2012). However, a select few microorganisms, including *Bacteroides fragilis*, are able to invade the mucus layer and reside on the mucosal surface through immune modulation and utilisation of host-derived substrates (Figure 1.2; Round *et al.*, 2011). Privileged sites such as the colonic crypts also display a distinct microbiome dominated by *Firmicutes*, with *Actinobacteria*, *Bacteroidetes* and *Alphaproteobacteria* also present (Saffarian *et al.*, 2019). The availability of mucin-derived substrates is one of several mechanisms through which the composition of the microbiome is regulated by the host, which also include secretion of bile acids, antimicrobial peptides and Immunoglobulin A (IgA) antibodies (Staley *et al.*, 2017; Peterson *et al.*, 2007).

1.3.2 The composition of the colonic microbiome is key to its function and gut homeostasis

Current estimates place the number of bacterial species present in the colon at >1000, with the combined colonic microbiome encoding 100 times the number of genes present in the human genome (Bäckhed *et al.*, 2015). The Human Microbiome Project revealed that there is significant variation in gut microbiota composition between

individuals, both in the species present and their relative abundance (The Human Microbiome Consortium, 2012). This poses a challenge to microbiome research, as the apparent lack of a 'healthy' reference microbiota composition means that it is difficult to study how the microbiota shifts during disease (McBurney *et al.*, 2019).

Recent studies focussing on microbial ecological succession have provided some indication of the factors which regulate microbiome composition, with initial colonisation occurring at birth. Depending upon the mode of delivery, neonates are exposed either to the maternal vaginal microbiota, or non-maternal environmental bacteria (Neu and Rushing, 2011). The placenta, despite previously being thought to be sterile, is now known to be colonised by microorganisms, and harbours a microbiota similar to that of the oral cavity (Aagaard *et al.*, 2014). These microorganisms are not thought to be conferred to the infant, although this is hotly contested (Perez-Muñoz *et al.*, 2017). After birth, the composition of a child's microbiota fluctuates dramatically during the first year of life (Bäckhed *et al.*, 2015). Many factors, including chance environmental encounters and early exposure to antibiotics exert significant influence over colonic microbiome composition; however, the most impactful factor is breast feeding (Stewart *et al.*, 2018).

Human breast milk contains approximately 10^3 bacteria per millilitre, and these bacteria directly populate the infant gut (Fernández *et al.*, 2013). The oligosaccharides present in breast milk provide the primary source of bacterial substrate in early life, promoting a gut microbiota dominated by *Bifidobacteria* (Turroni *et al.*, 2019). Formula-fed infants display increased microbiome diversity, and this has been associated with an increased occurrence of adverse outcomes, including diarrhoea (Ho *et al.*, 2018). This is in contrast to the adult microbiome where a lack of diversity is associated with disease (Shanahan, 2013). Microbiota

composition remains relatively stable in healthy infants until they are weaned, where the introduction of the diverse substrates found in solid foods promotes the establishment of a more complex microbiota (Bäckhed *et al.*, 2015). Thus, an individual's 'normal' microbiota composition is largely dependent on these formative years (Bokulich *et al.*, 2016). The adult gut microbiota is far more stable, with 60% of species persisting for longer than 5 years (Faith *et al.*, 2013) and is dominated by *Bacteroides* and *Firmicutes*, with other notable phyla including *Actinobacteria* and *Proteobacteria* (Lloyd-Price, Abu-Ali and Huttenhower, 2016).The impact of breast feeding demonstrates how diet modulates microbiota composition, and this remains true in adults, with consistent dietary patterns shaping the microbiome over time (David *et al.*, 2014). Host genetics are also known to play a significant role. Studies comparing the composition of the microbiota in monozygotic and dizygotic twins have demonstrated that the microbiotas of monozygotic twins display much greater concordance (Goodrich *et al.*, 2014).

1.3.3 Immune system development occurs alongside microbiome maturation

The stability of the microbiota composition is critical to gut homeostasis, and throughout the establishment of the microbiota in infants, the immune system develops in parallel (Martin *et al.*, 2010). This has been demonstrated in mouse models whereby germ-free animals display improper development of myeloid and T-cell populations (Williams *et al.*, 2006). During the maturation of the immune system, tolerance to commensal microorganisms is gained.

The gut-associated lymphoid tissue (GALT) is the largest immune system in the human body, and the gut microbiota provides a constant source of stimulation (Arrazuria *et al.*, 2018). Specialised epithelial cells within the mucosa, the microfold

cells, constantly collect and transfer antigens from the colonic lumen to the underlying GALT, which are then disseminated throughout the systemic circulation via dendritic cells (Martin *et al.*, 2010). This process is critical to the development of tolerance to ingested compounds, including food proteins, and the numerous microbial antigens acquired from food-borne microorganisms and the microbiota. Toll-like receptors on the surface of epithelial cells act as pattern recognition receptors (PRRs), and bind microbe-associated molecular patterns, which may be present on the bacterial surface or secreted into the environment (Chu and Mazmanian, 2013).

Tolerance to commensal microorganisms is developed during early life, and is aided by the presence of specific bacteria (e.g. *B. fragilis*) within the microbiome which suppress inflammation, and trigger the differentiation and migration of T-cell populations (Mazmanian *et al.*, 2005). Although the contribution of the microbiome to immune development is not yet fully understood, relationships between microorganisms and immune cell populations have been established. Segmented filamentous bacteria (SFB) are key to the maturation of Th17 cells in rodents (Ivanov *et al.*, 2009), and SFB populations have recently been discovered in the human gut, suggesting a similar role (Jonsson *et al.*, 2020). Presentation of antigens from SFB and other bacteria such as *B. fragilis* promotes the secretion of anti-inflammatory IL-10 by regulatory T cells (Round and Mazmanian, 2010). The development of this tolerance therefore permits mutualism between the host and a range of bacterial species, which as previously mentioned vary greatly between individuals according to genetic and environmental factors.

The development of this tolerance to commensal microorganisms is essential to gut health, and impacts all aspects of gut homeostasis, beginning at the stem cell level. It was recently demonstrated that the presence of bacterial antigens in the stem

cell compartment initiates a feedback loop whereby antigen-presenting Lgr5+ CSCs show enriched expression of major histocompatibility complex class II (MHCII) protein, allowing them to interact with T regulatory cells (Biton *et al.*, 2018). During infection, Th1 and Th2 cells stimulate CSC differentiation, at the expense of self-renewal, in order to facilitate barrier repair. After this immune response T_{reg} cell numbers are increased, which inhibits CSC differentiation and allows stem cell numbers to be replenished (Tanoue, Atarashi and Honda, 2016). Where this complex tripartite relationship between intestinal stem cells, inflammation and microorganisms is disturbed, there is a significant impact on gastrointestinal health. For example, the presence of immunogenic microbial products at the mucosal surface leads to the production of IL-23 by resident myeloid cells (Grivennikov *et al.*, 2012). This facilitates an IL-17 immune response, which has been linked to numerous forms of cancer including colorectal, as reviewed by Zhao *et al.*, (2020).

It is clear the homeostasis between the microbiome and mucosa is reliant upon stringent immune mechanisms and the recognition of commonly encountered microbial antigens. This however seems at odds with the significant variation in microbiota composition observed between individuals. Remarkably, functions carried out by microbiomes of differing composition are able to be conserved due to the presence of a 'core microbiota phenotype' (Turnbaugh *et al.*, 2009). Despite significant variation in microbiome compositions between individuals, the functional phenotype remains largely consistent. This core microbiota phenotype is the result of the conservation of bacterial genes across species, which allows the microbiota to perform the functions essential for gut homeostasis in an average individual (Turnbaugh and Gordon, 2009; Zeevi *et al.*, 2019). Many of these functions are concerned with the metabolism of ingested compounds.

1.3.4 Functions of the gut microbiome

The majority of dietary carbohydrates, proteins and fats are absorbed in the jejunum of the small intestine (Kiela and Ghishan, 2016). Non-digestible compounds, including resistant starch, reach the colon intact and are metabolised by resident gut microbes (Flint *et al.*, 2012). The contribution of microbial products to nutrient and energy homeostasis is substantial. Germ-free mice, lacking a gut microbiome, become underweight compared to wild-type individuals fed the same diet (Bäckhed *et al.*, 2004). However, reconstitution with a conventional microbiota results in a significant increase in body mass, even where food is restricted (Bäckhe*d et al*., 2004). In addition, the wealth of enzymes encoded by the colon microbiome allows for the metabolism and synthesis of products that would be otherwise unavailable to the host. The colon microbiota contributes to protein catabolism via two major mechanisms. Direct proteolysis by bacteria converts ingested proteins into amino acids, which are either incorporated into bacterial cells as they proliferate or absorbed by the colonic epithelium (Yao, Muir and Gibson, 2016). Microbes also ferment proteinaceous compounds, producing harmful gaseous derivatives (Macfarlane, Cummings and Allison, 1986). These products, which include ammonia, hydrogen sulphide (H_2S), phenols, indols and N-nitroso compounds, exert genotoxic effects, linking the consumption of high quantities of protein to DNA damage (Hughes, Magee and Bingham, 2000).

It is estimated that bacterial carbohydrate fermentation in the colon accounts for 10% of an individual's total energy requirements (McNeil, 1984). The most abundant products of fermentation in the colon are short-chain fatty acids (SCFAs), namely butyrate, propionate and acetate (Bird *et al.*, 2010). Between 90-95% of SCFAs produced in the colonic lumen are absorbed by colonocytes, and they exert

multiple effects on gut and systemic health. Acetate and propionate are primarily transported and metabolised outside of the colon, in peripheral tissues and the liver, respectively (Wong *et al.*, 2006). However, butyrate is metabolised by the colonic mucosa, with colonocytes generating approximately 70% of their adenosine triphosphate (ATP) through butyrate metabolism (Roediger, 1980). Furthermore, butyrate is also essential for gut homeostasis, and impacts a myriad of gut functions including proliferation, apoptosis, motility and differentiation by regulating gene expression as a histone deacetylase inhibitor (HDACi; Hamer *et al.*, 2008).

Ingested dietary compounds constitute the major source of substrate for the gut microbiota, and as discussed, the resulting metabolites have a significant impact on host health. However, the interaction between endogenous host products and the microbiota is also significant. Bile acids are secreted from the gall bladder into the lumen of the small intestine, where they aid in the digestion and absorption of lipids (Lefebvre *et al.*, 2009). Although 95% of bile acids are reabsorbed in the distal ileum, the remaining 5% enters the colon, where they exert a selection pressure on the colon microbiota. Bile acid antimicrobial activity has been documented against many constituents of the microbiota, including *Lactobacilli,* and *Bifodobacteria* species, through the induction of membrane damage and disruption of transmembrane electrical potential, which is important for essential cellular functions including ATP synthesis (Kurdi *et al.*, 2006). Many microbiota constituents possess bile salt hydrolases, which reduce bile acid toxicity by cleaving their conjugated amino acids (Song *et al.*, 2019). These free amino acids are incorporated into bacterial cells, whereas the deconjugated bile acids are reabsorbed through the mucosa. However, some of these deconjugated bile acids are targeted for further modification and transformation into secondary bile acids by bacteria possessing 7α-dehydroxylating

enzymes (Staley *et al.*, 2017). Secondary bile acids, including lithocholic acid and ursodeoxycholic acid, are produced via these mechanisms and are thought to cause DNA damage in the mucosa, although the precise mechanism remains unclear (Bernstein *et al.*, 2009).

Finally, competitive exclusion of pathogens is another important role carried out by the colonic microbiota. Depending on dietary composition, the average individual will ingest approximately 10^6-10^9 bacteria each day (Lang, Eisen and Zivkovic, 2014). Many of these food-borne bacteria are potential gastrointestinal pathogens, including *Salmonella, Shigella,* and *Campylobacter* species, as well as Enteropathogenic *Escherichia coli* (EPEC; Hara-Kudo *et al.*, 2013). The colonic microbiome is able to suppress these pathogens through secretion of metabolites. For example, butyrate has been shown to directly downregulate the transcription of virulence factors in several *Salmonella enterica* serovars (Gantois *et al.*, 2006). In addition, the fierce competition for both space and resources in the intestinal microbiota also excludes these pathogens. Loss of this competitive exclusion through disruption of the microbiome (e.g. through antibiotic use) is associated with colonisation by pathogens such as *Clostridium difficile*, resulting in persistent diarrheal disease (Mullish and Williams, 2018). Therefore, in addition to the direct beneficial effects of metabolite production, the microbiota indirectly protects and maintains gut homeostasis and limits the possibility for infection.

1.3.5 Perturbance of the microbiome has significant effects on health

It is evident that the gut microbiome is critical to the development and maintenance of homeostasis in the colon, and that the composition of the microbiota is finely balanced within each individual. Alterations to this balance, such as those caused by diet, antibiotic use, or infection, can have significant effects on both gut and systemic

health. Shifts in the relative abundance of species from those normally present in an individual, or shifts in the metatranscriptome of established species, can lead to colonic dysbiosis (Petersen and Round, 2014). Dysbiosis is characterised by the loss of beneficial microorganisms, colonisation by and outgrowth of potential pathogens, and a reduction in species diversity; these shifts may occur independently or together, altering the relative abundance of species present in the microbiome (Figure 1.3; David *et al.*, 2014). These changes compound one another over prolonged time periods, as changes in species abundance lead to altered metabolite levels and immune responses which further modulate microbiota composition. Furthermore, the proximity between bacteria in the mucus layer fosters nutritional relationships and dependencies between species (Flint *et al.*, 2007). Solubilisation of mucin-bound compounds and production of SCFAs are but two mechanisms through which bacteria in the colon cross-feed with other species (Schroeder, 2019).

Due to the interdependence of microorganisms in the colon, small shifts in relative metabolite levels or intestinal immunity can compound to cause drastic changes in the microbiota composition, potentially leading to the loss of entire species and altering the phenotypes of other organisms (Sonnenburg *et al.*, 2016). Conservation of bacterial genes across a wide range of species facilitates species-level variation amongst individuals without altering the microbiome phenotype. However, there are non-redundant genes and relationships present between an individual and specific species within their microbiome (Qin *et al.*, 2010). For example, a recent metabolic analysis of the microbiome has revealed that at least 434 metabolic pathways are present in a single species, and 91.6% of these pathways were present in all individuals tested (Visconti *et al.*, 2019). Thus, altered abundance of microbiota species may predispose altered levels of bacterial antigens, abnormal concentrations

of bacterial metabolites, outgrowth and loss of microbiome residents, and colonisation of the colon by alien species. This results in dysbiosis, which is associated with chronic low-grade inflammation (Petersen and Round, 2014). Both the environmental changes observed in dysbiosis and the action of specific bacterial species, which may be altered in this environment, have been linked to numerous diseases. These include obesity, inflammatory bowel disease (IBD) and Type 2 diabetes (Aron-Wisnewsky *et al.*, 2019; Tomasello *et al.*, 2014) . Recently, the role of the gut microbiome in colorectal cancer (CRC) has gained increased attention.
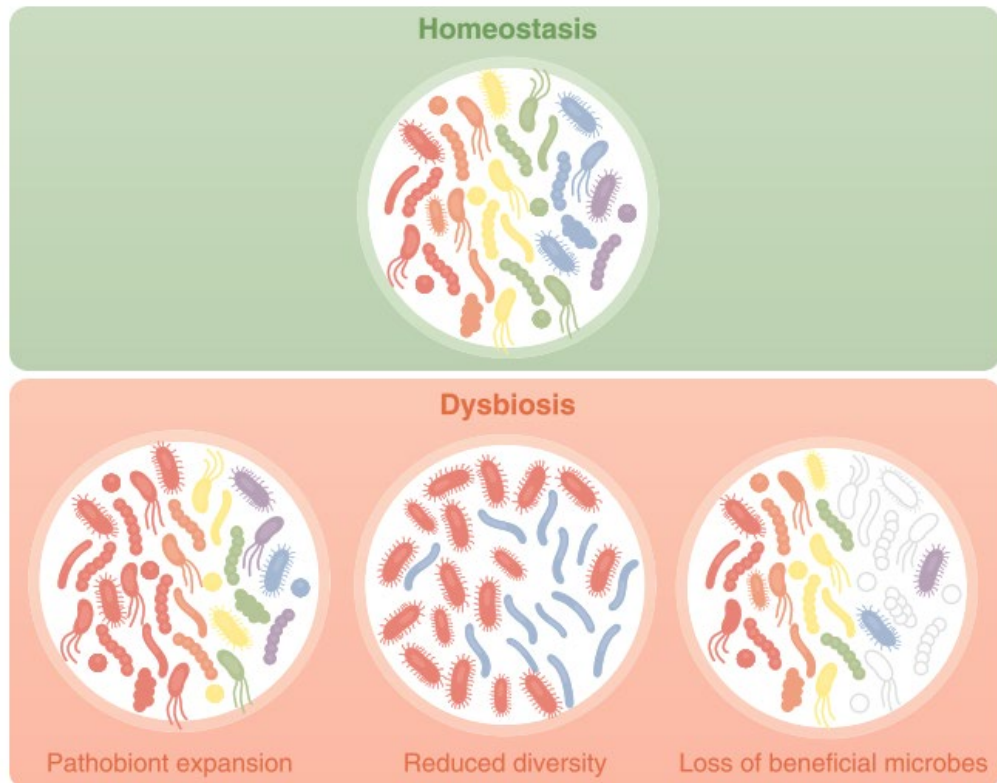
Figure 1.3: Microbiota changes during colonic dysbiosis. In healthy individuals, the colonic microbiome is composed of a diverse range of species which perform functions beneficial to host health. Dietary and lifestyle factors alter the colonic microenvironment leading to disrupted microbial growth. These alterations in microbiota composition lead to chronic inflammation. Expansion of pathobionts causes direct virulence against the mucosa. Reduced diversity and loss of beneficial microbes leads to the loss of non-redundant influences on host health and microbiota stability, leading to further divergence from homeostasis. These dysbiotic changes may occur separately or concurrently, having profound effects on host health. Reproduced with permissions from Petersen and Round (2014).

1.4 Colorectal cancer

1.4.1 Epidemiology of colorectal cancer

Colorectal cancer (CRC) is the third most common malignancy worldwide, accounting for 11% of all cancer cases, and is the second leading cause of cancer-related death (Sung *et al.* 2021). CRC incidence correlates positively with human development factors, and is therefore predicted to rise globally, with the largest increases expected in developing low/middle income countries (Arnold *et al.*, 2017). CRC prognosis is highly dependent on tumour stage at diagnosis. In the UK, >90% of individuals diagnosed with CRC at stage I will experience disease-free survival over 5 years. However, 55% of cases are diagnosed at stage III or IV, with the 5-year survival rate of these individuals being 65% and <10%, respectively (Cancer Research UK, 2021). The low survival rate of late-stage CRC patients can be attributed to a paucity of effective treatment options, difficulty in detecting metastases and disease recurrence. In addition, the formation of secondary tumours, particularly in the lymph and liver nodes, is common in CRC, further complicating treatment (Ganesh and Massagué, 2021). Therefore, improving early detection and our understanding of disease progression is a major obstacle to improving CRC survival.

CRC is one of the few cancers routinely screened for in many countries (Navarro *et al.*, 2017). In the UK, routine faecal screening is currently offered from age 60 onwards. As expected, the majority of CRC cases in the UK are diagnosed in individuals aged over 50, a trend which is reflected worldwide (Cancer Research UK, 2021). However, global incidence of CRC in individuals under the age of 50 is increasing, with a rise of 47% between 1975-2015 (Rawla, Sunkara and Barsouk, 2019). In populations that are not offered cancer screening, the burden of diagnosis often falls upon the patient's ability to detect and report symptoms, with common signs

of CRC including rectal bleeding and abdominal pain (Del Giudice *et al.*, 2014). This is problematic, as a systematic review of CRC symptomology found that many of the symptoms reported in published guidelines are not consistently present in CRC patients (Adelstein *et al.*, 2011). Of these symptoms, rectal bleeding and weight loss are not indicative of precancerous polyps, meaning that the appearance of detectable symptoms often corresponds to advanced disease (Adelstein *et al.*, 2011). Therefore, there is a clear need to develop our understanding of both the factors contributing to the global rise in CRC incidence, and how these factors can be used to inform current methods of cancer detection and diagnosis. As the development of CRC normally occurs over several decades, ample opportunity is present for intervention if our understanding of CRC progression is improved.

An estimated 30% of CRC cases can be attributed to hereditary factors (Kwak and Chung, 2007). Of these cases, 5-10% are caused by known inherited conditions, including familial adenomatous polyposis (FAP) and hereditary non-polyposis colorectal cancer (HNPCC). Although the primary genetic defects predisposing these conditions have been identified, the remaining hereditary cases remain poorly understood. Nevertheless, the overwhelming majority of cases (>70%) are sporadic in nature (Sung *et al.*, 2021). For sporadic CRC to develop, successive accumulation of pro-oncogenic mutations in a single colon cell or its progeny, often occurring over several decades, is required (Fearon and Vogelstein, 1990). The increase in CRC incidence in younger populations suggests that modifiable factors are accelerating this process, with diet and the microbiome being heavily implicated.

1.4.2 The Multistep Nature of Cancer

The development and progression of sporadic CRC has been well described by the Multistep Nature of Cancer model (Fearon and Vogelstein, 1990). This model states

that 3-7 'hits' – mutations to key oncogenes or tumour suppressor genes – are required for cancer to develop (Figure 1.4). Many of these early mutations have been well-defined in colorectal tumorigenesis, and correspond to morphological and histological changes (Vogelstein and Kinzler, 1993). Early mutations cause the epithelium to become hyperproliferative, with subsequent mutations leading to the development of small adenomas, which increase in size and accumulate driver mutations before developing into carcinoma *in situ* and eventually culminating in an invasive, metastatic tumour (Vogelstein and Kinzler, 1993). A mutation leading to upregulated Wnt signalling is the most common event initiating CRC development, and is found in 93% of colorectal tumours (The Cancer Genome Atlas Network, 2012). The most common of these mutations is an inactivation of the adenomatous polyposis coli (*APC*) tumour suppressor gene, which is seen in >80% of sporadic cases (Fearnhead, Britton and Bodmer, 2001). This loss of *APC* leads to the inactivation of the β-catenin destruction complex, allowing cytosolic β-catenin to translocate to the nucleus resulting in the stimulation of Wnt target genes (Munemitsu *et al.*, 1995).
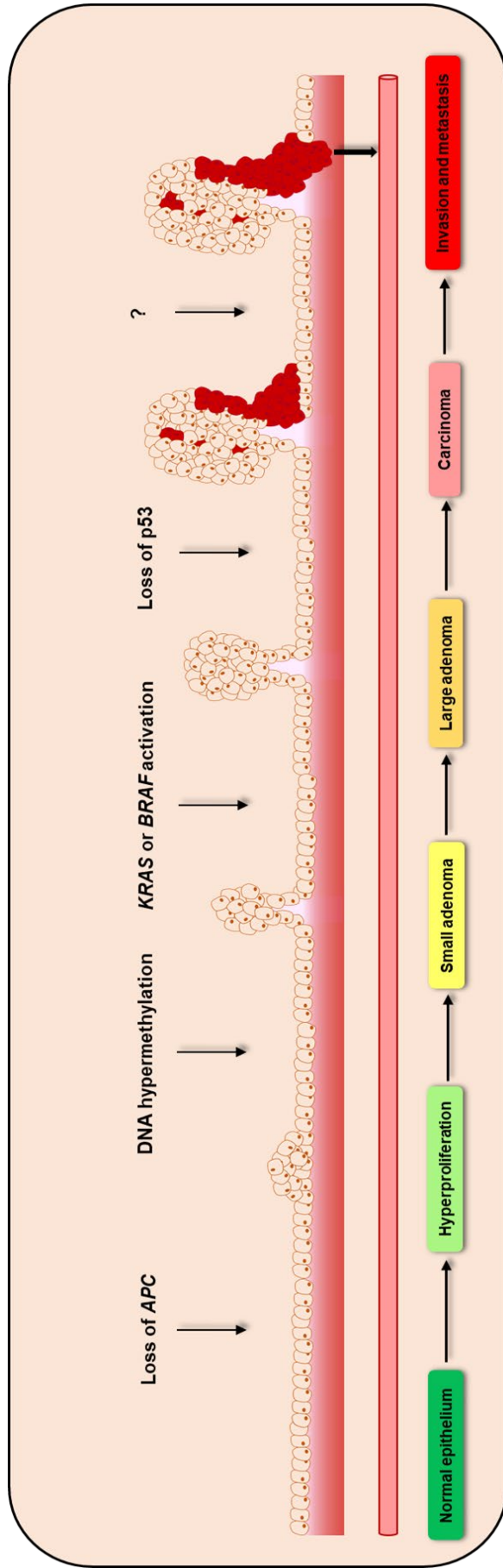
Figure 1.4: The multistep nature of cancer. For CRC to develop, successive 'hits' are required whereby key oncogenes and tumour suppressor genes become mutated. In sporadic CRC, an inactivating mutation in the *APC* tumour suppressor gene is thought to be the initial driver of tumorigenesis. As successive mutations accumulate, well-defined phenotypic changes in the epithelium are seen whereby increased proliferation gives rise to adenomas which eventually progress into cancer. Invasion and metastasis are the final stages of CRC progression; however, this process has not been linked to any specific genetic event. In an average individual this process is thought to occur over several decades, and many healthy individuals display aberrant crypt foci and small adenomas which never progress to cancer. Environmental factors, including the gut microbiome, are thought to accelerate this process in addition to promoting metastasis. Figure adapted from Vogelstein and Kinzler (1993).

Loss of *APC* has been associated with chromosomal instability (CIS), whereby alterations to the number and structure of chromosomes occur in tumour cells due to the reduced binding of microtubules which impairs their stability (Pino and Chung, 2010). CIS is the most common form of genomic instability seen in CRC, and accounts for the majority of sporadic cases (Lengauer, Kinzler and Vogelstein, 1997). Genomic instability, characterised by a high frequency of mutations and genome altering events, drives cancer formation as the baseline mutation rate in normal cells is insufficient to drive tumorigenesis (Loeb, Loeb and Anderson, 2003). The literature is currently conflicted as to whether CIS is a prerequisite to tumorigenesis, or whether the initial oncogenic mutations occur at a normal mutation rate, and are then selected for due to the growth advantages that they confer (Sieber, Heinimann and Tomlinson, 2003). Other mechanisms of genomic instability are also present in CRC, including microsatellite instability and CpG island hypermethylation (CIMP; Li and Martin, 2016).

Early CRC diagnosis was primarily based upon characteristics such as CIS, microsatellite instability, and the presence of key mutations such as *BRAF* and *KRAS*. However, over the past decade there has been a push for the categorisation of CRCs into more comprehensive subtypes (Cancer Genome Atlas Network, 2012; Budinska *et al.*, 2013). In 2015, a global consortium of CRC researchers and clinicians used gene expression-based algorithms to devise four new CRC subtypes dubbed the 'consensus molecular subtypes' (CMS), with the aim of providing a classification system with greater biological interpretability (Guinney *et al.*, 2015). The prognostic value of these classifications in predicting disease outcome has since been confirmed, and their main features are summarised in Table 1.1 (Mooi *et al.*, 2018).

Table 1.1. The abundance and characteristics of colorectal cancer consensus molecular subtypes, adapted from Mooi *et al.*, 2018.

CMS1 MSI Immune	CMS2 Canonical	CMS3 Metabolic	CMS4 Mesenchymal
14%	37%	13%	23%
MSI, CIMP high, hyper-mutation	SCNA high	Mixed MSI, SCNA low, CIMP low	SCNA high
BRAF mutations		*KRAS* mutations	
Immune infiltration and activation	WNT and MYC activation	Metabolic deregulation	Stromal infiltration, TGF-β activation, angiogenesis
Worse survival after relapse			Worse relapse-free and overall survival

During the development of CRC, the hyperproliferative epithelium gives rise to increasing numbers of aberrant crypt foci (ACF). These ACF were first discovered in the murine colon, where methylene blue staining and histological examination revealed the presence of larger, thicker crypts which stained darker than adjacent normal crypts (Bird, 1987). ACF were then demonstrated to be inducible by administering carcinogenic compounds to rats, and after their discovery in human CRC patients were designated as CRC precursors (Tudek, Bird and Bruce, 1989; Pretlow *et al.*, 1991). The number of ACF present in the colon increases at each stage along the adenoma-carcinoma sequence; however, ACF are also found in non-CRC populations. An analysis of tissue samples collected from colonoscopy patients found that <5 ACF were present in individuals under the age of 38, which gradually increased to >10 ACF in individuals aged 77 (Kowalczyk *et al.*, 2018). Interestingly, the researchers also found that consumption of a low-fibre diet was associated with

increased prevalence of ACF. Accumulation of genetic damage in ACF leads to the formation of benign adenomas or polyps, which in most sporadic cases protrude from the epithelium into the colonic lumen. Colorectal adenomas mostly display a tubulovillous morphology, although purely tubular and villous morphologies are also present, with varying levels of dysplasia (Calderwood, Lasser and Roy, 2016). Like ACF, colorectal adenomas are also present in the general population, with 25% of men and 15% of women undergoing a colonoscopy displaying detectable polyps (Corley *et al.*, 2013). However, these adenomas only progress to CRC in 10% of cases. The presence of supposed cancer precursors in the healthy population is perhaps not surprising given the high basal levels of proliferation found in the colon. As described by Vogelstein and Kinzler, the adenoma-carcinoma sequence normally occurs over several decades, and therefore it appears that non-genomic factors may control the emergence of ACF and the speed at which ACF progress to adenomas, and subsequently adenocarcinomas (Vogelstein and Kinzler, 1993).

In adenomas, normal cellular migration is reversed, with a reduction in migration towards the colonic lumen and an increase in migration towards the crypt base leading to an aberrant accumulation of cells (Moss *et al.*, 1996). Colorectal adenocarcinomas arise from continued, uncontrolled cell division and increasing dysplasia in adenomas (Fearon and Vogelstein, 1990). The tumour suppressor gene *TP53* is the most commonly mutated gene across all human cancers (Kandoth *et al.*, 2013). The loss of *TP53* inhibits apoptosis in tumour cells and is thought to be a major driver towards malignant transformation in CRC (reviewed by Nakayama and Oshima, 2019). In addition to apoptosis resistance, it is well established that neovascularisation is essential to cancer progression. Early studies of tumour xenografts in mice demonstrated that tumours struggled to reach a size greater than 2-

3mm in diameter in poorly vascularised regions, and subsequent studies found that apoptosis can be triggered in large tumours by suppressing angiogenesis (Muthakkaruppan, Kubai and Auerbach, 1982; Holmgren, O'reilly and Folkman, 1995). Therefore, increased signalling through pro-angiogenic molecules including vascular endothelial growth factor (VEGF) expression is important for tumour progression, and has been suggested as a hallmark of malignancy (Hanahan and Weinberg, 2011). Indeed, *VEGF-1* is upregulated in 50% of all CRC cases, with minimal expression in colorectal adenomas or healthy tissue (Bendardaf *et al.*, 2008). In addition to *TP53* and *VEGF*, mutations in multiple other genes have been linked to malignant transformation in CRC including *Bcl-2*, *TGF-β* and *MUC-1*. However, *TP53* is also frequently mutated in adenomas. Furthermore, Mamlouk and colleagues revealed that mutation of malignancy-associated genes, including *TP53*, is present in paired adenomas and carcinomas which display wildly differently morphologies and levels of dysplasia (Mamlouk *et al.*, 2020). In a study by Wolff *et al.* of 2,204 gene mutations present in 18 CRC patients, only 52 genes were mutated in colorectal carcinomas but were not found to be mutated in adenomas, further undermining genomic alterations as the sole drivers of tumour progression in CRC (Wolff *et al.*, 2018).

Clearly the acquisition of these key driver mutations, and the speed in which successive mutations occur, are significant determining factors as to whether cancer develops, and so factors influencing these genetic changes are of great importance. Many individuals will harbour benign colorectal tumours asymptomatically and will never go on to develop CRC. In patients who develop carcinoma *in situ* but are diagnosed at the earliest stage, 5-year survival rates are greater than 90% (Cancer Research UK, 2021). However, prognosis for patients at advanced stages rapidly

worsens, which can be primarily attributed to the presence of metastases and the difficulty in diagnosing and treating them. Metastasis is responsible for 90% of all cancer-related deaths, and approximately 20% of CRC patients present with synchronous metastases upon diagnosis (van der Geest *et al.*, 2015). However, no specific mutations have been identified as being drivers of invasion and metastasis. This has led to the hypothesis that late-stage disease is primarily driven by the tumour microenvironment (TME).

1.4.3 The unique tumour microenvironment of CRC governs disease progression

Throughout the progression of the adenoma-carcinoma sequence, the tumour microenvironment becomes progressively perturbed. For example, inflammatory and immune cells infiltrate tumours, and the resulting inflammatory signals have been implicated in epithelial-mesenchymal transition (EMT) which facilitates metastatic spread (Bates, Deleo and Mercurio, 2004). The architecture of the vasculature, extracellular matrix and lymphatic system are also greatly modified, physically interacting with the tumour and facilitating the provision of nutrients and the dissemination of invasive cancer cells (Ogino *et al*., 2018). Alterations to the microenvironment are common to almost all tumours, and in CRC these alterations also impact the gut microbiota, which in turn predisposes further changes to the microenvironment in a feedback loop (Li *et al.*, 2020). This two-way interaction between the TME and the microbiota is present throughout the adenoma-carcinoma sequence, and the metagenome-wide associated study performed by Feng *et al.* (2015) confirmed that as CRC develops, compositional and functional changes to the microbiota occur in parallel.

The degree to which the development of cancer influences the composition of the microbiota and vice versa is not well understood, and therefore requires further study. What is clear is that many of the factors known to modulate microbiota composition have also been linked to CRC risk. Diet has emerged as the greatest CRC risk factor, with a 'Western diet' being extensively linked to an increased risk of CRC development (Terry *et al.*, 2001; Mehta *et al.*, 2017; Wu *et al.*, 2004). This diet is characterised by a high intake of red and processed meat, high-fat dairy products and refined foods, whilst being low in fibre and complex carbohydrates (Adlercreutz, 1990). Diet and other environmental factors, as opposed to genetic differences between populations, are thought to explain differences in CRC incidence between geographical regions. American-born Asian men display significantly higher CRC risk than their Asian-born counterparts (Flood *et al.*, 2000). Similarly, migrants to high-income countries from poorer regions assume the CRC risk of their adoptive country within a single generation (Shuldiner, Liu and Lofters, 2018). This adoption of CRC risk is thought to be directly attributed to acculturation and adoption of native dietary patterns. A study of Mexican immigrants to the United States spanning three generations observed correlations between increased CRC risk and dietary change, with both the biggest increase in CRC risk and degree of dietary change observed within the first two generations (Monroe *et al.*, 2003).

In individuals consuming a Western diet, both high meat consumption and a lack of dietary fibre contribute to cancer risk. The relationship between a high consumption of processed and red meats and CRC was sufficient for the World Health Organisation to classify these food products as Group 1 and Group 2 carcinogens, respectively, in 2015 (WHO, 2015). Tumorigenesis caused by high meat consumption is attributed to the production of DNA-damaging metabolites produced following

protein breakdown by microorganisms (Yao, Muir and Gibson, 2016). Furthermore, lack of fibre consumption and subsequent decrease in bacterial fermentation products (e.g. butyrate) has been linked to poor gut health, intestinal dysbiosis and the development of intestinal tumours (Schulz *et al.*, 2014).

In colonocytes, butyrate is primarily metabolised as an energy source. Oxidation of butyrate provides an alternative source of acetyl coenzyme A, which is incorporated into the tricarboxylic acid cycle (Roediger, 1982). Colonocytes which metabolise butyrate are therefore less dependent on glucose from the underlying lamina propria, and have downregulated glycolysis compared to cells not found in the colon (Leschelle *et al.*, 2000). This constant turnover of butyrate prevents its accumulation in the cytosol and therefore prevents it from translocating to the nucleus where it exerts effects on various signalling pathways, notably Wnt signalling, as an HDACi (Hamer *et al.*, 2008). In contrast to normal cells, tumour cells preferentially ferment glucose, even in the presence of abundant oxygen for aerobic respiration in what is described as the "Warburg effect" (Warburg, Wind and Negelein, 1927). Although this represents a far less efficient means of generating ATP, it allows tumours to produce and metabolise small molecules for cell proliferation, such as reduced nicotinamide adenine dinucleotide phosphate (NADPH), more rapidly (Vander Heiden, Cantley and Thompson, 2009). This switch to aerobic glycolysis leads to the accumulation of butyrate in the cytosol and its subsequent translocation to the cell nucleus.
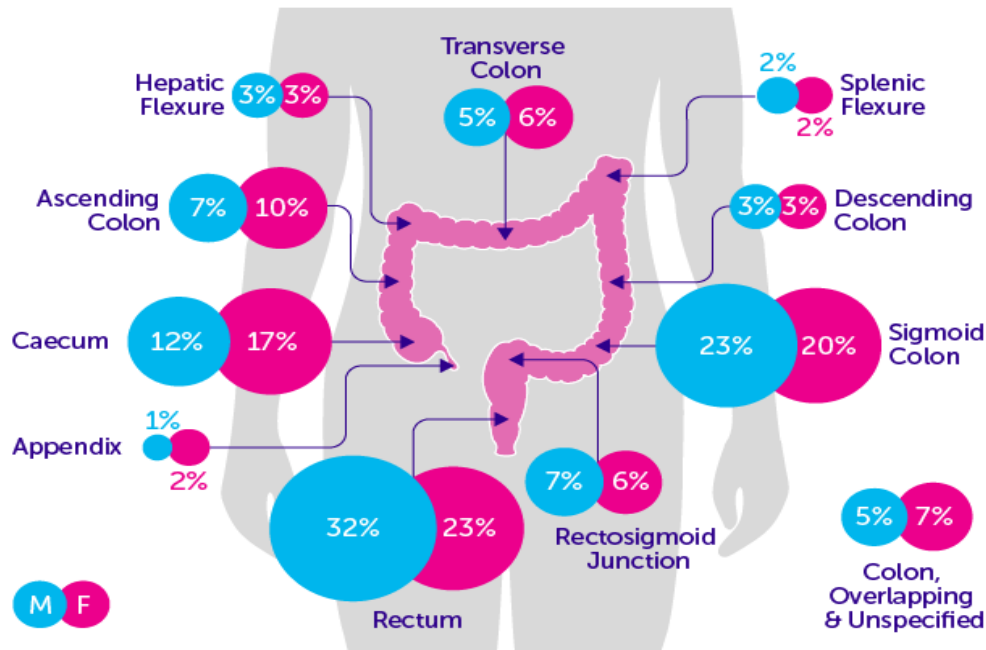
Figure 1.5: Anatomical distribution of colorectal tumours. The overwhelming majority of colorectal tumours are diagnosed distal to the splenic flexure. Tumours in the sigmoid colon, rectosigmoid junction and rectum account for 62% and 49% of tumours in men and women respectively. Tumours in the distal colon are more commonly characterised by chromosomal instability and display a polypoid morphology described by the adenoma-carcinoma sequence (Baran *et al.,* 2018). Figure reproduced and modified with permission from Cancer Research UK.

Once in the nucleus, butyrate regulates DNA transcription through HDACi activity. It has been shown to cause cell cycle arrest and apoptosis in colorectal tumour cells, promote the differentiation of stem cells, supress inflammation through NF-κB inhibition, and promote barrier function by upregulating Claudin-1 transcription (Hague *et al.*, 1993; Segain *et al.*, 2000; Wang *et al.*, 2012). It is for these reasons that butyrate significantly contributes to the protective effect of dietary fibre and is associated with a decreased risk of cancer development, particularly in the distal colon (Kunzmann *et al.*, 2015). The concentration of butyrate is highest in the proximal colon, where the most fermentable substrate is available (Pryde *et al.*, 2002). Strikingly, the incidence of colorectal tumours is highest in the sigmoid colon and rectum, where the concentration of butyrate is lowest (Figure

1.5). Therefore, altering the availability of fibre is one mechanism through which diet exerts a significant effect on cancer risk through the action of microbiota.

1.5 The link between the microbiome and cancer

1.5.1 Colorectal cancer patients exhibit distinct microbiota populations

In healthy individuals the composition of the gut microbiota is relatively stable, with one study finding that approximately 60% of species persisted for the entire 5 year measurement period (Faith *et al.*, 2013). Despite this, drastic alterations to diet, and other factors which impact the colonic microenvironment, have been demonstrated to rapidly alter microbiota composition within a matter of days. David *et al.* (2014) showed that diets high in animal products increased the abundance of bile-tolerant genera including *Biliphila* and *Bacteroides*. In contrast, *Firmicutes* including *Roseburia* and *Eubacterium* were enriched in individuals consuming plant-based diets. Where dietary patterns are sustained, long term drifts in microbiota composition occur, potentially leading to dysbiosis. As tumours develop, changes in mucus secretion, bleeding and lowering of pH in the colonic lumen further impact microbiota composition (Chénard *et al.*, 2020).

In CRC, microbial sampling of faeces, tumour biopsies and adjacent normal tissues has revealed distinct alterations in microbiota composition between CRC patients and healthy controls (Flanagan *et al.*, 2014; Mori *et al.*, 2018; Saffarian *et al.*, 2019). In a study comparing the faecal microbiome of patients diagnosed with hyperplastic polyps, low and high-risk adenomas and adenocarcinomas, Mori *et al.* (2018) identified stage-specific bacterial populations associated with each stage of the adenoma-carcinoma sequence. For example, *Firmicutes* and *Actinobacteria* were found at a high frequency in patients with pre-neoplastic lesions; whereas *Enterobacteriaceae* and *Escherichia* amongst others were found in patients with malignant lesions (M*ori et al*., 2018). A more recent study which used 16S rRNA sequencing to compare mucosa and crypt-associated microorganisms between CRC

patients and healthy controls revealed that *Fusobacterium nucleatum* and *B. fragilis* were enriched in CRC patients with both left and right-sided tumours (Saffarian *et al.*, 2019). These compositional shifts can also be linked to CSMs, highlighting the relationship between microorganisms and CRC disease progression, with oral pathogens including *F. nucleatum* and *Porphorymonas gingivalis* being found at greater abundance in CSM1 patients, and *Selenomas* species being more common in CSM2 cohorts (Purcell *et al.*, 2017).

Understanding microbiome alterations has the potential to revolutionise CRC screening; however, studying the involvement of bacteria in CRC presents some challenges. Early studies employed culture-based approaches to identify species present in faecal samples. However, the advent of 16S rRNA sequencing revealed that less than 30% of the gut microbiota has been cultured *ex vivo* (Qin *et al.*, 2010). Furthermore, faecal samples have long been preferred to direct mucosal biopsies due to their non-invasive nature. However, direct comparison between mucosal and faecal microbial communities has revealed key differences in both composition and function (Zoetendal *et al.*, 2002). While levels of *Firmicutes* are similar between faecal and mucosal communities, the mucosal microbiota is enriched in *Bacteroides* species, which provide immune system-regulatory functions, and carotenoid-producing species such as *Akkermansai muciniphila* (Vaga *et al.*, 2020). In contrast, *Bifidobacterium* are far more prevalent in faecal samples. Despite these challenges, some insight into the compositional microbiota changes seen during dysbiosis has been gained. Dysbiotic microbiotas are generally thought to exhibit a relative decrease in the number of *Firmicutes,* which are sensitive to changes in carbohydrate intake, corresponding with an expansion of the *Bacteroides* population (Mori *et al*., 2018). Crypt-associated microbiomes are also altered in CRC patients when compared to

healthy controls, exhibited by the overrepresentation of pathobionts (Saffarian *et al.*, 2019). This potentially exposes the sensitive stem cell niche to genotoxic compounds produced by both the pathogenic bacteria and infiltrating immune cells.

To date, many trends in microbiome composition have been linked to CRC aetiology, with individual species, or specific strains, emerging as potential carcinogenic pathogens. The species relevant to this body of work shall be discussed in detail below.

1.5.2 *Bacteroides fragilis*

Bacteroides species account for roughly 50% of the microorganisms found in faeces, and are also present in abundance at the mucosal surface (Wexler, 2007). They are anaerobic, Gram-negative rods and play important roles in host immunity and nutrition, as reviewed by Wexler (2007). *Bacteroides fragilis,* which comprises approximately 0.5% of the gut microbiota, is found in association with host mucus, and as previously mentioned is able to reside within colonic crypts (Huang, Lee and Mazmanian, 2011). *B. fragilis* is considered a symbiotic species, and is thought to promote the development of CD4$^+$ T-cell populations (Mazmanian *et al.*, 2005). Despite this, *B. fragilis* is also an adept opportunistic pathogen, and is the most frequently isolated anaerobic bacterium found in sepsis patients (Tan *et al.*, 2017). *B. fragilis* possesses multiple virulence factors, and approximately 10% of adult individuals harbour enterotoxigenic *B. fragilis* (ETBF), which is associated with diarrheal disease (Zhang and Weintraub, 1999). The virulence of these enterotoxigenic strains has been attributed to the *Bacteroides fragilis* toxin (BFT), a zinc-dependent metalloproteinase which causes diarrhoea by promoting ileal and colonic secretion of sodium and chloride ions (Myers *et al.*, 1985).

Despite *B. fragilis* being isolated from the stool of healthy controls in roughly equivalent amounts, using polymerase chain reaction (PCR) to analyse cultured stool isolates, Toprak *et al.* (2006) discovered that BFT was present in 38% of CRC patients compared to just 12% of healthy controls. ETBF virulence through the BFT is attributed to an increase in host epithelial cell membrane permeability. BFT cleaves E-cadherin, resulting in numerous downstream pro-tumorigenic effects (Rhee *et al.*, 2009). E-cadherin is an integral component of the zonula adherens, the integral junctions between epithelial cells (Hartsock and Nelson, 2008). Degradation of E-cadherin also releases β-catenin from the intracellular domain, which is a key effector of the Wnt signalling pathway. Subsequent downstream signalling promotes colonocyte proliferation through the c-myc pathway (Wu *et al.*, 2003). This pro-tumorigenic effect is observed at BFT concentrations of as little as 5×10^{-10} mol/L (Wu *et al.*, 2003). Furthermore, BFT has been shown to induce persistent colitis in murine models, as the reduced barrier function and subsequent stimulation of inflammatory cytokine secretion drastically alters the colonic microenvironment (Rhee *et al.*, 2009). This chronic inflammation further contributes to CRC-promoting DNA damage.

1.5.3 Escherichia coli

Escherichia coli, also a Gram-negative rod, is the most abundant facultative anaerobe found in the gastrointestinal tract, and is easily isolated from both faecal and mucosal samples (Wassenaar, 2018). It is an important commensal, and one of the first species to colonise infant gastrointestinal tracts; however, some strains possess virulence factors which are associated with intestinal and extra-intestinal disease (Benno, Sawada and Mitsuoka, 1984). The species can be generally divided into four main phylotypes, A, B1, B2 and D, with almost all commensal species belonging to group

A. Virulent strains are most commonly found in group B2, with a small proportion in group D (Herzer *et al.*, 1990).

Pathogenic *E. coli* strains have been linked to an increased risk of CRC, although whether these strains are causative in nature, or are incidentally increased in CRC patients is unclear (Wassenaar, 2018). Many strains of enteropathogenic *E. coli* (EPEC) possess virulence genes encoding cyclomodulin toxins, which are capable of causing double-stranded DNA breaks in eukaryotic cells (Nougayrède *et al.*, 2005). Several cyclomodulin genes are expressed by EPEC strains, including cytolethal distending toxin, cytotoxic necrotizing factor and cycle inhibiting factor. The prevalence of cyclomodulin-positive EPEC isolated from CRC patients varies by study; however, colibactin, encoded by the polyketide synthesis (*pks*) locus has emerged as the most significant EPEC toxin in CRC (Dalmasso *et al.*, 2015). Strains of colibactin-producing EPEC are significantly more prevalent in CRC patients, despite these strains showing lesser adhesive properties than commensal strains (Buc *et al.*, 2013).

Initial studies on the interaction between cyclomodulins and the colonic epithelium indicated that the CRC-promoting effect was limited to cells in direct contact with the toxin (Nougayrède *et al.*, 2006). However, more recent studies have revealed that intestinal epithelial cells treated with *pks+ E. coli* develop a senescence-associate secretory phenotype (SASP) characterised by the secretion of inflammatory cytokines, proteinases and chemokines (Secher *et al.*, 2013). These cells were able to promote double-stranded DNA breaks and cellular proliferation in neighbouring cells through a bystander effect, primarily through the induction of hepatocyte growth factor signalling (Cougnoux *et al.*, 2014). Therefore, it is now clear that colibactin is able to contribute to tumour growth without the requirement for direct cell contact.

The importance of colibactin expression to the mechanism of *E. coli* tumour promotion remains unclear. The *pks* locus is expressed by other members of the *Enterobacteriaceae;* however, as of yet no other family member has been linked to CRC (Putze *et al.*, 2009). The relevance of *pks+ E. coli* to CRC may be attributed to the ability of these strains to colonise and invade the epithelium of CRC patients. In one study, intracellular *E. coli* were recovered both from within tumours and on matched normal tissue in >90% of patients with colorectal adenomas or carcinomas, but were absent from healthy controls (Swidsinski *et al.*, 1998). Interestingly, invasive *E. coli* were found to be more frequently isolated from mice fed a Western diet (Martinez-Medina *et al.*, 2014). However, the *pks+* locus is also found in some probiotic strains, particularly *E. coli* Nissle 1917 (EcN). This strain is used commercially to treat a number of gastrointestinal conditions, including Crohn's disease and ulcerative colitis, and has been demonstrated to produce anti-inflammatory effects and prevent the colonisation of pathogens including *Salmonella* (Deriu *et al.*, 2013). Remarkably, invasive *E. coli* (including EcN) preferentially locate and invade tumour tissue even when injected intravenously, which has the potential to be exploited as a cancer diagnostic/therapy delivery tool (Brader *et al.*, 2008). However, it has been demonstrated that mutation of the *pks* locus in EcN abrogates its probiotic activity, the mechanism of which is still unclear (Olier *et al.*, 2012). Our understanding of the contribution of *pks+ E. coli* strains is further hindered by the inability to determine the structure of colibactin (Faïs *et al.*, 2018), and it is clear that further work is needed to elucidate the role of *E. coli* in CRC.

1.5.4 *Enterococcus faecalis*

Enterococcus faecalis is an early coloniser of the gastrointestinal tract. Belonging to the *Firmicutes*, *E. faecalis* is a Gram-negative, facultative anaerobe coccus which

despite being predominantly found in the gut is also isolated from the oral cavity and vaginal mucosa (Alhajjar *et al.*, 2020). It is one of the most commonly isolated bacteria from stool, usually in the range of 10^5-10^7 colony forming units per gram, and like *E. coli* Nissle, strains of *E. faecalis* such as EC-12 are frequently used as probiotics (Sghir *et al.*, 2000). During early gut microbiome development, *E. faecalis* plays a crucial role in suppressing pathogen-associated immune responses through its regulation of the inflammatory cytokines IL-8 and IL-10 via MAPK signalling pathways (Wang *et al.*, 2014). Despite these positive effects, elevated levels of *E. faecalis* are isolated from the stool of CRC patients, leading to some speculation as to its role in cancer development (Balamurugan *et al.*, 2008).

E. faecalis is thought to contribute to CRC development and progression through the production of reactive oxygen species (ROS) and other genotoxic metabolites. It is one of relatively few species of bacteria which produces extracellular superoxide (O_2^-), and also produces hydrogen peroxide (H_2O_2; Huycke, Abrams and Moore, 2002). Hydrogen peroxide freely diffuses across epithelial cell membranes, and has been demonstrated to induce epidermal growth factor receptor (EGFR) signalling (Boonanantanasarn *et al.*, 2012). As previously mentioned, chromosomal instability is a key mechanism through which genetic changes accumulate in CRC. Superoxide production by *E. faecalis* has been demonstrated to promote chromosomal instability in mammalian cells, both directly and through the induction of the COX-2 pathway in neighbouring macrophages (Wang and Huycke, 2007). In addition to the production of ROS, *E. faecalis* also produces and secretes a matrix metalloproteinase toxin, gelatinase. Action by gelatinase increases epithelial barrier permeability, allowing microorganisms and other luminal contents to access the underlying stroma, causing inflammation in the colon (Steck *et al.*, 2011). Mouse studies of E-cadherin

have identified cleavage sites for gelatinase, suggesting that this toxin could also contribute to Wnt signalling as previously discussed.

1.5.5 *Fusobacterium nucleatum*

The obligate anaerobe *Fusobacterium nucleatum* is a Gram-negative spindle-shaped rod and is not considered to be part of the normal gut microbiota. In a study of inflammatory bowel disease (IBD) related microorganisms, Huh and Roh found that *F. nucleatum* was present in less than 3% of samples taken from a mix of IBD and non-IBD subjects, highlighting its low prevalence in the normal population (Huh and Roh, 2020). Instead, *F. nucleatum* is an important constituent of the oral microbiome, where it is a strong biofilm former and acts as a bridge between stricter anaerobes at the mucosal surface and facultative anaerobes in the apical plaque layer (Okuda *et al.*, 2012). Despite being a member of the healthy oral microbiome, *F. nucleatum* is implicated in periodontal disease, as well as extra-oral conditions including adverse pregnancy outcomes, IBD and appendicitis (Han *et al.*, 2000; Wang *et al.*, 2013; Loening-Baucke *et al.*, 2012).

An increased prevalence of *F. nucleatum* in CRC patient biopsies was first reported in 2012 (Castellarin *et al.*, 2012), and was surprising, given that *F. nucleatum* was thought to primarily be an opportunistic pathogen of the oral cavity. To this date, the precise mechanism of translocation to the gut is unknown. Due to the concordance between oral and intestinal microbiota populations an oral-gut route is predicted (Ding and Schloss, 2014). However, the translocation of *F. nucleatum* to non-gastrointestinal tumours suggests that hematogenous spread is possible (Parhi *et al.*, 2020). *F. nucleatum* is more prevalent in IBD and CRC patients than healthy controls, and it is thought to remodel the gut microbiota in CRC by promoting the infiltration of

myeloid-derived immune cells which trigger erroneous immune responses against the microbiota (Kostic *et al.*, 2013).

F. nucleatum abundance is increased 4-fold in CMS1 CRC, however, it is found in lower proportions in CSM2 and CSM3 populations compared to healthy controls (Purcell *et al.*, 2017). Furthermore, there are conflicting reports on the association of *F. nucleatum* with various stages of CRC. *F. nucleatum* is over-represented in CRC patients, and has been frequently isolated from both lymph node and liver metastases, suggesting that this bacterium is relevant to late-stage CRC progression and is able to travel to secondary sites (Bullman *et al.*, 2017). *F. nucleatum* was also demonstrated to be enriched in colorectal adenomas (Flanagan *et al.*, 2014). Furthermore, isolation of *F. nucleatum* from stool samples has been shown to diagnose both colorectal adenomas and adenocarcinomas at high sensitivity and specificity (Wong *et al.*, 2017). Conversely, other research groups did not find any significant association between *F. nucleatum* and colorectal adenomas, and have suggested that the overabundance of *F. nucleatum* in CRC can be simply attributed to its adaptation to favourable microenvironment conditions (Amitay *et al.*, 2017). The increased abundance of *F. nucleatum* in IBD patients suggests a contribution to chronic colonic inflammation (Strauss *et al.*, 2011). Indeed, colonisation with *F. nucleatum* has been linked with the upregulation of numerous inflammatory pathways, including the secretion of inflammatory cytokines, infiltration of immune cells into colon tissue and stimulation of humoral immunity (Wu, Li and Fu, 2019).

F. nucleatum possesses multiple virulence mechanisms; however, its surface adhesin, FadA, is thought to facilitate direct pro-tumorigenic activity. FadA is found on the surface of *F. nucleatum* cells and regulates its attachment and invasion of the colonic epithelium (Guo *et al.*, 2020). Like BFT, FadA binds to E-cadherin, and is

thought to contribute to CRC progression by stimulating proliferation, tumour migration and epithelial-mesenchymal transition by modulating Wnt signalling through E-cadherin cleavage (Rubinstein *et al.*, 2013). Interestingly, *F. nucleatum* invasion of colorectal adenocarcinoma cells did not appear to correlate with relative levels of E-cadherin expression in multiple cell lines, suggesting that the interaction of FadA with E-cadherin is not the limiting factor on *F. nucleatum* association with the epithelium (Rubinstein *et al.*, 2013). Further studies from the Rubinstein research group have revealed that *F. nucleatum* modulates Wnt signalling through a secondary, albeit still E-cadherin-dependent mechanism, by upregulating Annexin A1 expression (Rubinstein *et al.*, 2019). Therefore, despite growing evidence suggesting a causal role for *F. nucleatum* in CRC progression, further research is required.

1.6 Models of bacterial contribution to colorectal cancer

The concept of cancer as a microbial disease is not novel. It is estimated that approximately 30% of all cancerous diseases can be attributed to microorganisms. For example, *Helicobacter pylori* is thought to be responsible for approximately 89% of non-cardia gastric cancer cases (Plummer *et al.*, 2015). The strong associations between infectious agents and cancer have allowed for novel interventions to prevent and diagnose cancerous diseases. Successful examples of this include vaccinations against human papillomavirus (primarily for the prevention of cervical cancer), and the urea breath test for gastric cancer screening (Chrysostomou *et al.*, 2018; Rahim *et al.*, 2019). Recent research has demonstrated the potential for antimicrobial adjuvant therapy in the treatment of CRC. For example, small molecule inhibitors of the colibactin synthesis enzyme ClbP effectively eliminated DNA damage and tumour proliferation induced by colibactin-producing *E. coli* in murine models (Cougnoux *et*

al., 2016). However, before microbial treatment can be fully utilised in CRC treatment some obstacles remain.

Unlike other microbe-associated cancers, the contribution of microorganisms to CRC initiation and progression is thought to be polymicrobial. Interpersonal variation in gut microbiota composition means that microbiome studies conducted on different populations, and in different geographical areas, often report different species of bacteria in association with CRC (Osman *et al.*, 2018). It is currently unclear whether these differences can be attributed to physiological differences (e.g. diet, host genetics) or inconsistencies between handling of samples and method of analysis. In addition to regional differences between CRC patients, temporal associations between oncogenic bacteria and CRC patients are also apparent. Many species isolated in abundance from tumour biopsies are absent, or found at significantly reduced abundancies in paired off-tumour tissue from the same patient (Flemer *et al.*, 2017). Similarly, colorectal adenomas appear to display microbial populations distinct from those of advanced tumours (Mori *et al.*, 2018). This supports the notion that just as microbial factors are able to influence tumour progression, the tumour microenvironment associated with different stages of CRC progression favours the growth of specific bacterial populations. In order to reconcile these issues in understanding the role of bacteria in CRC, several conceptual models have been put forward.

1.6.1 The alpha-bug hypothesis

One of the key issues with associating specific species of bacteria with CRC is that many candidate pathogens, despite being found in greater abundance in cancer patients, are still relatively low in abundance in the microbiome as a whole. The alpha-bug hypothesis (Sears and Pardoll, 2011) posits that certain low-abundance species

exert pro-oncogenic effects sufficient to remodel the microbiome and directly initiate cancer formation (Figure 1.6A). It was first proposed based upon studies of ETBF, and a key tenet of this model is the tripartite relationship between the microbiome, the colonic mucosa, and the immune system. Therefore, a key characteristic of potential alpha-bug species is the ability to exert an influence on the immune system and modulate the composition of the gut microbiota. In the multiple intestinal neoplasia (Min) mouse model, ETBF colonisation was sufficient to drastically alter the tumour formation phenotype (Wu *et al.*, 2009). Tumour development in Min mice is normally restricted to the small intestine, and animals usually perish within six months. In contrast, ETBF colonised mice developed multiple large tumours in the distal colon, the formation of which correlated with ETBF-induced colonic inflammation, and which led to premature death after 3 months (Wu *et al.*, 2009).

Prospective alpha-bugs must be able to exert their effects despite being at relatively low abundance in the microbiome. As previously stated, *B. fragilis* constitutes less than 1% of the gut microbiota, yet its regulation of Th17 immune cells is sufficient to cause dysbiosis (Sears and Pardoll, 2011). This is in contrast to 'dominant pathogens' which outcompete the native microbiota during their pathogenesis (Hajishengallis, Darveau and Curtis, 2012). For example, during infection *Salmonella enterica* subsp. *enterica* serovar Typhimirium, suppresses the microbiota by triggering IL-10 mediated inflammation and rapidly becomes the predominant species in the gut after 4 days in murine models (Stecher *et al.*, 2007). Other potential alpha-bugs in CRC include the previously discussed enteropathogenic *E. coli* and *E. faecalis* in addition to *Streptococcus gallolyticus* subsp. *gallolyticus* (Yu and Fang, 2015).

Despite the strengths of the alpha-bug, or more recently coined 'keystone pathogen' hypothesis, some issues remain when applying it to CRC. The alpha-bug model does not account for asymptomatic colonisation by the candidate organism, which in the case of ETBF occurs in 5-30% of individuals (Sears, 2009). Furthermore, several CRC-associated pathogens are often absent from malignant adenocarcinomas, suggesting that these species are not responsible for driving late-stage disease (Bundgaard-Nielsen *et al.*, 2019).

1.6.2 The Driver-Passenger model

The divergence between bacterial species present at different stages of the adenoma-carcinoma sequence, and between on-tumour and off-tumour samples, is explained by the Driver-Passenger Model (Tjalsma *et al.*, 2012). This theory, which has been frequently cited since its publication, suggests that as the tumour develops and progresses, its associated microbiota also experiences shifts in its composition. Therefore, bacteria linked to CRC can be divided into two categories: Driver or Passenger (Figure 1.6B). Driver bacteria are members of the core microbiome which possess pro-oncogenic properties and are able to directly initiate tumorigenesis, whether through induction of inflammation or direct virulence against the colonic mucosa. Due to natural variation in individuals' microbiota compositions, some individuals will naturally have a higher proportion of driver species than others, and these differences can be exacerbated by lifestyle factors such as consuming a Western diet. The action of these driver species both directly and indirectly modifies the tumour microenvironment. For example, increased epithelial barrier permeability caused by previously mentioned toxins produced by ETBF or *E. faecalis* may provide additional niches for the colonisation of other species. For example, *S. gallolyticus* poorly attaches to and invades colon epithelial cells *in vitro*, but readily attaches to and forms

biofilms on collagen fibres. Hyperplasia induced by other species, and eventual tumour development greatly alters the microenvironment and may expose collagen fibres for colonisation by *S. gallolyticus* (Boleij *et al.*, 2011). Mucosal metabolism also greatly affects bacterial viability, whereby colorectal tumours display distinct metabolic profiles from healthy epithelium (Hirayama *et al.*, 2009). The altered availability of substrates to mucosa-resident bacteria applies a selection pressure, favouring bacteria adapted to survive in this new environment. As a result, the initial driver bacteria are eventually outcompeted, and replaced by passenger bacteria which are better adapted.

Passenger bacteria have a competitive advantage in the tumour microenvironment and are therefore enriched at tumour sites, these may include *F. nucleatum*, which is a poor coloniser of the healthy gastrointestinal tract. However, unlike driver bacteria, passenger bacteria may not be pro-oncogenic, and may simply be commensal or even exert anti-tumour effects. For example, *Coriobacteriaceae* are enriched in CRC samples, despite being considered probiotic species (Drewes *et al.*, 2017). This notion has been supported by further studies which have successfully categorised Passenger bacteria according to whether they exert a pro-inflammatory or anti-inflammatory effect (Geng *et al.*, 2014). Many commensal passenger species may not directly contribute to colorectal carcinogenesis, but may have altered transcription in the tumour microenvironment that contributes to dysbiosis (Hajishengallis and Lamont, 2016).

Passenger populations themselves are considered transient, as they too are subject to being outcompeted as the tumour progresses further. This poses a problem with the application of microbiome studies to CRC diagnosis and treatment, as Driver species may only be active during the very early stages of disease development.

However, this model does explain the conflicting results of studies analysing the role of specific species in CRC. Tjalsma *et al*. also conclude that the role of an individual species as a cancer driver or passenger may be determined at a strain-specific level and be temporally associated with disease stage. Therefore, investigation of how bacteria interact with tumour cells of different stages would be beneficial in understanding this relationship.

Although the driver-passenger model expands upon the alpha-bug hypothesis, it raises further questions about the role of specific species in CRC development and progression. Firstly, it is not yet possible to conclude that initial Driver species are completely lost from the microbiota population, as despite advances in the sensitivity of sequencing technologies it cannot be guaranteed that bacteria do not remain present outside the limits of detection (both spatially and analytically). It is also important to consider that changes in the abundance of certain species are not enough to confirm a positive or negative effect on CRC. For this reason, RNA sequencing, in addition to tried and tested DNA sequencing has become more common. This allows investigators to determine whether species identified as being associated with tumours are transcriptionally active (Castellarin *et al*., 2012). It has been demonstrated that biofilms isolated from CRC patients and matched healthy controls exert pro-oncogenic effects after injection into murine models (Tomkovich *et al.*, 2019). This suggests that the phenotypic switches, such as biofilm formation, may play a role in tumour initiation. However, the realistic feasibility of being able to undertake longitudinal microbiome studies to monitor the development of CRC in previously healthy populations means that definitively identifying bacteria that are cancer initiators remains challenging.

(A) Alpha-bug (2011) and keystone pathogen (2012) models

Proliferation of alpha-bug or keystone pathogen (e.g., ETBF)

Colonic microbiota remodeling

Th17 inflammation

Genotoxicity

Host genetic defect

Epithelium proliferation

Pro-carcinogenic state

Hyperplasia

Carcinoma

(B) Dysbiosis (passenger bacteria emerging, e.g., Fusobacteria)

Host genetic defect

Colonization of driver bacteria (e.g., ETBF, Enterobacteriaceae)

Genotoxicity

Accumulation of germline and somatic mutations

Epithelium proliferation

Epithelium transition

Th17 inflammation

Epithelial barrier defect

Dysbiosis (passenger bacteria emerging, e.g., Fusobacteria)

Passenger bacteria gradually increase

Passenger bacteria outcompete driver bacteria

Bacterial community shift during CRC initiation

Normal epithelium

Hyperplasia

Adenoma

Carcinoma

Figure 1.6: A comparison of the 'Alpha-bug' and 'Driver Passenger' models of bacterial contribution to colorectal cancer. (A) An 'alpha-bug'/'keystone pathogen' such as ETBF induces tumorigenesis through a variety of mechanisms. Remodelling of the microbiota and induction of Th_{17} cell inflammation form a feedback loop, which combined with direct genotoxicity (e.g. through BFT) and host mutations contributes towards a pro-carcinogenic colonic microenvironment. (B) In the Driver-Passenger model, remodelling of the microbiota by a Driver species such as ETBF facilitates colonisation by Passenger bacteria e.g. *Fusobacteria*. These Passenger bacteria may be pro-tumorigenic or may be commensal. Eventually, sustained changes to the colonic microenvironment mean that initial driver species are outcompeted by passenger bacteria which are better adapted to the tumour microenvironment. Figure reproduced from Li *et al.* (2017). License number: 5055851464946.

46

1.7 The two-way relationship between the adenoma-carcinoma sequence and the microbiome

Studies have shown that distinct microbial populations can be associated with specific stages of tumorigenesis (Sheng *et al.*, 2019). These stages are characterised by changes in the tumour microenvironment, colonic structure, tumour metabolism, immune status, and hypoxia. It is therefore possible that certain bacterial species exert differing effects on tumours at different points during tumour development. Nakatsu *et al.* (2015) demonstrated that low-grade dysplasia, high-grade dysplasia, early-stage CRC, and late-CRC each have distinct microbial communities. For example, *Fusobacterium*, *Parvimonas* and *Gemella* species increased in CRC relative to polyps displaying low-grade dysplasia, whereas levels of *Bacteroides* remained similar throughout (Nakats*u et al*., 2015). Therefore, an acute understanding of the role of specific species at each stage of the adenoma-carcinoma sequence is required. A key aspect of sporadic colorectal tumorigenesis is the development of cancer over long time periods, usually decades, and many individuals will harmlessly develop colorectal polyps which never progress into cancer (Pan *et al.*, 2020). As previously discussed, development of invasive disease from a benign adenoma has not been linked to any specific mutation. For example, genomic alterations such as inactivating mutations of *TP53*, while largely considered to be markers of malignant carcinoma are also seen in late-stage adenomas (Mamlouk *et al.*, 2020). Therefore, bacteria able to promote the benign-malignant transformation may have the greatest impact on cancer development.

Studying the induction of CRC by bacteria from normal epithelium is difficult *in vitro*, and *in vivo* models are not without issue. It has previously discussed how, although valuable, studies focussing on the changing population dynamics of bacterial species in CRC patients do not necessarily demonstrate a causal relationship. Mouse

models of tumorigenesis allow the investigation of much of the CRC timeline, from tumour formation to death, but are not without their nuances. Due to natural intertumour variation, it is often not possible to investigate metastasis in mouse models before the tumour burden causes death (Khanna and Hunter, 2005). In addition, research on the role of bacteria in CRC using mouse models has primarily been performed using germ-free animals, which possess no indigenous microbiota (Gordon and Pesti, 1971). Immune system development in these animals is therefore severely compromised, as is the tripartite relationship between the immune system, inflammation and CRC described as critical by the alpha-bug model (Sears and Pardoll, 2011). For this reason, some researchers have begun to incorporate conventional mice into their studies, to great success. Wong *et al.* (2017) demonstrated that gavage of stool samples from CRC patients into germ-free or conventional mice induced intestinal carcinogenesis, demonstrating that the microorganisms present in the stool were able to induce dysbiosis even when challenged by an established microbiome.

Tumour cell lines provide a cheaper, low-maintenance, high throughput alternative to mouse models, and allow for the direct study of single species interactions with colorectal tumour cells. However, culturing and maintaining benign adenoma cells *in vitro* has generally proved challenging. Despite this, Paraskeva *et al.* (1989) succeeded in isolating a number of colorectal adenoma cell lines. In particular, the RG/C2 cell line was isolated from a tubular adenoma and is relatively easy to manipulate *in vitro*. Unlike adenocarcinoma cell lines, RG/C2 cells display many features indicative of normal colonic mucosa, including microvilli and the formation of mucin droplets (Paraskeva *et al.*, 1989). As previously discussed, changes in cell surface architecture and the presence of mucin are likely to significantly impact

interactions of bacteria with tumour cells. In addition, the authors note that RG/C2 cells are clonogenic, which is indicative of high malignant potential. Therefore, comparison of this cell line with conventional adenocarcinoma cells enables investigation into the relationship between bacteria and distinct stages of CRC development, potentially identifying mechanisms through which bacteria initiate benign-malignant transformation.

1.8 Aims and Objectives

1.8.1 Aims

Through comparison between RG/C2 cells and well-established colorectal adenocarcinoma cell lines, this body of work aims to investigate how specific bacteria interact with colorectal adenoma and adenocarcinoma cells *in vitro*. Once these interactions have been characterised, this study will aim to determine the impact of these bacteria on cellular behaviours associated with tumour progression. Specifically, the ability of bacteria to moderate tumour cell proliferation, metabolism and motility shall be measured, with a focus on comparing how bacteria may alter these behaviours differently in benign and malignant cells.

1.8.2 Objectives

- To develop an *in vitro* co-culture model which allows the treatment of colorectal tumour cells with bacteria under standard aerobic culture conditions. In addition, to quantify levels of bacterial association with tumour cells during co-culture.

- To assess the effects of bacterial infection on tumour cell yield and apoptosis, and investigate how any differences observed in these characteristics can be explained by bacteria-induced changes to the cell cycle.

- To investigate how tumour cell metabolism differs between adenoma and adenocarcinoma cell lines by investigating the Warburg effect using the Seahorse Xf analyser to measure oxidative phosphorylation and glycolysis both under standard culture conditions and when challenged with bacteria. Similarly, selected ion flow tube mass spectrometry (SIFT-MS) shall also be used to compare the volatile metabolites produced by adenoma and adenocarcinoma cells *in vitro*, and how these metabolites are altered during infection by bacteria.

- To assess how infection with bacteria impacts tumour cell motility using both wound healing and transwell filter migration and invasion assays. The mechanisms behind observed effects shall be explored, firstly by targeting commonly upregulated pathways e.g. Wnt signalling, and secondly by using qPCR microarrays to rapidly quantify the effect of bacteria on the expression of EMT and cell motility-related genes. The results of these qPCR assays will indicate which mechanisms may be involved in bacterial contribution to tumour progression

Chapter 2: Materials and Methods

2.1 Culture of colorectal cell lines

Throughout this study, the principle cell lines used were the colon adenoma epithelial cell line RG/C2 (Paraskeva *et al.*, 1989) and the colon adenocarcinoma epithelial cell line HCT116 (Brattain *et al.*, 1981). The RG/C2 cell line was derived from a tubular adenoma isolated in a sporadic CRC patient, and unlike most colorectal adenoma cell lines is clonogenic at early passages. The HCT116 cell line was initially isolated from a colonic carcinoma, and is tumorigenic in athymic nude mice. Both cell lines display a polygonal appearance *in vitro* typical of epithelial cells. A summary of all colorectal cell lines used throughout this project is presented in Table 2.1. All cell lines were kindly donated by Professor Ann Williams at the University of Bristol, UK.

Table 2.1: Colorectal tumour cell lines used throughout this study.

Cell line	Commercial reference	Derivation	Origin	Notable mutations and CMS	Reference
HCT116	*ATCC CCL-247™*	Colorectal adenocarcinoma	Adult male (age not specified)	*KRAS, PIK3CA* – CMS4	Brattain *et al.* (1981)
HT29	ATCC HTB-38™	Colorectal adenocarcinoma	44-year-old female	*APC, BRAF, p53,PIK3CA* – CMS3	Fogh and Trempe (1975)
RG/C2	*N/A*	Colonic tubular adenoma	59-year-old male	Wild type for *APC, KRAS, BRAF* and *p53* – CMS unknown.	Paraskeva *et al.* (1989)
SW480	*ATCC CCL-228™*	Duke's B colorectal adenocarcinoma	50-year-old male	*APC, KRAS, p53* – CMS4	Fogh *et al.* (1989)

2.1.1 Standard tissue culture procedure

Cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) – high glucose (*Sigma-Aldrich Ltd., Gillingham, UK*) in T25cm^2 filter-capped tissue culture flasks (*Corning® Amsterdam, The Netherlands*). Media used to culture adenoma cells was supplemented with 20% heat-inactivated foetal bovine serum (FBS; *Fisher Scientific UK Ltd., Loughborough, UK*), 4 mM L-glutamine (*Sigma-Aldrich Ltd.*), 0.2 U/ml human insulin (*Sigma-Aldrich Ltd.*), 1 μg/ml hydrocortisone 21-hemisuccinate (*Sigma-Aldrich Ltd.*) and 200 U/ml Penicillin-Streptomycin (*Thermo-Fisher Scientific, Massachusetts, USA*). Media used to culture adenocarcinoma cell lines was supplemented with 10% FBS, 2 mM L-glutamine and 200 U/ml Penicillin-Streptomycin. Cells were maintained in a humidified incubator at 37°C and 5% CO_2 (*NuAire DH Autoflow™, NU 500 series, Triple Red Ltd., UK*) and media was changed twice weekly.

To subculture, cells were passaged once they reached approximately 80-90% confluency by washing in 5 ml of Dulbecco's A phosphate buffered saline (PBS; *Thermo-Fischer Scientific*), followed by trypsinisation with 3 ml of 0.1% w/v trypsin-EDTA solution (*Sigma Aldrich Ltd.*) in PBS for 10 minutes at 37°C and 5% CO2. Detached cells were collected using an equal volume of fresh complete media and centrifuged at 1,000 rpm for 10 minutes to remove trypsin (*OptimaXP centrifuge, Beckman Coulter Life Sciences, Indianapolis, USA*). The resulting cell pellets were then resuspended in an appropriate volume of fresh cell culture media and seeded into new flasks. Adenoma and adenocarcinoma cell lines were split at ratios of 1:3 and 1:10, respectively.

2.1.2 Cell line cryopreservation

To create frozen stocks of colorectal tumour cell lines, sub-confluent cells were trypsinised as described in Section 2.1.1, followed by centrifugation at 1,000 rpm for 10 minutes. Cells were resuspended in an appropriate volume to give approximately 2×10^6 cells/ml, and 900 µl of this cell suspension was aliquoted into a cryovial, to which 100 µl of dimethyl sulfoxide (DMSO; *Sigma-Aldrich Ltd*.) was added. Cryovials were placed into a 'Mr. Frosty' cell freezing container (*Nalgene®, Sigma-Aldrich Ltd*.) and stored at -80°C. After 24 hours, cryovials were deposited into a liquid nitrogen freezer and kept at -136°C. For retrieval, cryovials were removed from liquid nitrogen and warmed to 37°C. The cell suspension was centrifuged at 1,000 rpm for 10 minutes and resuspended in fresh medium to remove residual DMSO. The contents were then added to a T25 flask containing 5 ml of fresh medium. After retrieval from liquid nitrogen storage, cells were grown to 80-90% confluency and allowed to recover for two passages before use in experiments.

2.1.3 Cell enumeration

To seed experiments, cells were trypsinised and resuspended as previously described. According to Hague *et al*. (1993), colonic tumour cells cultured *in vitro* detach from the monolayer as they undergo apoptosis. Therefore, all cells collected during trypsinisation will be viable. Thus, the total viable cell count was achieved by counting a 10 µl sample of the cell suspension using a haemocytometer. Once counted, cells were centrifuged for 10 minutes at 1,000 rpm and resuspended in an appropriate volume of cell culture medium, to remove the trypsin solution, and seeded into T25 flasks or tissue culture plates (*Corning®*) at the required density.

2.2 Bacterial culture

A summary of the bacterial reference strains used throughout this study is shown in Table 2.2.

Table 2.2. Bacterial strains used throughout this study.

Species	Strain reference	Oxygen requirement
Bacteroides fragilis	ATCC 43858	Strict anaerobe
Escherichia coli Nissle	Ardeypharm	Facultative anaerobe
Enterococcus faecalis	ATCC 19433	Facultative anaerobe
Fusobacterium nucleatum	ATCC 25586	Strict anaerobe

2.2.1 Bacterial culture media and culture conditions

For routine non-selective culture, all bacterial strains were cultured using Brain Heart Infusion (BHI) agar (CM3116) or in BHI broth (CM3115; *Oxoid Ltd., Heysham)*. Agar plates were supplemented with 5% defibrinated horse blood (*TCS Biosciences Ltd., Buckingham, UK*). Dehydrated media was added to deionised water and autoclaved at 121°C for 15 minutes. For long term storage, bacterial strains were stored at -80°C using Micro-bank cryopreservation beads (*Pro-Lab Diagnostics, UK*). For revival, the stock was allowed to return to room temperature, and a sterile inoculating loop was used to streak the bacteria onto a fresh agar plate. Both strict and facultative anaerobic bacteria were used throughout this study (see Table 2.2). With the exception of transformed isolates used for plasmid extraction, all bacteria were cultured under anaerobic conditions in an A95 anaerobic workstation (*Don Whitley Scientific Ltd., Bingley, UK*). The gaseous composition in the workstation was calibrated to be 80% Nitrogen, 10% CO_2 and 10% Hydrogen. Cultures were incubated at 35°C and 75%

humidity. Prior to use, bacteriological media was reduced by incubating under anaerobic conditions for 1 hour.

2.2.2 Agar culture

Single colonies were obtained by using a sterile inoculating loop to streak bacteria across a pre-reduced agar plate. Aerobic cultures were incubated at 37°C for 24 hours, whereas anaerobic cultures were incubated for 48 hours.

2.2.3 Planktonic culture

To culture bacteria in the liquid phase, a single colony from an agar plate was isolated using a sterile loop and used to inoculate 5 ml of BHI broth, which was pre-reduced before use with anaerobic bacteria. Aerobic cultures were incubated at 37°C while shaking at 225 rpm, whereas anaerobic cultures were not shaken due to the unavailability of a shaking anaerobic incubator. All broth cultures were incubated for approximately 24 hours. Where required, broth cultures were adjusted to a desired concentration of colony forming units (cfu) per millilitre by reading absorbance at 600 nm using a Jenway Model 6715 UV/Visible Scanning Spectrophotometer (*Cole-Parmer, Saint Neots, UK*). A 0.1 OD_{600nm} culture was found to contain approximately 1×10^8 cfu/ml for each test species.

2.2.4 Bacterial enumeration

To accurately determine the number of cfu in a bacterial suspension, the Miles and Misra technique was used (Miles, Misra and Irwin, 1938). A 1 in 10 serial dilution of bacterial broth culture was prepared to a dilution factor of 10^{-7} in PBS. Dilutions were gently vortexed to ensure homogenisation, and a 10 µl aliquot of each suspension was pipetted onto an agar plate and incubated for the appropriate time period (see Section 2.2.2) to allow distinct colonies to form. Following incubation, the appropriate dilution

for counting was selected, where between 3 and 30 isolated colonies were visible, and this count was used to determine the total cfu/ml in the original sample. A calibration curve of cfu against optical density was created and the cfu present in subsequent cultures was determined by interpolating using this curve.

2.3 Quantification of bacterial adhesion and invasion into tumour cells

To infect cells with viable bacteria, tumour cells were seeded into the appropriate cell culture plate and grown to confluent monolayers to best mimic *in vivo* conditions. Overnight bacterial cultures were centrifuged at 10,000 xG and resuspended in serum-free DMEM. Tumour cell monolayers were washed three times with PBS, and bacteria were added at the desired multiplicity of infection (MOI). After initial infection, cells were treated with 300 μg/ml gentamicin (*Sigma-Aldrich Ltd.*) and 200 μg/ml metronidazole (*Sigma-Aldrich Ltd.*) in serum-free DMEM to eliminate extracellular bacteria. Tumour cells were then washed a further three times with PBS and supplemented with fresh media before use in downstream experiments. The general procedure for tumour cell infection is illustrated in Figure 2.1.

2.3.1 Gentamicin protection assay

To quantify the number of bacteria attaching to or invading tumour cells, the gentamicin protection assay was used (Edwards and Massey, 2011). Tumour cells were grown to confluent monolayers in 96-well plates. Overnight cultures of bacteria were resuspended in serum-free DMEM and added to the monolayers at an MOI of approximately 10,000:1. Co-cultures were then incubated for 4 hours at 37°C and 5% CO_2. Viable counts were performed using bacteria suspended in serum-free DMEM under these conditions to ensure no negative effect on bacterial viability (Supplementary Figure 1). To quantify the total number of bacteria associated with the monolayer, cells were washed three times with sterile PBS, and incubated with 1% w/v saponin (*Sigma-Aldrich Ltd.*) in serum-free DMEM for 30 minutes to lyse tumour cells and release the bound bacteria. Bacterial invasion was quantified by incubating the monolayers with serum-free DMEM containing 300 μg/ml gentamicin (*Sigma-Aldrich Ltd.*) and 200 μg/ml metronidazole (*Sigma-Aldrich Ltd.*) for 1 hour to kill

extracellular bacteria before lysis with saponin. Following lysis, serial dilutions were prepared in PBS and 10 µl aliquots were used to inoculate agar plates in triplicate (see section 2.2.4). Plates were incubated under anaerobic conditions and counted after 24 or 48 hours as appropriate. The number of adherent bacteria was calculated by subtracting the number of invasive bacteria from the total associated bacteria. The concentrations of antibiotics used were confirmed to be effective against the four species tested (Supplementary Figure 2). In addition, solutions of 1% saponin in serum-free DMEM were tested to ensure there was no negative effect on bacterial viability (Supplementary Figure 3).

2.3.2 Determining intracellular bacterial persistence

To determine whether intracellular bacteria were able to survive, and replicate, within colorectal tumour cells, monolayers were infected with bacteria as described in section 2.3.1. Following antibiotic treatment to eliminate extracellular bacteria, fresh medium was added to each well, and cells were incubated under standard conditions. At 24-hour intervals, monolayers were washed three times with PBS, before being lysed with saponin. The resultant lysates were serially diluted, plated onto agar plates, incubated, and counted to enable bacterial enumeration (see section 2.2.4). Addition of penicillin/streptomycin to the medium was used to prevent the outgrowth of any facultative anaerobes which may have been released from the monolayer during experimentation.
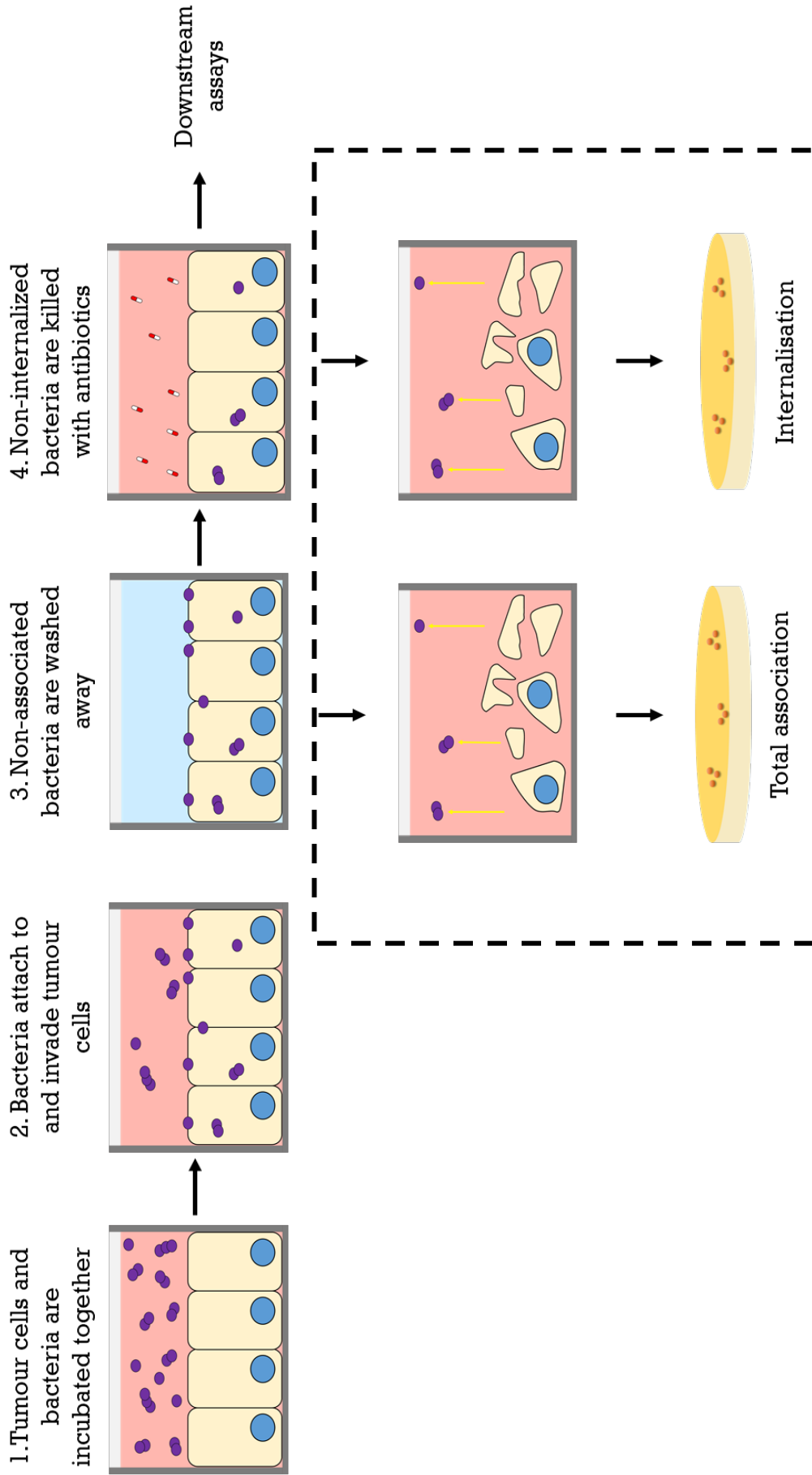
Figure 2.1: Method of tumour cell infection. 1. Tumour cells were cultured as confluent monolayers (or sub-confluent where required) and bacteria were added at a desired MOI. 2. Co-cultures were incubated for 4 hours to allow bacterial interaction with tumour cells. 3. Non-associated bacteria were washed away with PBS. 4. Cells were then treated with antibiotics to eliminate bacteria adhered to the cell surface. Cells containing intracellular bacteria were then used for downstream assays. For the gentamicin protection assay, cells were lysed after washing or after antibiotic treatment, and bacteria were enumerated on agar plates (dashed box).

2.4 Imaging of colorectal tumour cells and their interactions with bacteria

2.4.1 Phase contrast microscopy

To visualise cell morphology, tumour cells were cultured in 12-well plates at the desired seeding density and imaged using a Nikon TE300 inverted microscope and analysed using Image-Pro Plus software (*Meyer Instruments, Texas, USA*).

2.4.2 Scanning electron microscopy

To visualise the interaction of bacteria with the tumour cell surface, 10 mm coverslips were sterilised with 70% ethanol and placed at the bottom of a 24-well plate. HCT116 cells were seeded into the wells at a density of $5x10^5$ cells/cm^2. Cells were then infected with bacteria for 4 hours, as described in Section 2.3, at an MOI of 1,000:1. After infection, coverslip-associated cells were fixed in 4% glutaraldehyde (*Sigma-Aldrich Ltd.*). Following fixation, tumour cells were then washed three times with PBS and subsequently dehydrated using increasing concentrations of ethanol, followed by hexamethyldisilazane (*Acros Organics, Geel, Belgium*). Coverslip-associated cells were then sputter coated using an Emscope SC500 gold sputter coating unit and imaged using an FEI Quanta 650 field emission scanning electron microscope (*Thermo-Fisher Scientific Ltd.)*.

2.4.3 Confocal laser scanning microscopy

Intracellular bacteria were imaged by staining bacterial cells with carboxyfluorescein (CFSE; *Sigma-Aldrich Ltd.*). Bacteria were pelleted by centrifugation at 10,000 xG, washed three times with PBS, resuspended in 5 μM CFSE, and incubated in the dark at 37°C and 200 rpm for one hour. After staining, bacteria were washed three times with PBS and added to sub-confluent RG/C2 or HCT116 cells grown on glass

coverslips at a MOI of 1,000:1 and incubated for 4 hours under standard conditions. Co-cultures were then washed three times with PBS to remove unbound bacteria before fixation.

Samples were simultaneously fixed and permeabilised using a solution of 4% paraformaldehyde (*Sigma-Aldrich Ltd.*) and 0.2% Triton X-100 (*Sigma-Aldrich Ltd.*) in PBS for 10 minutes at room temperature. Cells were blocked with 1% bovine serum albumin (*Sigma-Aldrich Ltd.*) and stained with Alexa-Fluor 555 phalloidin (*Fisher Scientific UK Ltd.*) for 1 hour. Subsequently, cells were counterstained with 4',6-Diamidino-2-phenylindole (DAPI; *Sigma-Aldrich Ltd.*) for 10 minutes. Coverslips were carefully removed from the wells and mounted onto glass slides using Vectashield® Antifade Mounting Medium for Fluorescence (*Vector Laboratories Ltd., United Kingdom*). Samples were stored in the dark at 4°C and imaged within 48 hours of preparation using a SP8 inverted confocal microscope (*Leica Biosystems, Milton Keynes, UK*).

2.5 Crystal violet biofilm formation assays

To determine how host products may impact biofilm formation in gut bacteria, the crystal violet biofilm assay was used (Merritt, Kadouri and O'Toole, 2005). Solutions of porcine mucin (*Sigma-Aldrich Ltd.*) were prepared in BHI broth at various concentrations before autoclaving. Overnight bacterial cultures were diluted 1:100 in either fresh BHI broth, or fresh BHI broth containing mucin, and 100 µl aliquots were pipetted into a 96-well plate and incubated under anaerobic conditions for 48 hours. Following incubation, biofilms were washed twice with sterile water and stained with 125 µl of 0.1% w/v crystal violet (*Sigma-Aldrich Ltd.*). Plates were then washed a further three times and allowed to air dry. To solubilise the crystal violet stain, 125 µl of 33% w/v acetic acid (*Sigma-Aldrich Ltd.*) was added to each well and incubated at room temperature for 15 minutes. One hundred microliters of this solution was then transferred to a new 96-well plate, and absorbance was read at 500 nm using an Infinite 200 microplate reader (*Tecan UK Ltd., Reading, UK*). The absorbance of un-inoculated wells was subtracted as a background reading, and the effect of mucin on biofilm formation was assessed by comparing absorbance to the control.

2.6 Monitoring cell yield and viability in infected tumour cells

2.6.1 MTS assay

To determine how contact with bacteria affected viability in RG/C2 and HCT116 tumour cells, the MTS assay was used. After addition of the MTS assay compound (*Abcam PLC, Cambridge, UK*), viable cells reduce the tetrazolium present in the assay reagent to formazan, inducing a detectable colorimetric change. Cells were grown to confluency in 96-well plates and infected with bacteria as described in Section 2.3. After treatment with 300 µg/ml gentamicin and 200 µg/ml metronidazole to eradicate extracellular bacteria, 180 µl of fresh media was added to each well. Twenty microlitres of MTS assay reagent (*Abcam PLC*) was added and the plate was incubated for one hour under standard conditions. Following incubation, absorbance at 500 nm was measured using a FLUOstar Omega microplate reader (*BMG Labtech Ltd., Bucks, UK*). Assay wells were blanked by subtracting the absorbance of wells which contained media and MTS reagent only.

2.6.2 The effect of bacterial infection on tumour cell yield

RG/C2 and HCT116 cells were seeded into 12-well plates at a seeding density of $4x10^4$ cells/cm^2 or $2x10^4$ cells/cm^2, respectively. Cells were incubated under standard conditions for 72 hours, before being treated with bacteria at a MOI of 10:1 or 100:1 as described in Section 2.4. Following antibiotic treatment (to eliminate extracellular bacteria), tumour cells were washed three times with PBS and fresh media was added. Plates were then incubated for a further 96 hours under standard conditions. To determine the number of apoptotic cells, floating cells present in the media were counted as described in Section 2.1.3. Cells which were adhered to the bottom of the well were considered viable and counted using a haemocytometer as described in Section 2.1.3.

2.7 Cell cycle analysis

To investigate whether observed changes in cell yield were as a result of alterations to the cell cycle, cellular DNA content was analysed via flow cytometry on samples prepared in parallel with cell yield assays.

2.7.1 Flow cytometry sample preparation and staining

To prepare samples for analysis via flow cytometry, cells were cultured and treated with bacteria as described in Section 2.6.2. After trypsinisation, cells were pelleted by centrifugation at 1,200 rpm for 5 minutes. Cells were then resuspended in 1 ml of 70% ethanol at 4°C and gently vortexed, before being incubated at 4°C for 30 minutes. Cells were then pelleted at 2,000 rpm for 5 minutes and stored in ethanol until staining.

To stain DNA for analysis, the ethanol supernatant was discarded, and cells were washed three times by resuspending in PBS and centrifuging at 2,000 rpm for 5 minutes. A solution of 100 μg/ml RNAse A (*Sigma-Aldrich Ltd.*) and 50 μg/ml propridium iodide (*Sigma-Aldrich Ltd.*) was prepared in PBS, and 400 μl of this solution was used to resuspend the cell pellets. Stained cells were stored at 4°C protected from light for at least 12 hours before running to allow the stain to develop. Cell cycle analysis of stained samples was performed using a BD LSRFortessa™ X-20 Flow Cytometer (*BD Biosciences, California, USA*). Forward scatter (FCS-A) and side scatter (SSC-A) were used to gate cell populations, excluding cellular debris and cell doublets. The proportion of the cell population in each cell cycle phase was then quantified by comparing the fluorescence of the propridium iodide stain and modelled using Modfit LT™ software.

2.8 Metabolic analysis

The Seahorse XFe24 Extracellular Flux Analyser (*Agilent Technologies Ltd., Stockport, UK*), hereafter referred to as 'the Seahorse', was used to measure the metabolism of colorectal tumour cells cultured *in vitro*. The instrument measures oxygen consumption and extracellular acidification within the media as indicators of cellular respiration and glycolysis, respectively.

2.8.1 Seahorse general assay procedure

Seahorse XF cell culture plates (*Agilent Technologies Ltd.*) were exclusively used for metabolic assays. Colorectal tumour cells were harvested as described in Section 2.1.3 and seeded into assay plate wells in a volume of 50 μl. According to manufacturer guidelines, four wells are used to normalise for background readings, therefore only media was added to these wells. Culture plates were incubated at room temperature (within a biological safety cabinet) for 1 hour to promote even cell distribution according to manufacturer guidelines, before being transferred to a 37 °C and 5% CO_2 humidified incubator. An additional 100 μl of fresh media was added after 8 hours of incubation. One day before the assay, 1 ml of Seahorse XF calibrant was added to each well of a Seahorse XF sensor cartridge, and this sensor cartridge was placed into a humidified CO_2-free incubator at 37°C.

Assay medium was prepared on the day of the assay. Seahorse XF DMEM medium was supplemented with 10 mM glucose (*Sigma-Aldrich Ltd.*), 1 mM Sodium pyruvate (*Fisher Scientific UK Ltd.*) and 2 mM L-glutamine, before being adjusted to a pH of 7.4. Treatment compounds were loaded into the appropriate injector port of the sensor cartridge. The sensor cartridge was then placed into the machine for validation. On the day of the assay, cells were washed three times with pre-warmed

Seahorse assay medium, and a final volume of 500 µl was added to each well. The cell culture plate was placed into a humidified CO_2-free incubator for 1 hour prior to the commencement of the assay. Following this incubation period, the cell culture plate was loaded into the machine and the assay was initiated.

2.8.2 Optimisation of seeding density and carbonyl cyanide-4 (trifluoromethoxy) phenylhydrazone (FCCP) concentration

Seeding density for each cell line was optimised by seeding cells at various densities into a Seahorse XF cell culture plate. The plate was then analysed using the Seahorse without injection of treatment compounds, and a seeding density was chosen where an increase in cell density resulted in a linear increase in oxidative phosphorylation.

The Seahorse sensor cartridge contains four injector ports which allow the real-time addition of compounds to the cell culture medium. By injecting compounds which alter cellular metabolism, the basal and maximal metabolic rates can be determined. Throughout this work, carbonyl cyanide-4 (trifluoromethoxy) phenylhydrazone (FCCP) was used to modulate oxidative phosphorylation, whereas Oligomycin was used to modulate glycolysis. Both compounds are provided within the Seahorse XF Energy Phenotype test kit.

Oligomycin was used at the manufacturer's recommended concentration of 1 µM after confirming that injection increased extracellular acidification rate (ECAR) be a factor of at least 2-fold in each cell line. The appropriate concentration of FCCP was determined for each cell line once the optimal seeding density had been determined. For all cell lines, the largest increase in oxidative phosphorylation was seen at an FCCP concentration of   1 µM.

2.8.3 Crystal violet normalisation

Data for the rates of oxidative phosphorylation and glycolysis were normalised against the total number of cells within each well. After assay completion, cells were fixed in 4% paraformaldehyde and stained with 100 µl of 0.1% crystal violet solution (*Sigma-Aldrich Ltd.*). Cells were then washed three times in PBS, before being lysed in 1% sodium dodecyl sulphate (SDS; *Sigma-Aldrich Ltd.*) at 100 rpm for 1 hour. The lysate was transferred to a 96-well plate and absorbance was measured at 595 nm. The absorbance of each well was compared to a standard curve and used to interpolate the number of cells in each well.

2.8.4 Measuring the impact of bacterial invasion on cellular metabolism

Twenty-four hours before measurement in the Seahorse, and a minimum of 24 hours after seeding, tumour cells were infected with bacteria as described in Section 2.3. Following bacterial infection, cells were washed three times with PBS and treated with 300 µg/ml gentamicin and 200 µg/ml metronidazole for 1 hour to eliminate extracellular bacteria. Cells were washed a further three times with PBS before fresh media was added. Cells were then returned to the incubator and cultured under standard conditions. The Seahorse assay was then performed as described in Section 2.8.1.

2.9 Analysis of volatile metabolites released by tumour cells

Volatile compounds present in the headspace of colorectal tumour cells cultured in T25 cell culture flasks were detected using selected ion flow tube mass spectrometry (SIFT-MS). Colorectal tumour cells were grown to confluency in T25 cell culture flasks and infected with bacteria at a MOI of 10:1 as described in Section 2.3, or treated with serum-free media as a control. Following infection and antibiotic treatment, fresh serum-free medium was added. Flask caps were then sealed with parafilm and flasks were incubated at 37°C for 30 minutes to allow accumulation of volatile compounds in the headspace. A 21-gauge needle was attached to the sampling arm of a Voice200 Ultra instrument (*Syft Technologies, New Zealand*) and passed through the parafilm-covered cap of the cell culture flask for sampling. The instrument was operated in full mass scan mode over a spectrum range of 15-200 m/z using H_3O^+, NO^+ and O_2^+ reagent ions to detect volatile product ion peaks. Three repeat scans were performed for each sample. To analyse data, a background subtraction was performed whereby T-tests were used to identify volatile product ion peaks which were significantly different from the control (blank media or uninfected cells). Hierarchical cluster analysis (HCA) and principal component analysis (PCA) were used to analyse significant peaks using IMB SPSS statistics for Windows version 26.

2.10 In vitro migration and invasion assays

2.10.1 Wound scratch assay

RG/C2 and HCT116 cells were seeded into 6-well plates at a seeding density of 1×10^5 cells/cm^2 and 5×10^4 cells/cm^2, respectively, and cultured until confluent monolayers were formed. Tumour cells were then infected with bacteria at an MOI of 10:1 as described in Section 2.3. Following infection and antibiotic treatment to eliminate extracellular bacteria, a sterile pipette tip (P200) was used to create a wound of approximately 1 mm across the diameter of the well. Cells were washed three times with PBS to remove cellular debris and fresh DMEM supplemented with 1% FBS was added. For each wound, phase contrast images of three separate locations were taken at 0 hrs and 24 hrs at a magnification of 400x. For each image, the surface area of the wound was measured using the MRI Wound Healing Tool for ImageJ software (*NIH, Maryland, USA*). The degree of wound closure was expressed as the percentage reduction in surface area of the wound observed within a 24-hour period.

2.10.2 Transwell filter migration assay

Millicell hanging cell culture inserts (*Millipore UK Ltd., Watford, UK*), with a pore size of 8 μm, were coated on the basal side with 10 μg/ml VitroCol® human collagen type 1 (*Advanced Biomatrix Inc., California, USA*) and left to dry in a biological safety cabinet overnight. Once dry, inserts were placed into 6-well plates, which were sealed with parafilm and stored at 4°C until needed. Colorectal tumour cells were cultured and infected with bacteria as described in Section 2.3. After elimination of extracellular bacteria, cells were detached with 0.48 mM versene solution (*Sigma-Aldrich Ltd.*) in PBS, before being counted and resuspended in Ca^{2+}-free DMEM (*Fisher Scientific UK Ltd.*). RG/C2 and HCT116 cells were then seeded onto the apical surface of the collagen coated inserts at seeding densities of 4×10^4 cells/cm^2 and

$2x10^4$ cells/cm^2, respectively. Ca^{2+}-free medium supplemented with 5% FBS was added to the lower compartment to act as a chemoattractant, and the fluid level was equilibrated between the upper and lower chambers. Cells were then incubated at 37°C for 24 hours. The design of the assay is illustrated in Figure 2.2.

Following incubation, inserts were removed from the 6-well plates and washed three times in PBS. Cells were then fixed in 100% ice-cold methanol (*Fisher Scientific UK Ltd.*) for 10 minutes, washed three times in PBS and stained with 1% w/v crystal violet solution. Inserts were washed a further three times with PBS to remove excess crystal violet, and the remaining cells on the apical surface of the insert were removed with a cotton swab. Cells embedded in the collagen matrix on the basal surface of the insert were considered migratory. The insert membrane was dissected with a scalpel and mounted onto a glass slide before being imaged under a microscope. Migrated cells were counted in ten separate fields of view at 200x magnification. Three independent experiments were performed, in each of which three separate inserts were counted in 10 separate images for each condition.

2.10.3 Transwell filter invasion assay

In vitro invasion was assessed using Corning® Biocoat™ Matrigel invasion chambers (*Sigma-Aldrich Ltd.*). In contrast to the transwell filter migration assays, the basal surface of the filter is coated in Matrigel®, and therefore cells must produce the necessary enzymes to degrade extracellular matrix in order to pass through the filter. Cells were grown to confluency in 6-well plates and treated with bacteria as described above. Matrigel invasion chambers were reconstituted by placing each chamber into a 6-well plate well and adding Ca^{2+}-free medium to the upper and lower compartments. Chambers were incubated at 37°C for 2 hours, after which the media was discarded, and cells were seeded onto the apical surface at a seeding density of $4x10^4$ cells/cm^2.

Media containing 5% FBS was added to the lower compartment, and cells were incubated for 24 hours at 37°C and 5% CO_2. After incubation, cells embedded in the Matrigel matrix were fixed, stained, and counted as described in Section 2.10.2.
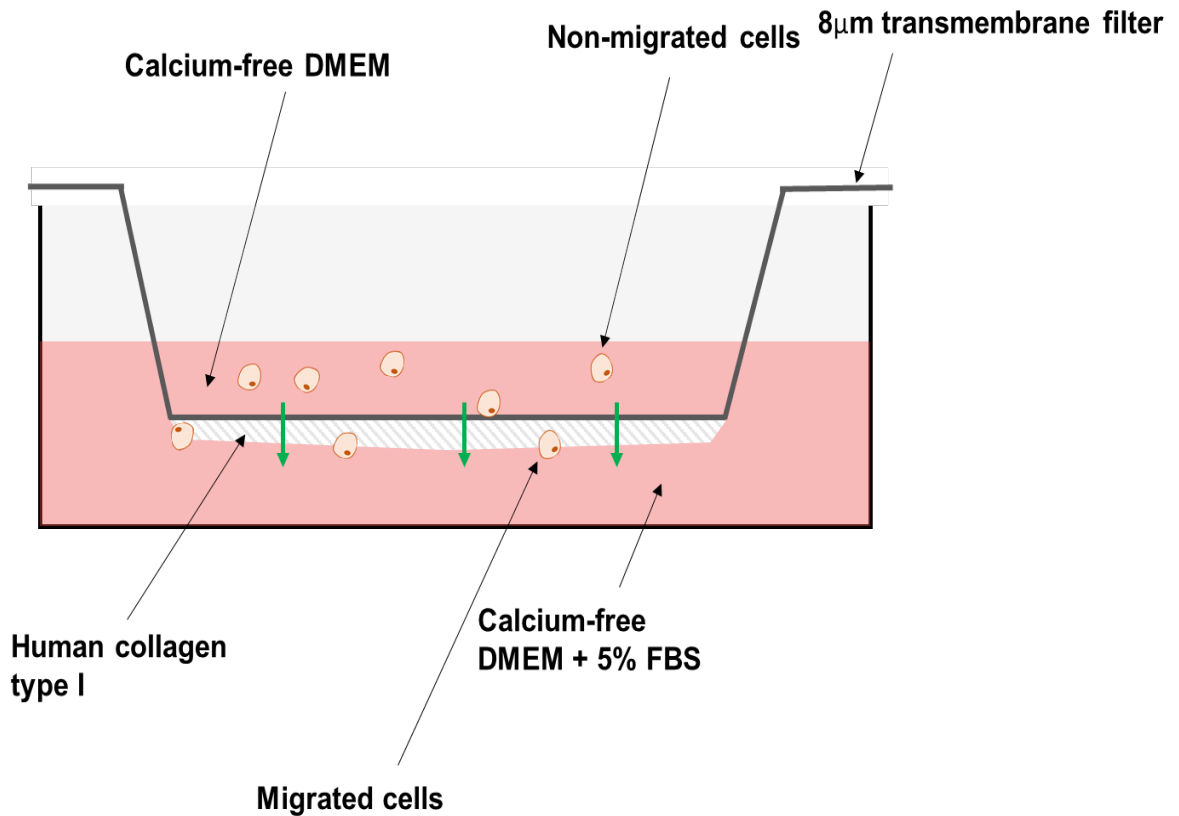
Figure 2.2: Transwell filter migration assay design. Transwell filter migration assays were performed using collagen-coated cell culture inserts with an 8 μm pore size. Cells migrated from the apical chamber through the pores and became embedded in the collagen matrix. These migratory cells were then stained and counted. Invasion assays were set up similarly, however inserts pre-coated with Matrigel were used.

2.11 Luciferase reporter assays

The TOPFLASH luciferase reporter assay was used to quantify Wnt signalling activity in colorectal tumour cells (Molenaar *et al.*, 1996). The TOPFLASH luciferase reporter plasmid contains numerous Wnt-sensitive Tcf binding sites coupled to a *fos* promoter, upstream of a firefly luciferase reporter. The TOPFLASH reporter therefore yields luciferase activity from its *fos* promoter, and its Tcf binding sites. The control reporter, FOPFLASH, has mutated Tcf binding sights and thus only yields luciferase activity from its *fos* promoter. Therefore, the ratio of TOPFLASH luminescence to FOPFLASH luminescence provides a direct measurement of Wnt signalling-specific transcriptional activity. A third plasmid, containing a renilla luciferase reporter bound to a cytomegalovirus promoter was used to control for transfection efficiency. All plasmids were kindly donated by Dr. Alexander Greenhough (*UWE*) and are summarised in Table 2.3.

Table 2.3: Plasmids used for the TOPFLASH assay.

Construct	Function	Selective agent
CMV Renilla	Control for transfection efficiency	Ampicillin
TOPFLASH	LEF/TCF responsive reporter	Ampicillin
FOPFLASH	TOPFLASH control reporter	Ampicillin

2.11.1 Bacterial transformation

Chemically competent *E. coli* DH5α recipient cultures were kindly provided by Dr. Tim Craig (*UWE*), and were subsequently transformed with donor plasmid. Prior to transformation, 100 µl aliquots of competent bacterial cells were thawed on ice. Five

microliters of donor plasmid were added to the bacteria, and incubated on ice for 30 minutes, before heat shocking bacteria for 90 s at 42°C. Bacterial cultures were diluted 1:1 in Super Optimal broth with Catabolite repression (SOC) outgrowth medium (*New England Biolabs Ltd., Hitchin, UK*) and incubated at 37°C and 200 rpm in a shaking incubator for 1 hour. Following incubation, serial dilutions were plated onto Luria broth (LB; *Oxoid Ltd.*) agar plates containing 100 μg/ml ampicillin (*Fisher Scientific UK Ltd.*). Non-transformed cultures were also inoculated onto ampicillin plates to ensure sensitivity to ampicillin was present in the initial culture.

2.11.2 Plasmid purification from transformed *E. coli*

Once bacteria had been transformed, the Qiagen Plasmid Maxi Prep Kit (*Qiagen, Germany*) was used to according to the manufacturer's protocol to extract plasmid DNA. Once extracted, plasmid DNA was dissolved in pH 8.0 Tris-EDTA buffer (*Sigma-Aldrich Ltd.*) and stored at 4°C. The yield and purity of extracted plasmid DNA was determined using the Qubit dsDNA high-sensitivity assay (*Thermo-Fisher Scientific*).

2.11.3 Tumour cell transfection

Transfection was performed by diluting 3.6 μg of either TOPFLASH or FOPFLASH plasmid, combined with 0.36 μg of CMV-Renilla plasmid in 500 μl of OptiMem™ medium (*Fisher Scientific UK Ltd.*). In a separate tube, 18 μl of Lipofectamine 2000 transfection reagent (*Invitrogen, California, US*) was diluted in 500 μl of OptiMem, and this was then added to each plasmid mix and incubated at room temperature and 100 rpm in a shaking incubator for 20 minutes. Sub-confluent tumour cells were trypsinised and 1.8×10^6 cells were resuspended in OptiMem. The transfection mix containing either TOPFLASH or FOPFLASH plasmid, in addition to CMV-Renilla plasmid, and Lipofectamine 2000 was added to the cell suspension, and cells were

immediately seeded into 96-well plates at a seeding density of 1×10^5 cells/cm^2. For each cell line, cells transfected with TOPFLASH or FOPFLASH were seeded into a 96-well plate and incubated for 24 hours at 37°C and 5% CO_2.

2.11.4 Dual-Luciferase® Reporter assay

Once transfected, cells were grown to confluency before infection with bacteria as described in Section 2.3. The Dual-Luciferase Reporter Assay System (*Promega, Southampton, UK*) was then used to quantify luciferase activity from treated cells. Cells were washed three times with PBS and 20 μl of passive lysis buffer was added to each well. Plates were placed on a shaker at 200 rpm for 1 hour and complete cell lysis was visually confirmed using a microscope. Fifty microliters of LAR II reagent were added to each well of a black, clear-bottomed 96-well plate (*Corning®*). Ten microliters of lysate were added to the LAR II reagent, and Firefly luminescence was immediately measured using a FLUOstar Optima microplate reader (*BMG Labtech, Bucks, UK*) plate reader. Fifty microliters of Stop&Glo® reagent were then added to each well and Renilla luminescence activity was recorded. Wells containing untransfected cells were used to control for background luminescence and the average of these wells was subtracted from all readings. Catenin-related transcription was expressed as the ratio of luminescence of TOP:FOP, after normalising for cell number using the CMV reporter.

2.12 Western blots

2.12.1 Preparation of protein lysates

Protein samples were prepared from RG/C2 and HCT116 cells grown in 6-well plates and infected as described in Section 2.3. Following infection, cell monolayers were incubated for 24 hours under standard conditions. Plates were then placed on ice and cells were washed three times with ice-cold PBS, before the addition of 1 ml of ice-cold lysis buffer to each well. The formulation of all buffers and gels used for Western blotting can be found in Table 2.4. Cell monolayers were further disrupted using a cell scraper and the lysates were transferred into pre-cooled Eppendorf tubes and incubated at 4°C for 30 minutes, whilst being briefly vortex mixed approximately every 5 minutes. Lysates were then centrifuged at 12,000 xG for 20 minutes at 4°C. Supernatants were collected and transferred to fresh, pre-cooled Eppendorf ml tubes and stored at -80°C until required.

2.12.2 Estimation of protein concentrations

To determine protein concentration, the Bradford assay was used (Bradford, 1976). Bovine serum albumin (BSA) protein standards were prepared in sterile ddiH$_2$O at concentrations ranging from 0.1-1.5 μg/μl. Five microliters of protein standard or protein sample was added to the base of each well in a 96-well plate. Two hundred microliters of Bradford reagent (*Sigma-Aldrich Ltd.*) were then added to each well. The binding of protein to Coomassie dye present in Bradford reagent causes a colour change from brown to blue, and absorbance of samples after the addition of dye was measured at 595 nm using an Infinite 200 microplate reader. The concentration of protein within samples was estimated by interpolating with a BSA standard curve prepared by measuring the absorbance of known quantities of BSA diluted in ddiH$_2$O using the Bradford assay.

2.12.3 SDS-PAGE

Proteins were separated according to their molecular weight using SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Resolving gels were prepared to a concentration of 10% acrylamide and overlaid with a 4% stacking gel (see Table 2.5). Protein samples were prepared by adding 25% loading dye (*Sigma-Aldrich Ltd.)* and 0.05% μl β-mercaptoethanol (*Sigma-Aldrich Ltd.*) to 50 μg of protein, before denaturation in a 95°C water bath for 10 minutes. Samples were then loaded into the gel, leaving an empty lane between each sample. In addition, 5 μl of protein ladder was added to the leftmost lane (*Amersham Biosciences plc., Buckinghamshire, UK*). The gel tank (*Bio-Rad Laboratories, California, USA*) was then filled with running buffer (see Table 2.4). Samples were initially run at 100 V for 20 minutes, or until the dye front had passed through the stacking gel. Voltage was then increased to 150 V and samples were run for a further 90 minutes.

2.12.4 Wet transfer and antibody probing

Following electrophoresis, gels were removed from the tank and the stacking gel was removed. Proteins were transferred onto Amersham™ Hybond® 0.45 μm polyvinylidene difluoride (PVDF) blotting membrane which had been activated by briefly soaking in 100% methanol. PVDF membranes and the remaining resolving gel were sandwiched between layers of filter paper, followed by sponge, which had been soaked in transfer buffer (see Table 2.4). A roller was used to eliminate air bubbles which may have interfered with efficient transfer. Transfer was performed in a Bio-Rad transfer tank filled with transfer buffer (see Table 2.4) at 4°C for two hours at 50 V. Following transfer, membranes were washed three times for 5 minutes in TBST, before being blocked for 1 hour in blocking buffer (see Table 2.4). Primary antibodies were diluted in blocking buffer and added to membranes in sterile petri dishes, which

Table 2.4: A list of buffers used for Western blotting. All buffers were made up in diH$_2$O.

Buffer	Composition
Cell lysis buffer	150 mM NaCl
	1% Triton X-100
	50 mM Tris
	1x Protease inhibitor cocktail tablet per 10 ml
Running buffer	25 mM Tris
	190 mM Glycine
	0.1% SDS
Transfer buffer	25 mM Tris
	190 mM Glycine
	0.1% SDS
	25% Methanol
Tris-buffered saline (TBS) – pH 7.5	200 mM NaCl
	19 mM Tris
Tris-buffered saline + Tween (TBST)	200 mM NaCl
	19 mM Tris
	0.1% Tween-20
Blocking buffer	5% w/v dried skimmed milk in TBST
Resolving buffer – pH 8.8	1.5M Tris
Stacking buffer – pH 6.8	0.5M Tris

were then placed on a rocker and incubated overnight (16 hours) at 4°C. A summary of the primary antibodies used (and their dilutions) is presented in Table 2.6. After incubation with primary antibody membranes were washed a further three times with TBST and secondary antibody (see Table 2.7) prepared in blocking buffer, was added to the membrane. After incubation with secondary antibodies for one hour, membranes

were imaged using a Li-Cor Odyssey gel imager (*Li-Cor, Nebraska, USA*). Protein

bands were identified after exposure to fluorescent light at 600 or 700 nm for 2 minutes

and densitometry was used to determine relative protein concentrations.

Table 2.5: Gel compositions used for Western blots.

10% Resolving gel		4% Stacking gel	
diH$_2$O	6.4 ml	diH$_2$O	3.05 ml
Resolving gel buffer	2.5 ml	Stacking gel buffer	1.25 ml
30% Acrylamide	3.3 ml	30% Acrylamide	650 µl
10% SDS	100 µl	10% SDS	50 µl

Table 2.6: List of primary antibodies used to probe PVDF membranes from proteins of interest.

Antibody	Clone	Isotype	Supplier	Concentration	Band size
E-cadherin	34/E-cadherin	Mouse monoclonal IgG	Becton Dickinson Transduction Laboratories	1:500	120 kda
β-catenin	14/β-catenin	Mouse monoclonal IgG	Becton Dickinson Transduction Laboratories	1:500	92 kda
MMP-2	42-5D11	Mouse monoclonal IgG	Sigma-Aldrich Ltd.	1:500	72 kda
SVIL	B8C1	Mouse monoclonal IgG	Santa Cruz Biotechnology	1:500	240 kda
β-tubulin	5H1	Mouse monoclonal IgG	Becton Dickinson Transduction Laboratories	1:2000	50 kda

Table 2.7: Secondary antibodies used for protein visualisation and densitometry using the Li-Cor gel imager.

Antibody	Host	Reactivity	Supplier	Concentration
IRDye® 800CW	Donkey	Mouse	Li-Cor	1:15,000
IRDye® 680RD	Donkey	Mouse	Li-Cor	1:15,000

2.13 Quantitative PCR Microarrays
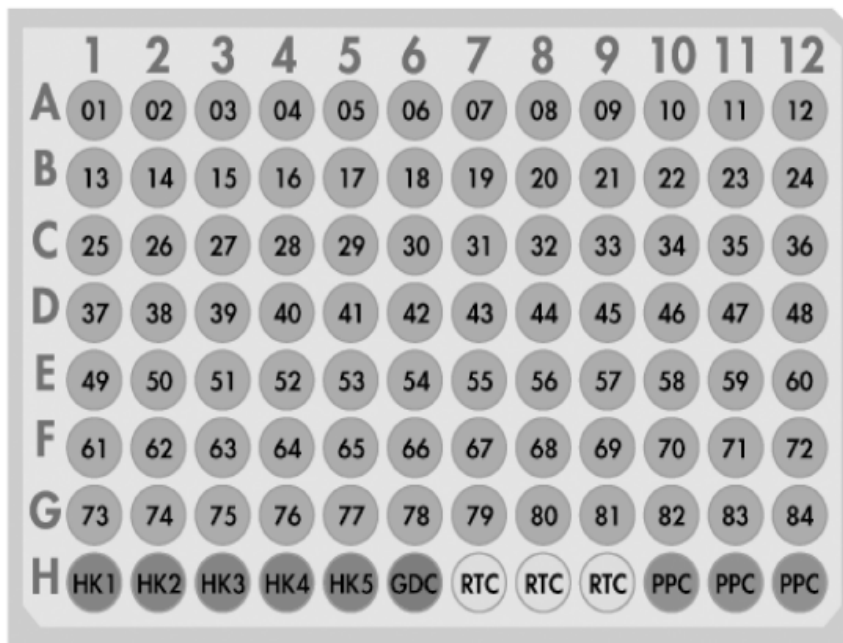
Changes in genes expression as a result of bacterial infection were analysed using quantitative polymerase chain reaction (qPCR). RG/C2 or HCT116 cells were grown to confluency in 6-well plates, before being infected with bacteria as described in Section 2.3. RNA was extracted and reverse transcribed into cDNA, which was then analysed using Qiagen RT² Profiler PCR Arrays (*Qiagen, Germany*).

2.13.1 RNA extraction and cDNA production

Following infection, RNA was isolated from colorectal tumour cells using the RNeasy Mini Kit (*Qiagen*). Before work with RNA commenced, the biological safety cabinet was sterilised with ultraviolet light overnight, and all surfaces were cleaned using RNAZap™ (*Thermo-Fisher Scientific Ltd.*). After isolation and elution of RNA using the RNeasy Mini kit, purity was assessed by running samples on a 1% agarose gel containing 0.5% ethidium bromide. An equal volume of loading dye was added to 5 µl of sample, which was run at 100 V until the dye front had travelled sufficiently though the gel. Gels were carefully transferred removed and imaged under ultraviolet light. The intensity of the 28S and 18S RNA bands were compared, with the intensity of the 28S band being approximately double that of the 18S band representing high quality RNA. The integrity of the RNA was also assessed by comparing absorbance at 260:280 nm using a Nanodrop™ spectrophotometer (*Thermo-Fischer*) which was also used to determine RNA concentration. After quantification, 0.5 µg of sample was immediately reverse transcribed into cDNA using the RT² First Strand Kit (*Qiagen*) according to the manufacturer's instructions. The concentration and purity of cDNA was analysed using a Nanodrop™ spectrophotometer. cDNA samples were stored at -20°C until required.

2.13.2 RT² Profiler PCR Arrays

Quantitative PCR was performed using the ABI StepOnePlus system (*Applied Biosystems, California, USA*). Arrays were used to investigate both cell motility and epithelial-mesenchymal transition, and the full lists of target primers present in each array are described in Supplementary Tables 1 and 2, respectively. SYBR Green Mastermix was added to entire cDNA samples as per the manufacturer's instructions and 25 µl of the PCR components were loaded into each well of an RT^2 Profiler PCR Array microplate (*Qiagen*). Plates were then loaded into the PCR machine and cycled once for 10 minutes at 95°C, followed by 40 cycles of 15 s at 95°C and 1 minute at 60°C. Threshold values were manually determined and kept constant across similar assays. The layout of these primers in each assay plate is illustrated in Figure 2.3. Data was analysed using the Qiagen GeneGlobe Data analysis centre.

Figure 2.3: RT² Profiler PCR array format. Wells A1 to G12 contained primers for individual genes. Wells H1-H5 contained a housekeeping gene panel to normalize array data **(HK1–5)**, full gene lists for both assay can be found in Supplementary Table 1 and 2. Well H6 contained a genomic DNA control **(GDC)**. Wells H7-H9 contained replicate reverse-transcription controls **(RTC)**. Wells H10-H12 contained replicate positive PCR controls **(PPC)**.

2.14 Statistical analysis

All data presented in graphical form represent the mean ± standard deviation of the mean. Three biological replicates were performed in triplicate for each assay. All data analysis and production of graphs was performed using GraphPad Prism 7 for Windows, with the exception of the analysis of SIFT-MS data which was performed using SPSS. Unless otherwise stated, non-paired T-tests were used to determine statistical significance, with a p-value of less than 0.05 considered to be statistically significant. Where data has been normalised prior to statistical analysis this is stated in the figure legend.

Chapter 3: Quantifying the interactions between gut bacteria and colorectal tumour cells

3.1. Introduction

In the healthy colon, homeostasis between the epithelium and the microbiota is maintained through spatial separation by host mucins, and the secretion of antimicrobial peptides (Schroeder, 2019). Bacterial colonisation is normally restricted to the outer mucus layer, as increasing concentration of antimicrobial peptides and decreasing mucin pore size nearer the mucosal surface restricts bacterial growth (Johansson *et al.*, 2008). However, specific species have become adapted to colonise this inner layer, as demonstrated by unique microbiota of the colonic crypts which includes *Bacteroides fragilis* and *Akkermansia muciniphila* (Swidsinski *et al.*, 2005; Donaldson, Lee and Mazmanian, 2015). The spatial organisation of the microbiota is altered in CRC, with mucosa-adherent microbial biofilms frequently identified on both tumour tissue and paired normal tissue (Kinzler *et al.*, 2014). These biofilms allow aberrant colonisation of the inner mucus layer and increase the frequency and duration of contact between bacteria and the epithelium (Johansson *et al.*, 2014). Furthermore, the formation of mucosal biofilms is thought to provide a niche for invasive species such as *F. nucleatum*, which are not found in biofilm-negative tissue, to colonise the colon (Dejea *et al.*, 2018).

A key aspect of colonic biofilm virulence is thought to be the increased prevalence of invasive microorganisms found within them. This includes *B. fragilis* (Nakano *et al.*, 2008), *E. coli* (Martin *et al.*, 2004) and *F. nucleatum* (Strauss *et al.*, 2011). This invasion of the epithelium, which includes direct invasion of the colonocytes and the underlying tissue, is associated with pro-tumorigenic effects.

Entry of *B. fragilis* and *F. nucleatum* into colonic epithelial cells is facilitated by the virulence factors BFT and FadA, respectively, which bind to and cleave E-cadherin (Saidi and Sears, 1996; Rubinstein *et al.*, 2013). Furthermore, a study of *F. nucleatum* strains isolated from IBD patients demonstrated that strains isolated from inflamed tissue possess greater invasive potential than those identified from non-inflamed tissue taken from IBD patients or controls (Strauss *et al.*, 2011).

Bacterial interactions with host cells, and the ability of bacteria to attach to and invade the epithelium is largely dependent upon cell surface architecture, which alters dramatically during oncogenesis (Weidle *et al.*, 2011). During the progression from healthy epithelium to adenocarcinoma, epithelial surface area, secretion of mucins, structure of cellular junctions and expression of surface proteins are all significantly altered (Filipe and Branfoot, 1974; Soler *et al.*, 1999; Weidle *et al.*, 2011; Sewda *et al.*, 2016). Therefore, interactions between bacteria and tumour cells are likely to change throughout the adenoma-carcinoma sequence (Vogelstein and Kinzler, 1993). Studying these interactions *in vivo* is challenging as both immune development and mucus secretion are severely compromised in germ-free models, which influences bacteria-epithelium interactions (Cebra, 1999; Johansson *et al.*, 2015). In addition, the high cost and time investment can be prohibitive to high throughput experiments. This chapter will describe the development of a simple 2D *in vitro* model for comparing the interactions of bacteria with RG/C2 and HCT116 monolayers. This model will then be used to quantify the level of attachment and invasion of tumour cells by CRC-associated bacteria.

3.2. Aims

The aim of this chapter was to investigate the differences in the interaction of bacteria with benign and malignant colorectal tumour cells.

3.3 Objectives

- To develop a laboratory model to study the infection of RG/C2 and HCT116 tumour cells, enabling quantification of any resultant downstream effects using the gentamicin protection assay. These included bacterial attachment, invasion, and intracellular persistence.

- To visualise bacteria-tumour interactions using confocal microscopy and scanning electron microscopy, to gain insight into bacterial binding to tumour cell surfaces and sub-cellular localisation.

- To determine how interaction with mucin alters bacterial growth, to inform the future development of a laboratory model more closely mimicking *in vivo* conditions.

3.4 Results

3.4.1. Benign and malignant colorectal tumour cells display distinct morphologies *in vitro*

When developing a co-culture model for colorectal tumour cells and gut bacteria, it was important to work with confluent monolayers to best imitate *in vivo* conditions. The ability of RG/C2 adenoma cells to form monolayers *in vitro* was compared with the adenocarcinoma cell lines HCT116, HT29 and SW480. Cells were seeded at equal seeding densities and imaged at 24-hour intervals using phase contrast microscopy (Figure 3.1). The three adenocarcinoma cell lines readily attached to the surface of the cell culture well. Clear evidence of cell division was visually observed within 24 hours, and confluent monolayers formed within 72 hours. In contrast, a visual increase in RG/C2 cells was not observed until 72 hours post-seeding. Furthermore, the visual inspection of plates revealed numerous floating cells and cellular debris, indicative of reduced viability. To determine the time necessary for RG/C2 cells to reach confluency, cultures were continued with media changes every 48 hours. After 2 weeks of continuous culture, RG/C2 cells formed large islands which eventually began to merge with adjacent colonies to form a confluent layer. In contrast to the adenocarcinoma cell lines which appeared to form single cell-thick monolayers, RG/C2 cultures appeared to grow in three dimensions.

Future experiments primarily used the RG/C2 and HCT116 cell lines. To further investigate the structural differences present in RG/C2 and HCT116 monolayers, cells were examined using confocal laser scanning microscopy (CLSM). The actin cytoskeleton was stained, and colonies were imaged in series at 0.3 μm intervals which enabled the creation of three-dimensional reconstructions. RG/C2 cells grew in layers 4-5 cells thick (Figure 3.2A), whereas HCT116 cells were

confirmed to grow as single cell-thick monolayers (Figure 3.2B). Confocal microscopy also revealed differences in individual cell morphology between RG/C2 and HCT116 cells. RG/C2 cells were rounded, and DAPI counterstaining showed that the nucleus dominated the intracellular space. In contrast, much larger angular cytoskeletons were observed in HCT116 cells which displayed numerous projections.
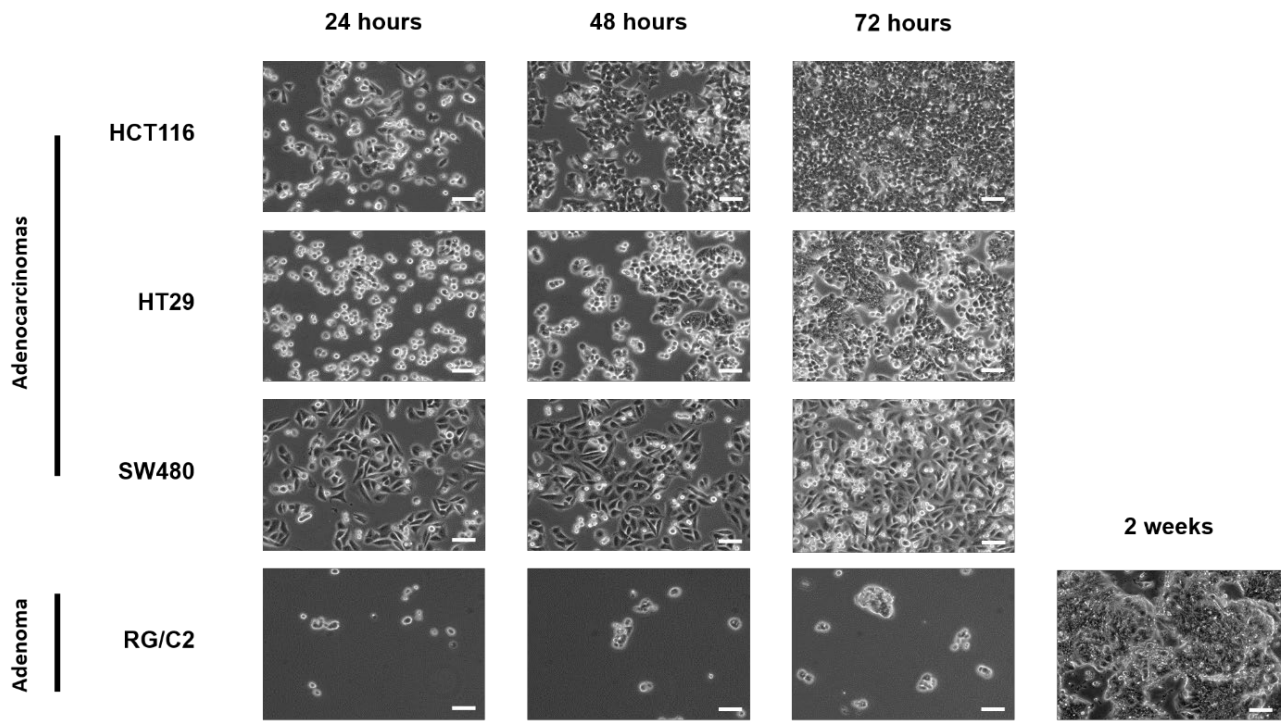
Figure 3.1: A comparison of the monolayer forming capabilities of colorectal adenocarcinoma and adenoma cell lines. Four colorectal tumour cell lines were seeded into 6-well plates at a seeding density of 4×10^4 cells/cm^2. Adenocarcinoma cell lines (HCT116, HT29, SW480) showed high viability after seeding, and rapidly formed confluent monolayers. RG/C2 cells formed smaller, isolated colonies, which grew at a significantly slower rate. Images were taken at 100 x objective using a Leica phase contrast microscope and are representative of three independent experiments. Scale bar represents 20 μm.
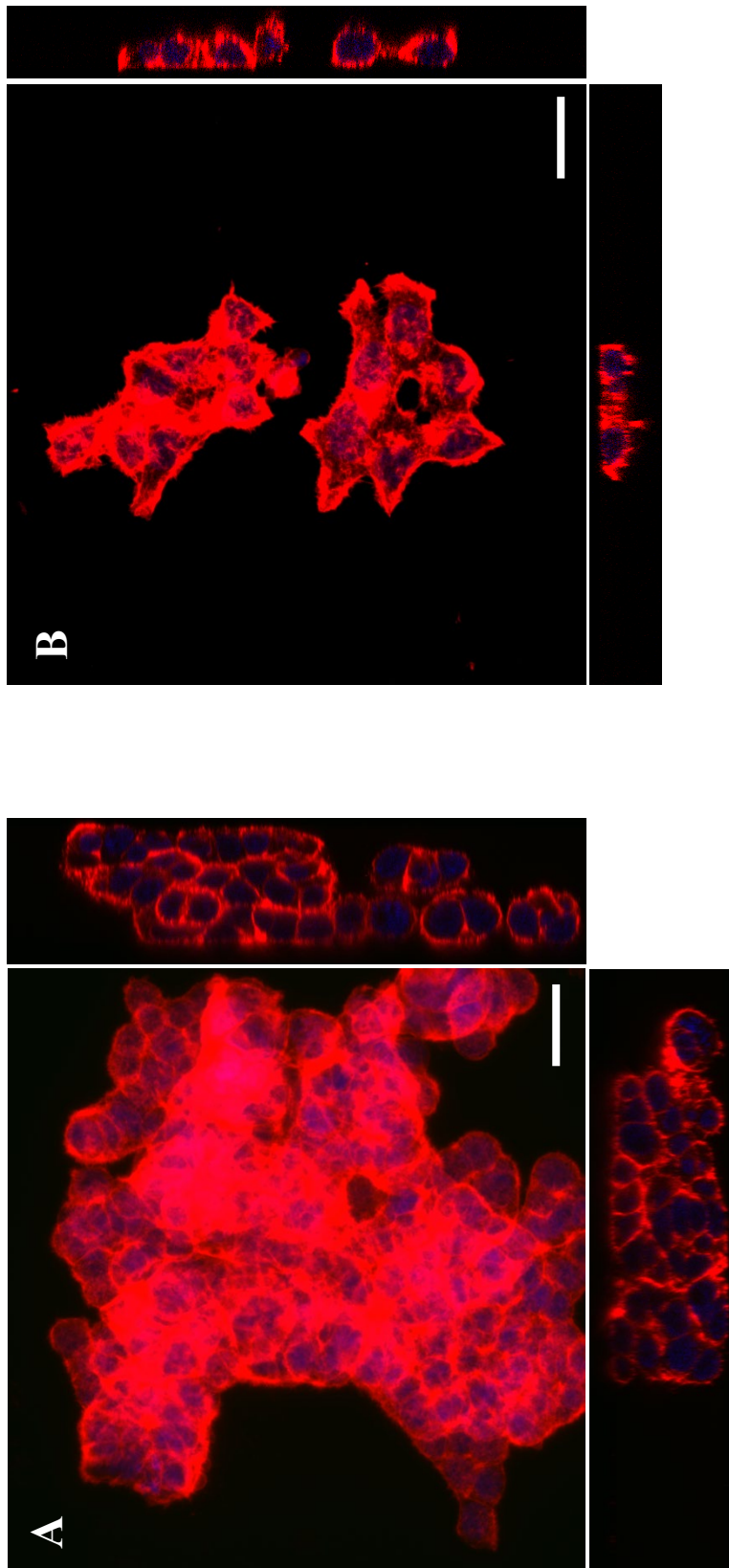
Figure 3.2: Confocal laser scanning micrographs of colorectal tumour cells. RG/C2 (A) and HCT116 (B) cells were cultured on glass coverslips. Cells were stained for actin using AlexaFluor™ 555 Phalloidin (red) and the cell nuclei were counterstained with DAPI (blue). Areas containing subconfluent colonies were imaged to best visualise cell morphology. Centre panels show the XY plane, bottom and side panels show the XZ and YZ slices, respectively. Scale bar represents 10 µm. Representative images from three individual experiments.

3.4.2. Co-culture with bacteria does not impact tumour cell viability

Before attempting to quantify bacterial interactions with tumour cell monolayers, the impact of bacterial infection on RG/C2 and HCT116 tumour cell monolayers was assessed using the MTS assy. The reduction of tetrazolium in the MTS assay reagent produces a soluble formazan dye and causes a detectable colour change. This reaction is performed by NADPH-dependent enzymes in metabolically active cells, and the resulting colour change provides an indication of cell viability.

Bacteria were added to confluent tumour cell monolayers at multiplicities of infection (MOIs) of 10:1, 100:1 and 1,000:1. Following a 4-hour incubation period, no significant reduction in cell viability was observed in RG/C2 or HCT116 cells treated with *B. fragilis*, *E. faecalis* or *F. nucleatum* (Figure 3.3). However, a reduction in cell viability was observed in RG/C2 cells co-cultured with *E. coli* Nissle (Figure 3.3B), with a 40% reduction observed at an MOI of 10:1 and 100:1. Conversely, absorbance in wells containing HCT116 cells increased as a result of bacterial co-culture, with significant increases observed in cells cultured with *E. coli* Nissle (Figure 3.3B) and *E. faecalis* (Figure 3.3C) at an MOI of 1,000:1. Over the course of the assay, monolayers were also visually inspected under a light microscope to monitor observable changes in cell morphology and appearance. However, no visible changes to the tumour cell monolayers were observed under any condition (data not shown). As the 1-hour MTS reagent incubation period would not have been sufficient for cell division to occur, the observed increases in tetrazolium reduction are hypothesized to be as a result of altered cellular metabolism in response to bacterial infection.
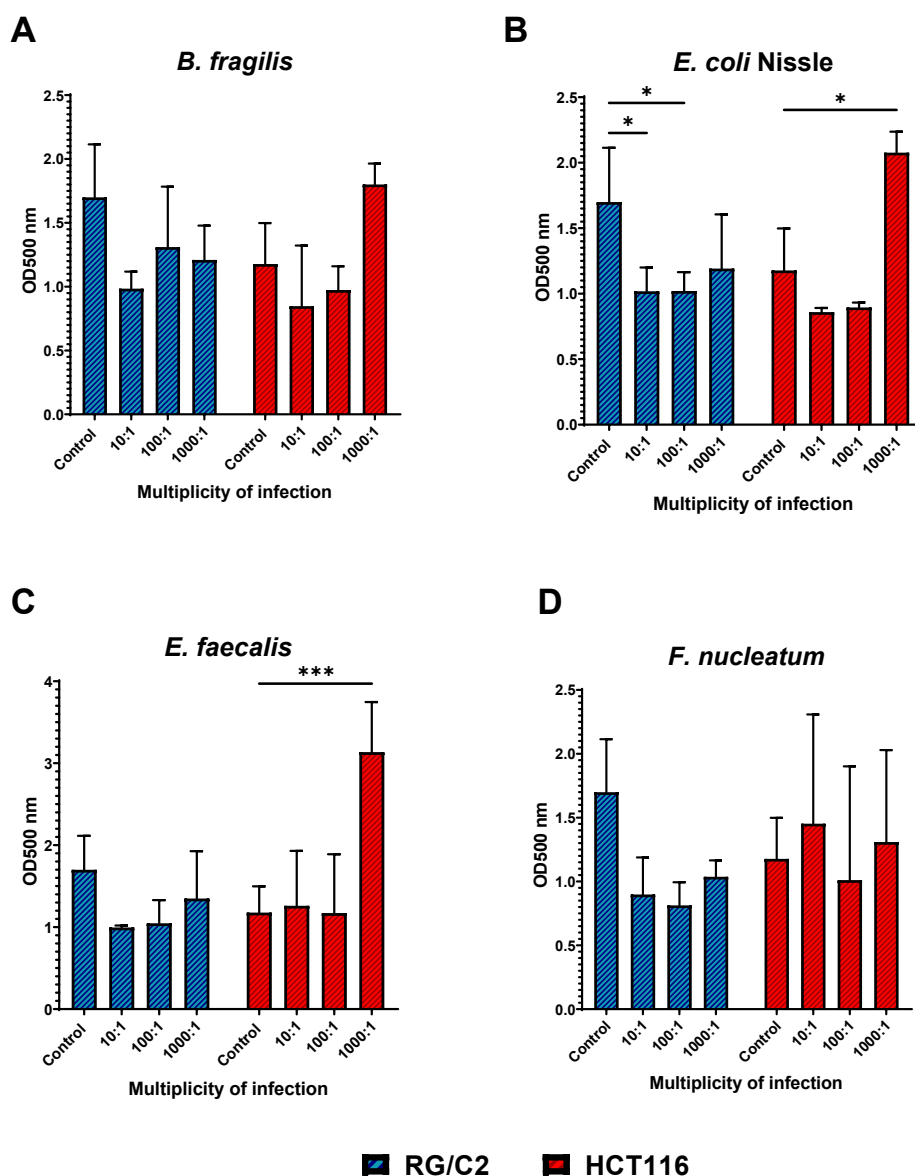
Figure 3.3: Viability of RG/C2 and HCT116 monolayers after bacterial co-culture. The effect of *B. fragilis* (A), *E. coli* Nissle (B), *E. faecalis* (C) and *F. nucleatum* (D) on cell viability was measured using the MTS assay. Data expressed as normalised absorbance at an optical density of 500 nm. Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate. (T-test: * = $p < 0.05$, *** = $p < 0.001$).

3.4.3. Bacterial attachment to colorectal tumour cell monolayers is dependent upon tumour cell type

To study the interactions between tumour cell monolayers and bacteria, RG/C2 and HCT116 cells were grown in 96-well plates and infected with bacteria at an MOI of 1,000:1. Prior to infection, confluency was confirmed by visual inspection under a light microscope. The survival and viability of the selected bacterial species in DMEM cell culture media under aerobic conditions was confirmed via serial dilution and Miles and Misra counting after a 4-hour incubation (Supplementary Figure 1). This revealed that all species remained viable for the entire 4-hour duration. Following this, the gentamicin protection assay was used to quantify bacterial attachment and invasion of tumour cell monolayers (see Materials and Methods 2.3). The total number of bacteria attached to the monolayer surface was calculated by subtracting intracellular bacterial cfu from the total associated bacterial cfu. Both attachment and invasion were expressed as the total cfu present in the whole lysate.

These results show that all four species readily adhered to the surface of RG/C2 and HCT116 tumour cells (Figure 3.4). Bacteria were present on the cell surface in the range of $1x10^4$-$1x10^7$ total cfu, with approximately $3x10^5$ tumour cells present in both RG/C2 and HCT116 monolayers. Therefore, the estimated number of bacteria per cell ranged from 0.03-33 across all species. Despite all bacteria being initially added to the culture at an MOI of 1000:1, replication of the facultative anaerobes *E. coli* Nissle and *E. faecalis* during aerobic incubation led to an increased bacterial load at the end of treatment. This may be responsible for the much higher total cfu when compared with the strict anaerobes *B. fragilis* and *F. nucleatum*, making direct comparison between bacterial species difficult. However, when analysing each species in isolation, clear differences were observed in bacterial attachment to RG/C2 and HCT116

monolayers. Attachment of *B. fragilis* (p = 0.0015), *E. coli* Nissle (p = 0.0049), *E. faecalis* (p = 0.050) and *F. nucleatum* (p < 0.001) was significant higher for RG/C2 monolayers compared to HCT116 monolayers. An increase of >50% was observed across each species, suggesting that the cellular characteristics specific to the benign RG/C2 cell line are able to promote the adhesion of bacterial cells to their surface.
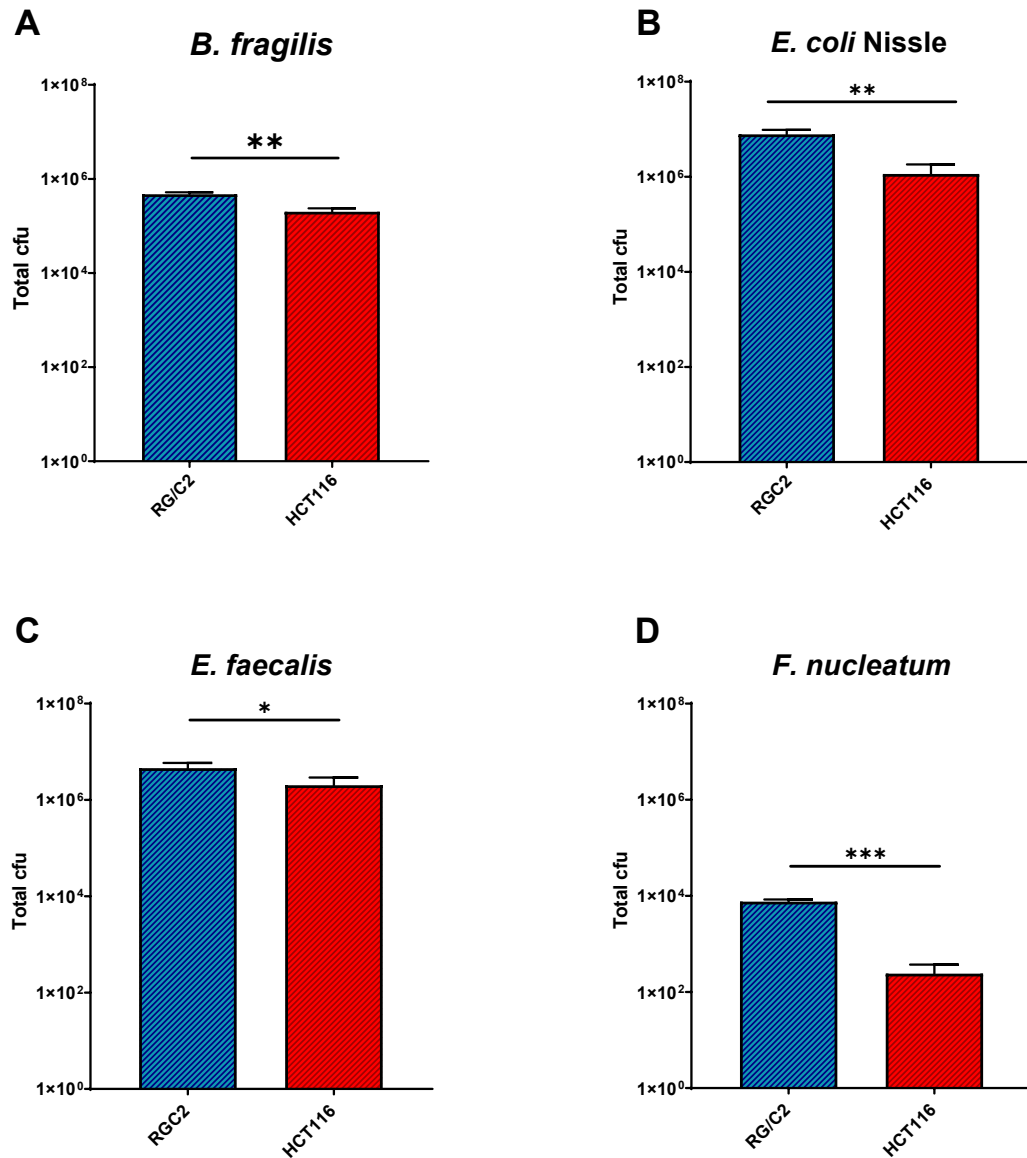
Figure 3.4: Attachment of bacteria to colorectal tumour cell monolayers. *B. fragilis* (A), *E. coli* Nissle (B), *E. faecalis* (C) and *F. nucleatum* (D) were co-cultured with tumour cell monolayers at an MOI of 1000:1 for 4 hours. After incubation, monolayers were washed with PBS to remove non-associated bacteria, lysed and serial dilutions were enumerated onto blood agar plates. Total attachment was calculated by subtracting intracellular bacteria from total associated bacteria. Bacterial numbers are expressed as total colony forming units present in the whole lysate. Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate. (T-test: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$).

3.4.4 Imaging the attachment of bacteria to HCT116 cells

To visualise the attachment of bacteria to tumour cells, HCT116 cells were grown to 60-70% confluency on 10 mm glass coverslips and were imaged via scanning electron microscopy (SEM). The dehydration process required for SEM sample preparation causes cell shrinking, which can result in the tearing of cells in regions of high cell density. Therefore, only regions containing clearly identifiable isolated cells were chosen for imaging. Figure 3.5 illustrates that these isolated HCT116 cells display a rounded morphology, with small projections on the upper cell surface providing a larger surface area. Longer, filament-like projections appear to be anchoring the cells to the slide surface. These structures are thought to be filopodia; however, many of these structures appear to have been damaged and/or distorted by dehydration.

All four species tested clearly adhered to the cell surface (Figure 3.5, white arrows), and this adhesion was strong enough to remain intact during washing, fixation, and dehydration. *B. fragilis* and *F. nucleatum* (Figure 3.5B and 3.5E) appeared to attach to HCT116 cells in small numbers, in contrast to the multicellular aggregates of *E. coli* Nissle and *E. faecalis* (Figure 3.5C and 3.5D, black arrows). These aggregates, in addition to the visible bacterial cells present on the glass surface, would suggest that co-aggregation between bacterial cells plays a key role in facilitating the attachment of these species in large numbers. However, it was also demonstrated that these species multiply under the experimental conditions (Supplementary Figure 1), and it is therefore possible that these aggregates are the result of bacterial cell division. *F. nucleatum* cells ranged in size from 5-30 μm, in contrast to *B. fragilis*, *E. coli* Nissle and *E. faecalis* cells which were approximately 1 μm in diameter. Furthermore, *F. nucleatum* attachment appeared to span multiple tumour cells (Figure 3.5E). Large *F. nucleatum* cells were observed in contact with

multiple tumour cells and the surface of the glass coverslip, whereas cells of approximately 5 µm in length were observed in full contract with the tumour cell surface. This suggests that individual cell morphology may be an important determinant of *F. nucleatum* interaction with epithelial cells. No rupture of HCT116 membranes or other visible signs of cell damage were observed in bacteria-treated samples.
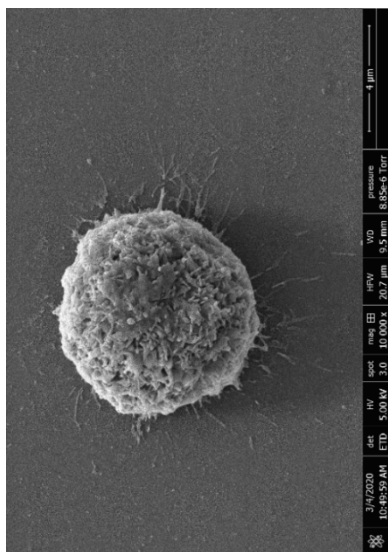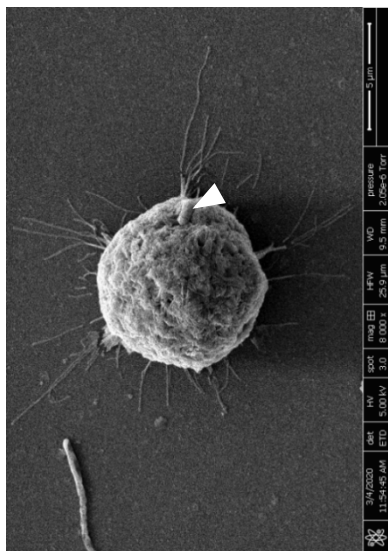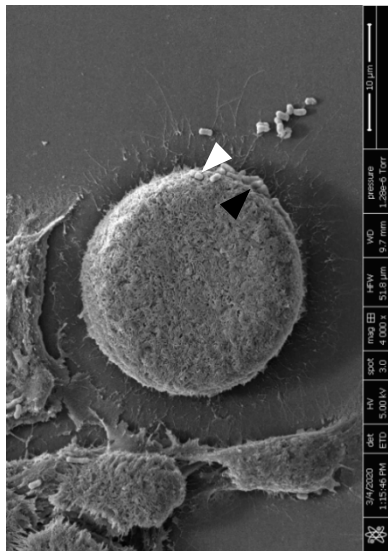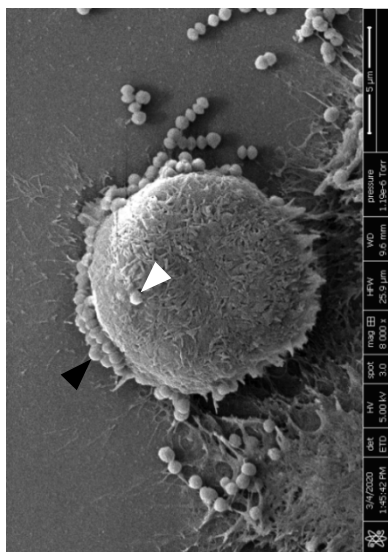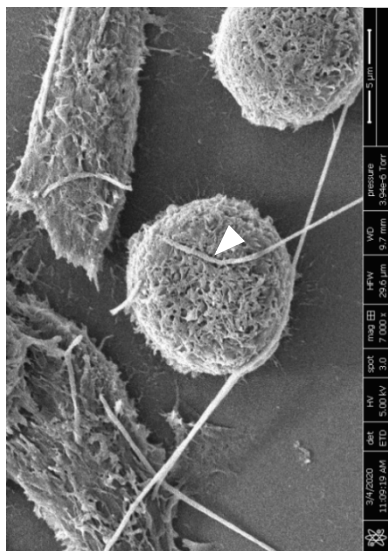
Figure 3.5: Scanning electron micrographs of HCT116 cells after co-culture with various species of gut bacteria. Representative images of isolated tumour cells demonstrating bacterial attachment (white arrows) and co-aggregation of *E. coli* Nissle and *E. faecalis* (black arrows). Cells were cultured on 1 mm glass coverslips at the bottom of a 24-well plate before 4-hour co-culture with bacteria. Samples were then fixed and dehydrated with increasing concentrations of ethanol and HMDS, before gold sputter coating and imaging.

3.4.5. RG/C2 and HCT116 colorectal tumour cells are readily invaded by gut bacteria

The gentamicin protection assay was also used to quantify the number of bacteria which invaded colorectal tumour cells. After a 4-hour infection period, monolayers were washed with PBS before being treated with 300 µg/ml gentamicin and 200 µg/ml metronidazole to eliminate extracellular bacteria. Antibiotic concentrations were based upon recommendations from the literature and optimised before beginning the assay (Rubinstein *et al.*, 2013; Supplementary Figure 2). All four species tested were able to invade both RG/C2 and HCT116 cells. Invasion of *B. fragilis* was higher in RG/C2 monolayers, whereas *E. coli* Nissle, *E. faecalis* and *F. nucleatum* showed significantly higher invasion when co-cultured with HCT116 monolayers (Figure 3.6). Invasion by all species was approximately 100-fold lower than attachment after a 4-hour infection period, with the exception of *F. nucleatum* where more intracellular bacteria were recovered than surface bacteria in HCT116, and numbers were comparable for RG/C2.

To visualise bacterial invasion, bacteria were pre-stained with carboxyfluorescein (CFSE) before being added to tumour cell cultures. CFSE effectively stained all four species tested, however, there was some variability both between species and between samples. The anaerobic species *B. fragilis* and *F. nucleatum* showed the best stain retention, whereas *E. coli* Nissle and *E. faecalis* staining was less clear and exhibited increased background (e.g. Figure 3.8D). In some cases, non-stained *E. coli* Nissle and *E. faecalis* where observed, evidenced by DAPI staining without CFSE detection (not shown). This could be attributed to incomplete

staining of the initial culture or bacterial cell division generating unstained daughter cells.

Figures 3.7-3.10 illustrate representative CLSM micrographs taken of infected tumour cells immediately after a 4-hour infection period. Intracellular bacteria were detected between the stacks stained with phalloidin, definitively proving that the bacteria are either intracellular, or are within the intercellular space. In all cases (e.g. Figure 3.7B, 3D reconstructions of Z-stacks demonstrated that bacteria are able to penetrate the cell colony to reach the basal surface. As described in Section 3.1, the morphology of RG/C2 and HCT116 cells differs, with RG/C2 cells growing in colonies 2-4 cells deep, in comparison to HCT116 cells which grow as single cell-thick monolayers. In infected RG/C2 cells, bacteria were detected at multiple locations within the cell colony, best illustrated by *F. nucleatum*-infected cells (Figure 3.10B). The increased cell density and intercellular spaces may represent an additional niche provided by benign tumour cells for bacteria, or could be an indicator of increased invasion time needed to reach the basal layer due to the density of the cell layer. In summary, CFSE staining provided an inexpensive and simple proof of the presence of intracellular bacteria, particularly for anaerobic species.
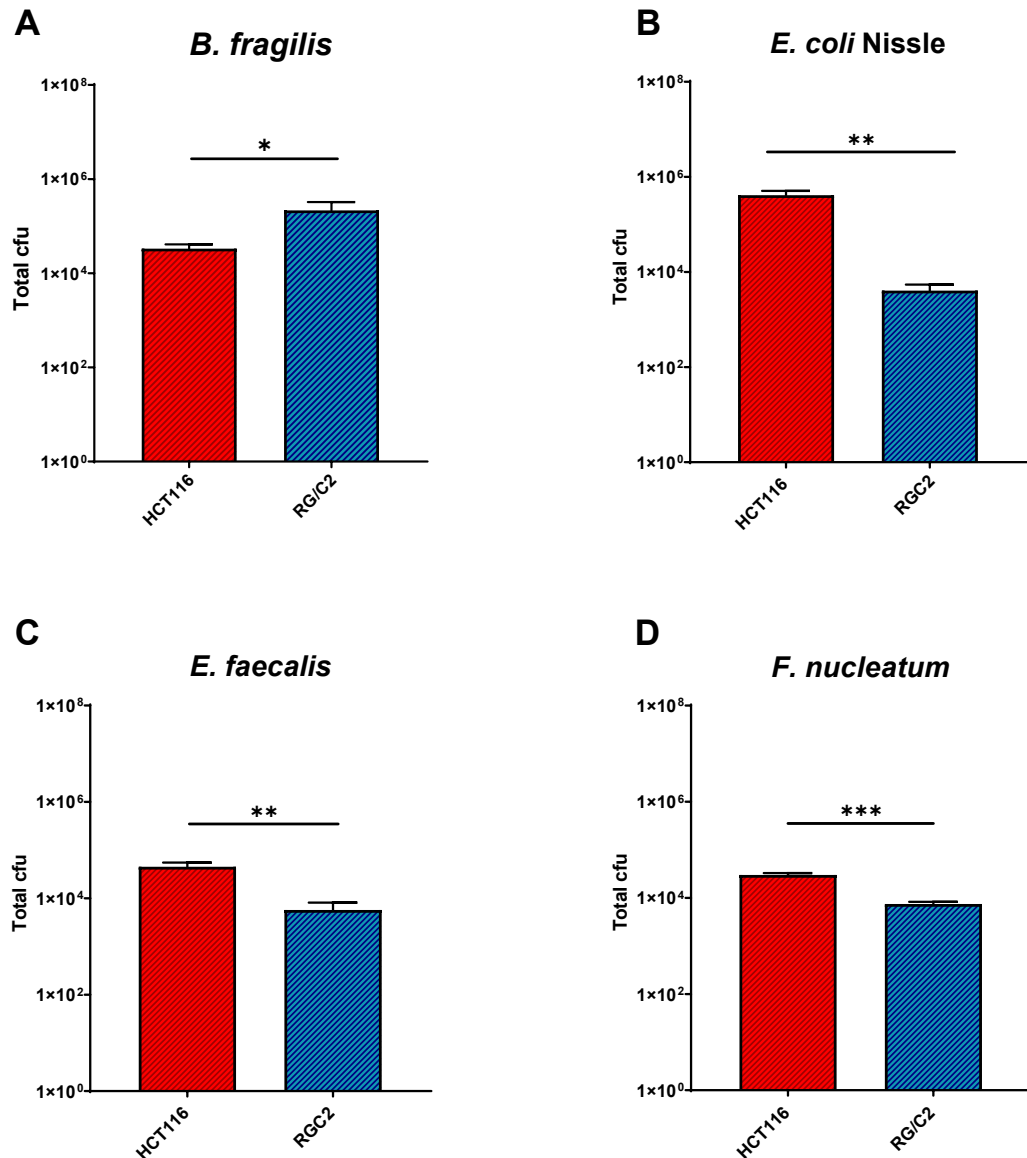
Figure 3.6: Invasion of gut bacteria in RG/C2 and HCT116 tumour cell monolayers. Cells were cultured with *B. fragilis* (A), *E. coli* Nissle (B), *E. faecalis* (C) and *F. nucleatum* (D) for 4 hours, before being washed and treated with 300 μg/ml gentamicin and 200 μg/ml metronidazole for 1 hour to eliminate extracellular bacteria. Monolayers were then washed and lysed, with serial dilutions being enumerated onto blood agar plates. Bacterial numbers are expressed as total colony forming units present in the whole lysate. Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate. (T-test: $* = p < 0.05$, $** = p < 0.01$, $*** = p < 0.001$).
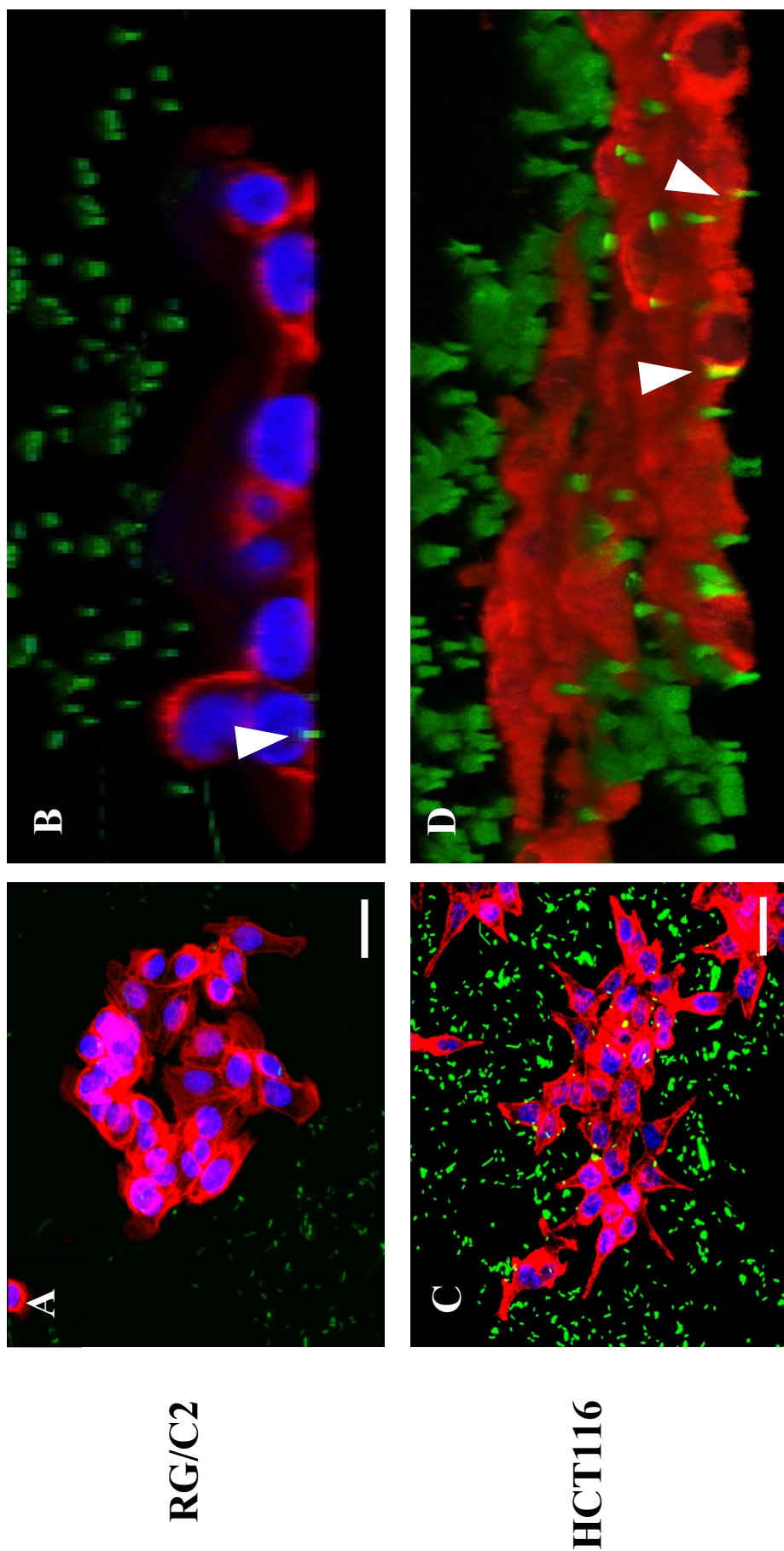
Figure 3.7: Confocal microscopy images of *B. fragilis* infection. Representative confocal Z-stacks of RG/C2 (A) and HCT116 (C) cells after infection with *B. fragilis*. ImageJ software was used to 3D model confocal Z stacks to better visualise intracellular bacteria (B and D). Blue = cell nuclei stained with DAPI, Red = actin cytoskeleton staining with phalloidin, Green = bacteria stained with CFSE. White arrows indicate intracellular bacteria. Scale bars represent 20 μm.
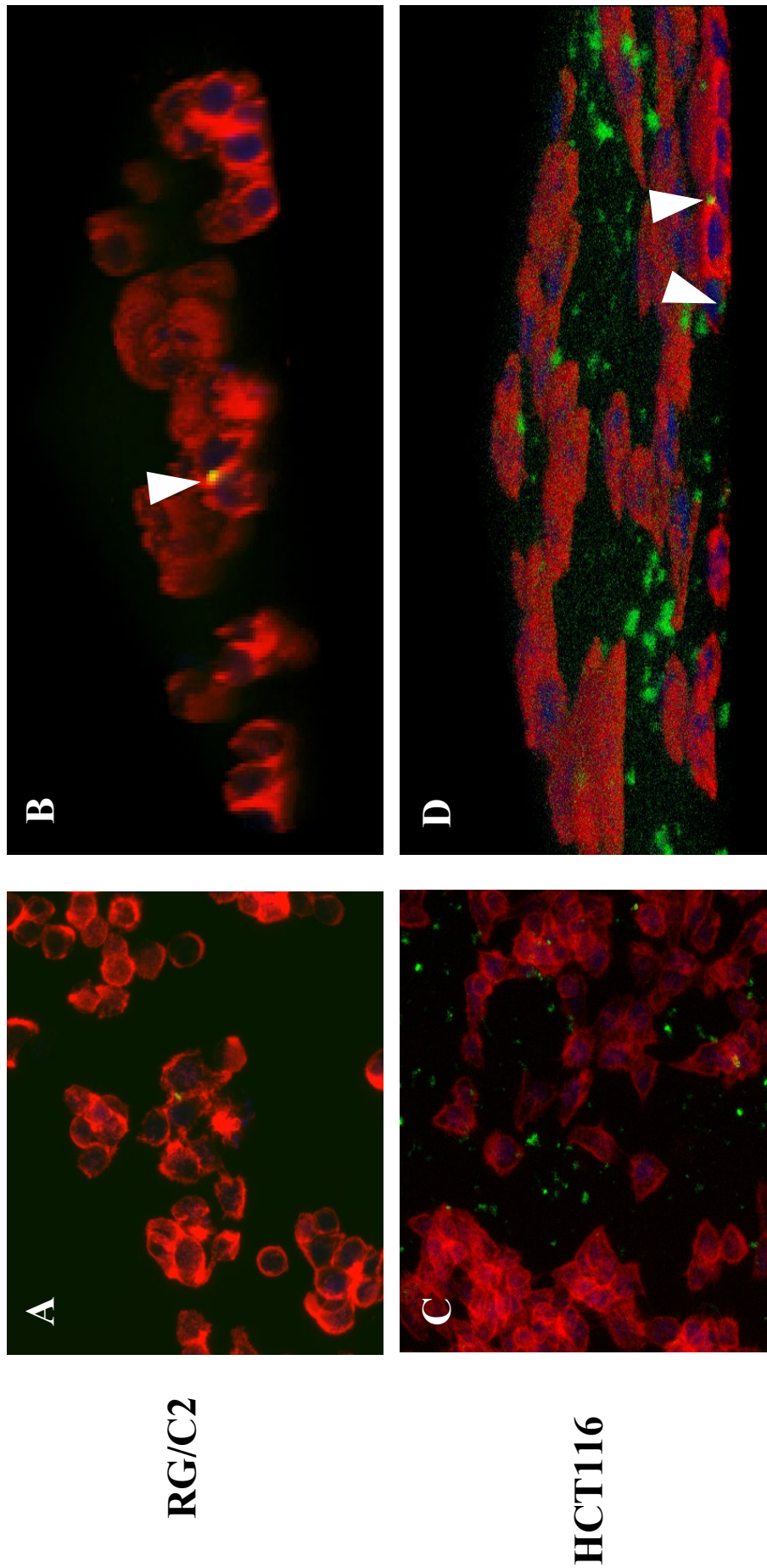
Figure 3.8: Confocal microscopy images of *E. coli* Nissle infection. Representative confocal Z-stacks of RG/C2 (A) and HCT116 (C) cells after infection with *E. coli* Nissle. ImageJ software was used to 3D model confocal Z stacks to better visualise intracellular bacteria (B and D). Blue = cell nuclei stained with DAPI, Red = actin cytoskeleton staining with phalloidin, Green = bacteria stained with CFSE White arrows indicate intracellular bacteria. Scale bars represent 20 μm.
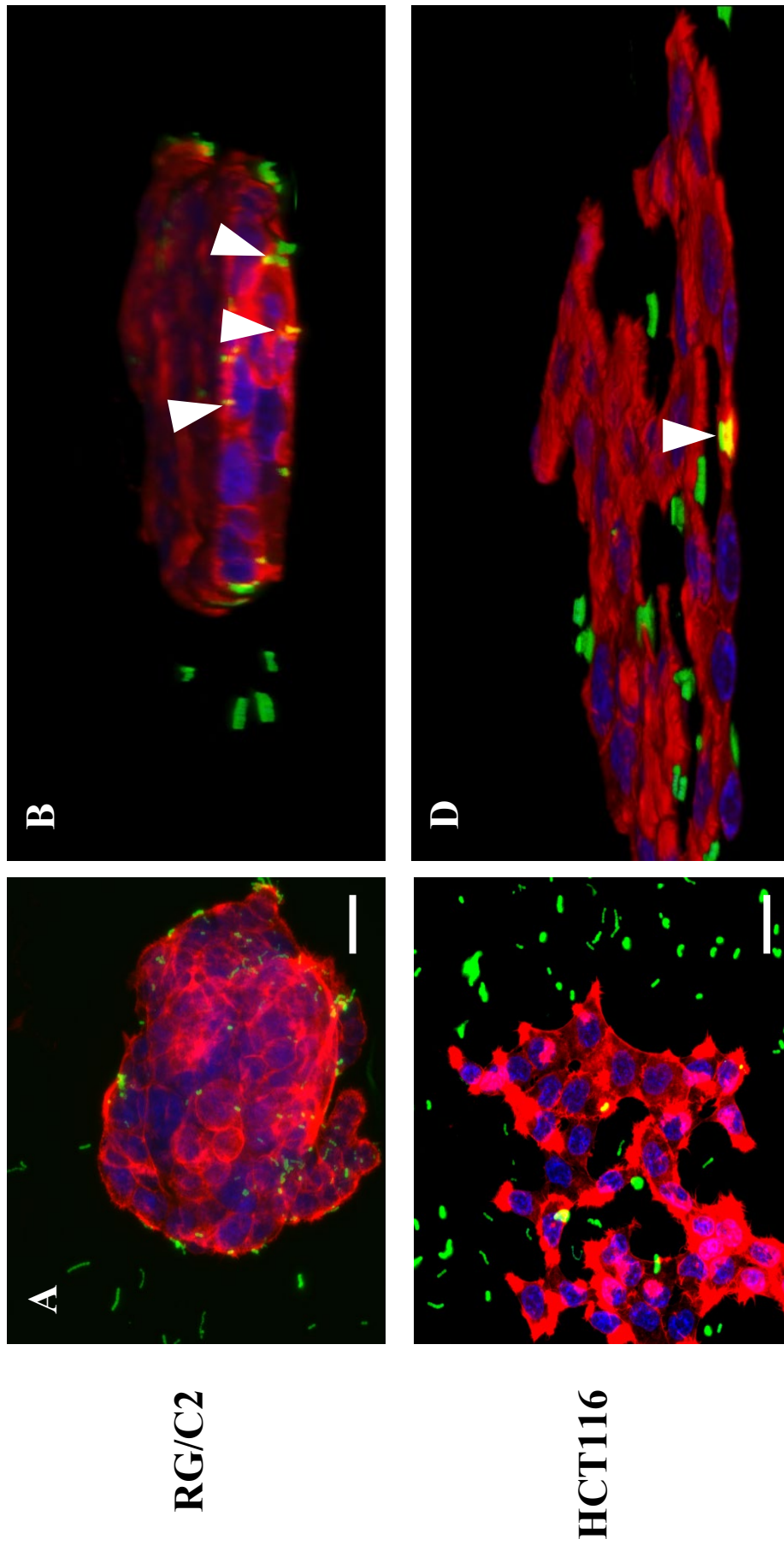
Figure 3.9: Confocal microscopy images of *E. faecalis* infection. Representative confocal Z-stacks of RG/C2 (A) and HCT116 (C) cells after infection with *E. faecalis*. ImageJ software was used to 3D model confocal Z stacks to better visualise intracellular bacteria (B and D). Blue = cell nuclei stained with DAPI, Red = actin cytoskeleton staining with phalloidin, Green = bacteria stained with CFSE White arrows indicate intracellular bacteria. Scale bars represent 20 μm.
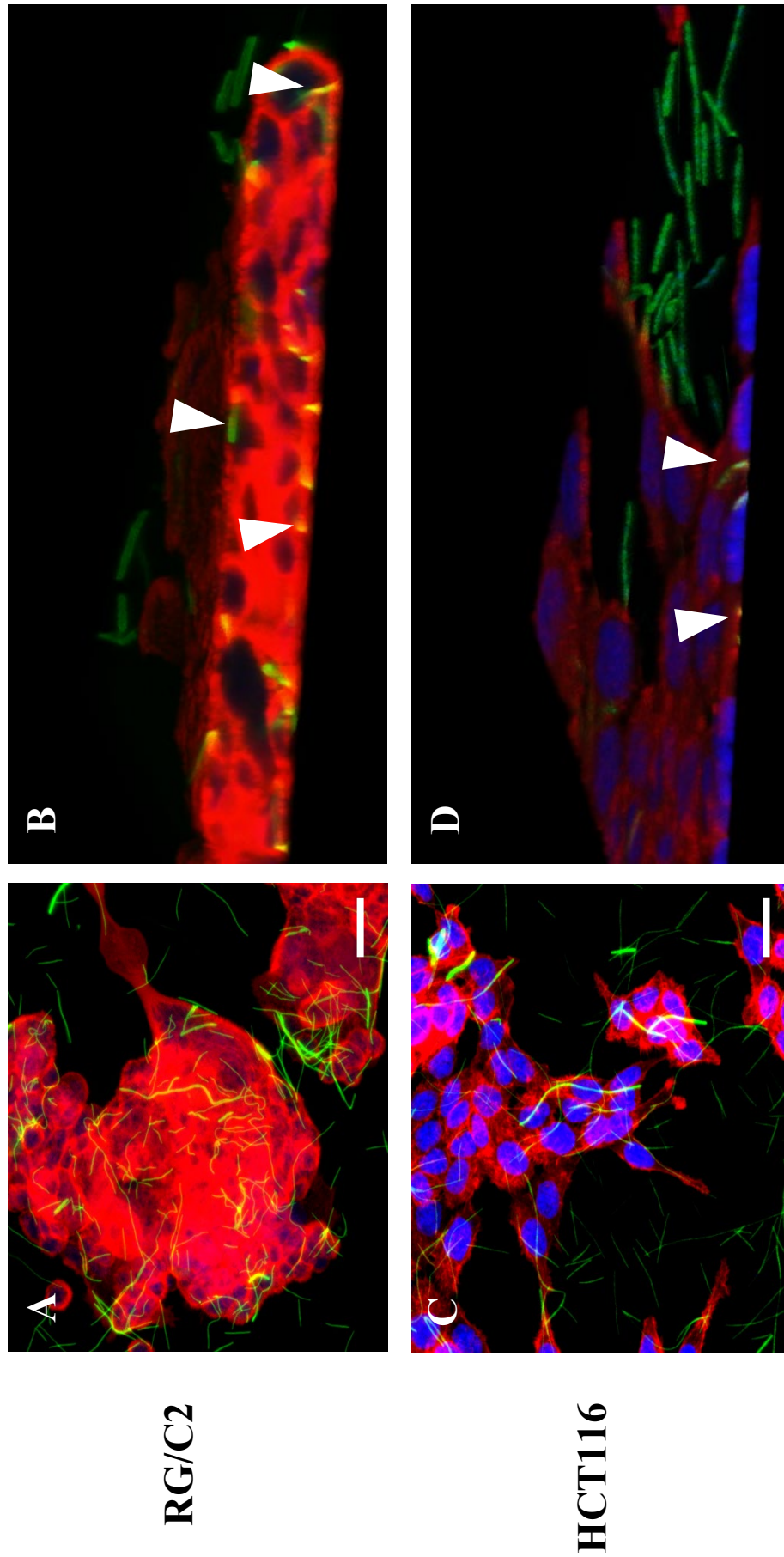
Figure 3.10: Confocal microscopy images of *F. nucleatum* infection. Representative confocal Z-stacks of RG/C2 (A) and HCT116 (C) cells after infection with *F. nucleatum*. ImageJ software was used to 3D model confocal Z stacks to better visualise intracellular bacteria (B and D). Blue = cell nuclei stained with DAPI, Red = actin cytoskeleton staining with phalloidin, Green = bacteria stained with CFSE White arrows indicate intracellular bacteria. Scale bars represent 20 µm.

3.4.6 Intracellular bacteria persist within colorectal tumour cells

Once invasion of tumour cell monolayers had been established, the ability of the bacteria to survive and remain viable within tumour cells over extended time periods was assessed. After co-culturing cell monolayers and bacteria for 4 hours, extracellular bacteria were eliminated with antibiotics. Fresh media was added, and total intracellular bacteria were quantified every 24 hours (see Materials and Methods 2.3). Monolayers were visually inspected for signs of damage/cell death, with bacterial infection having no visually observable effects on tumour cell morphology or appearance over the 96-hour incubation period.

Monolayers infected with *B. fragilis* experienced no significant change in bacterial load for the first 48 hours (Figure 3.10A). After this time, there was a significant decrease in the remaining viable bacteria in RG/C2 monolayers ($p = 0.022$). In comparison, there was no significant reduction in *B. fragilis* within HCT116 monolayers until 96 hours ($p = 0.013$). The total viable cfu of *E. coli* Nissle did not significantly change in either cell line over the course of the 96-hour incubation period (Figure 3.11B). Similarly, *E. faecalis* cfu did not significantly decrease in either cell line until 96 hours post-infection in RG/C2 cells ($p = 0.0004$). *F. nucleatum* exhibited the lowest overall survival of the tested species across both cell lines, with no viable cfu recovered in RG/C2 monolayers after 72 hours or in HCT116 monolayers after 96 hours (Figure 3.11D). With the exception of a non-significant increase *B. fragilis* cfu recovered from HCT116 monolayers between 24-72 hours, no increases in bacterial cfu were recorded for any species. This suggests that bacteria are not able to replicate within the tumour cells.

As highlighted in Figure 3.6, initial invasion of RG/C2 and HCT116 monolayers was statistically different in all four species, and this corresponds to the

0-hour time point presented in Figure 3.11. To determine whether the survival of intracellular bacteria was statistically different in RG/C2 and HCT116 monolayers, a two-way ANOVA was performed to analyse the relationship between bacterial survival and host tumour cell line, accounting for differences in total starting cfu. With the exception of *E. coli* Nissle, where the tumour cell line had no effect on bacterial survival ($p = 0.756$), survival of each species was significantly higher in HCT116 monolayers (p values illustrated on Figure 3.11). These results demonstrate that once infection of tumour cell monolayers has occurred, all species tested are able to survive within the intracellular space under standard cell culture conditions for a minimum of 24 hours. Therefore, viable intracellular bacteria will be present within tumour cells throughout the duration of all future downstream experiments.
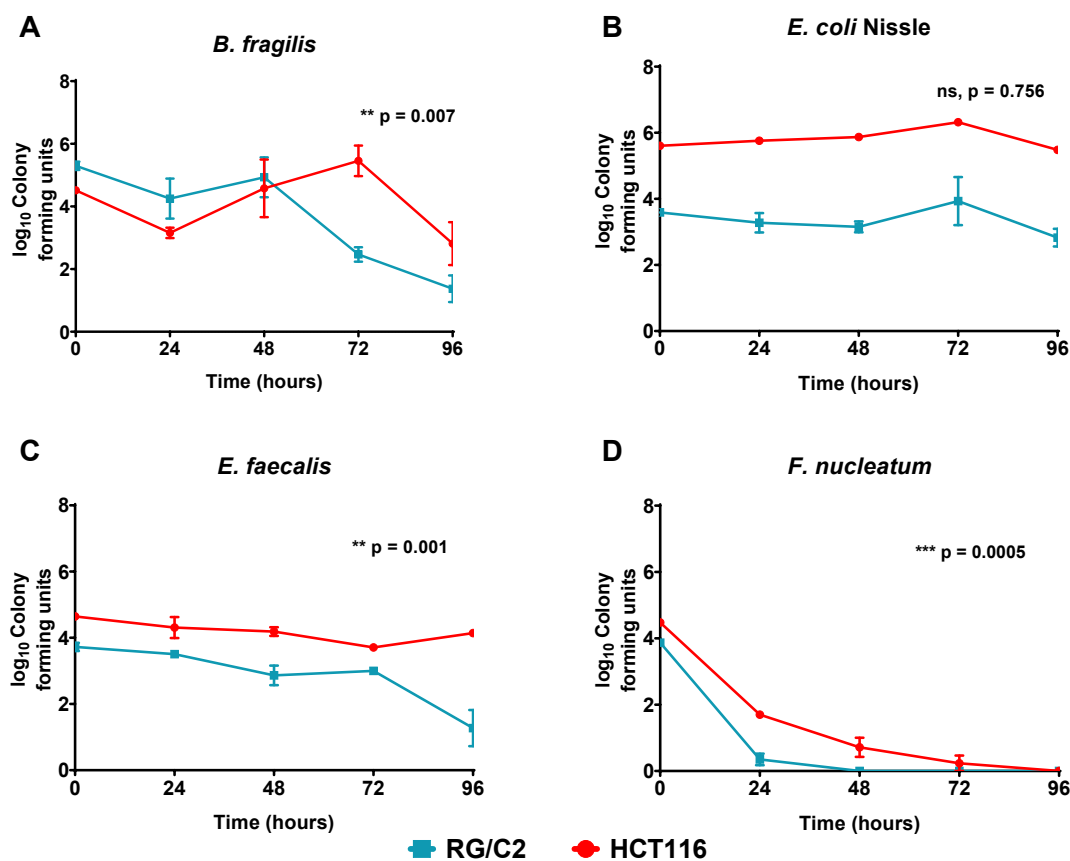
Figure 3.11: Persistence of bacterial within colorectal tumour cell monolayers. After initial infection with bacteria, co-cultures were continued in the presence of antibiotics to prevent the growth of extracellular bacteria. Monolayers were lysed at 24-hour intervals and viable bacteria enumerated onto blood agar plates. The survival of each species within RG/C2 and HCT116 monolayers was compared using a two-way ANOVA (inset). Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate. (Two-way ANOVA: $* = p < 0.05$, $** = p < 0.01$, $*** = p < 0.001$).

3.4.7. Gastrointestinal mucin has the potential to promote biofilm formation *in vitro*

In the colon, the epithelium is protected by two layers of mucus which serve as a substrate for bacterial growth and acts as a barrier to limit direct contract between the epithelium and the microbiota (Schroeder, 2019). To investigate whether the addition of mucins could promote a more *in vivo*-like phenotype in the tested bacteria, the effect of mucins on biofilm formation was quantified using the crystal violet biofilm formation assay (O'Toole, 2011). Three species, *B. fragilis*, *E. coli* Nissle and *F. nucleatum* displayed a dose-dependent increase in biofilm formation when grown in the presence of porcine mucin. In all cases, the increase in biofilm formation was statistically significant at a mucin concentration of 1.0%, compared to the 0% control (Figure 3.12). *F. nucleatum* biofilm formation was also significantly increased at lower mucin concentrations of 0.4, 0.6 and 0.8% (Figure 3.12D). Interestingly, in the absence of mucin, *F. nucleatum* formed weak biofilms under these experimental conditions. Mucin had no effect on the biofilm forming capabilities of *E. faecalis*, which despite having the highest absorbance readings showed a high of degree variability (Figure 3.12C). These data suggest that the relative abundance of mucin at the mucosal surface may promote bacterial biofilm formation but is not essential for all species.
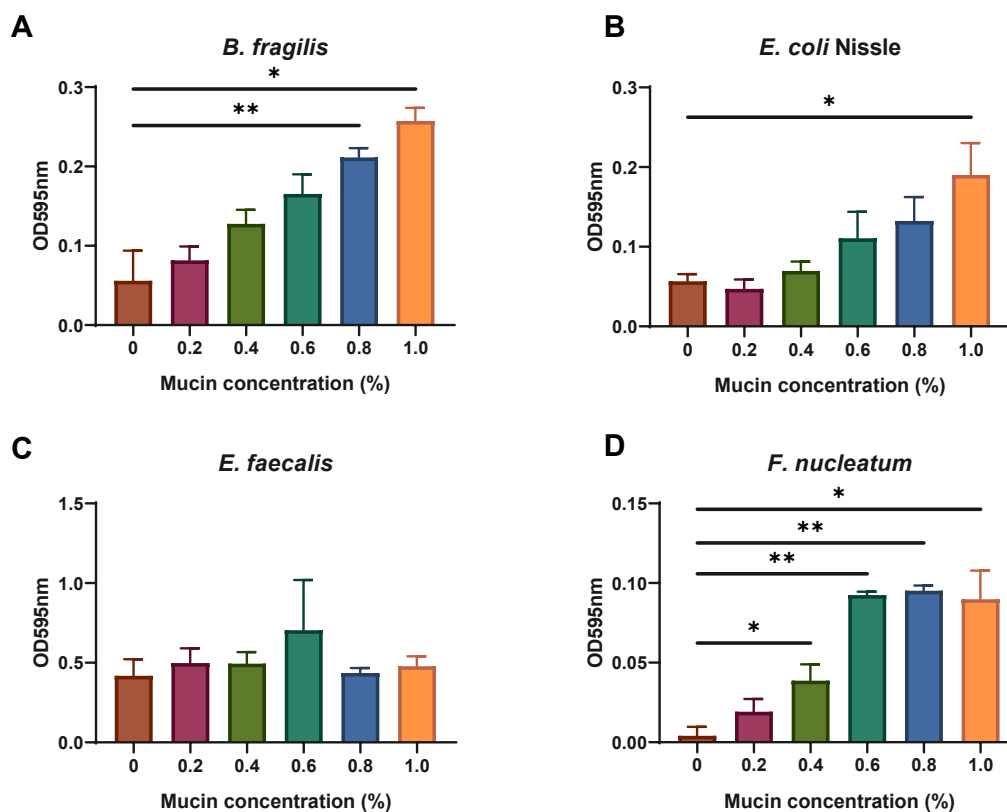
Figure 3.12: The effect of mucin on bacterial biofilm formation. *B. fragilis* (A), *E. coli* Nissle (B), *E. faecalis* (C) and *F. nucleatum* (D) were cultured in BHI broth containing increasing concentrations of porcine mucin type II. Biofilms were quantified using the crystal violet staining method after 48 hours of growth under anaerobic conditions. Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate. (T-test: * = $p < 0.05$, ** = $p < 0.01$).

3.5 Discussion

The adenoma-carcinoma sequence in CRC normally unfolds over several decades, and the colonic mucosa is in constant contact with the microbiome throughout this process (Vogelstein and Kinzler, 1993). CRC has been associated with changes in the relative abundance of microbiota species (dysbiosis), and colonisation by species that are not present in the normal colonic microbiota e.g. *F. nucleatum* (Chen, Domingue and Sears, 2017). In CRC bacteria are found in closer proximity to the mucosa, at increased abundance in privileged sites such as the colonic crypts, and directly within tumour tissue (Weinberg *et al.*, 2020). The formation of mucosal biofilms characterised by the presence of highly invasive species has also been implicated in CRC development (Drewes *et al.*, 2017). Bacterial invasion of the colonic epithelium involves the targeting of cell surface and junction proteins such as E-cadherin, which triggers pro-tumorigenic downstream effects (Guo *et al.*, 2020). Furthermore, the presence of invasive bacteria in direct contact with the epithelium has been shown to correlate with indirect mechanisms of microbe-induced tumorigenesis including induction of chronic inflammation and remodelling of the microbiota (Strauss *et al.*, 2011; Hajishengallis, Darveau and Curtis, 2012). Distinct microbial populations are associated with different stages of CRC progression (Mori *et al.*, 2018); and as CRC progresses, changes in cell architecture and the tumour microenvironment may alter the interactions of these invasive bacteria with the epithelium (Li *et al.*, 2020). The Driver-Passenger model suggests that these stage-specific changes may alter the phenotype of tumour-associated microorganisms (e.g. increased toxin production), and that specific species may act as drivers or passengers at different stages of CRC (Tjalsma *et al.*, 2012). Therefore, the interactions between bacteria and benign adenomas may be vastly different from those with malignant adenocarcinomas.

The main aim of this chapter was to investigate whether the interactions of three species associated with CRC and one probiotic species differed between the benign RG/C2 and malignant HCT116 cell lines. As demonstrated in Figure 3.1, there is a marked difference in the morphology of adenoma and adenocarcinoma cells cultured *in vitro*, suggesting that the bacteria-tumour relationships may shift during this transition due to the altered surface architecture and presence of cell surface molecules. Adenocarcinoma cell lines rapidly reached confluency in the culture vessel, growing in uniform two-dimensional layers. However, the RG/C2 cell line displayed slower growth, with irregular colonies. Initial adherence of RG/C2 cells to the surface of the culture plate was also lower, resulting in an extended lag time for cell growth. The confocal micrographs presented in Figure 3.2 illustrate how HCT116 cells remain as single-cell thick cultures regardless of their confluence. In comparison, RG/C2 cells readily grow outwards on the Z axis, often before spreading outwards to cover the surface of the culture vessel.

Before beginning to quantify how bacteria attach to and invade colorectal tumour cell monolayers, it was important to establish that the selected bacterial species remained viable under the experimental conditions. It was empirically determined that all test bacterial species could survive within DMEM cell culture medium (Supplementary Figure 1). Furthermore, suspension in cell culture media did not impact the viability of the obligate anaerobes during the 4-hour infection period. This is supported by current literature which states that both *B. fragilis* and *F. nucleatum* can tolerate oxygen at low concentrations. *B. fragilis* has been demonstrated to survive and be able to replicate at nanomolar levels of oxygen (Baughn and Malamy, 2004). Furthermore, tolerance to atmospheric oxygen has been suggested as a key

characteristic of *F. nucleatum* which enables it to act as a 'bridge organism' that allows stricter anaerobes to colonise the oral cavity (Okuda *et al.*, 2012).

The impact of bacterial infection on colonic tumour cell viability was also investigated. Tumour cell death during the gentamicin protection assay may cause inaccurate estimations of bacterial attachment and invasion; and it was also important to establish the feasibility of using infected tumour cells in downstream experiments. The MTS assay demonstrated that no significant reduction in tumour cell viability occurred when tumour cells were co-cultured with bacteria. This is surprising, as some adverse effect on tumour cell viability would be expected during bacterial infection. *B. fragilis, E. coli* Nissle and *F. nucleatum* are Gram negative bacteria, with outer cell walls containing lipopolysaccharide (LPS) which has been demonstrated to induce the production of pro-inflammatory mediators TNF-α and IL10 in Caco-2 monolayers (Stephens and von der Weid, 2020). Furthermore, production of LPS by *B. fragilis* is significantly higher under aerobic conditions (Chmelař *et al.*, 2016). *E. faecalis* is also able to induce cell damage through the production of genotoxic compounds e.g. superoxide (Wang *et al.*, 2008). The lack of a reduction in cell viability observed through the MTS assay, or via microscopic inspection, suggests that the tumour cells tested are not significantly impacted by the aforementioned factors, or that these factors are not apparent under these experimental conditions. It is however important to consider that the MTS assay is based upon cellular metabolism and is therefore an indirect measure of cell viability. It is possible that stress placed on tumour cell monolayers by bacterial infection may have triggered metabolic shifts, which may obscure the true effect on viability. To account for this, an additional measure of cell viability could have been performed such a CellTiter-Glo or Annexin V staining. Despite the limitations of the MTS assay, the lack of any obvious indicator of a

detrimental effect on tumour cell viability provided confidence in the suitability of the model for investigating bacteria-tumour cell interactions and the resulting effects on tumour cell behaviour.

Once it was established that both bacteria and tumour cells remained viable after 4 hours of co-culture under standard cell culture conditions, bacterial attachment and invasion of tumour cell monolayers were quantified. For all species tested, bacterial adhesion to colorectal tumour cells was significantly higher in RG/C2 co-cultures, although significant adhesion to HCT116 monolayers was also observed. Many of the species investigated in this study have previously been reported to bind to E-cadherin, expression of which is higher in colorectal adenomas than adenocarcinomas (Rees and Qualtrough, unpublished data). Furthermore, this increased attachment could be attributed to a possible larger surface area of RG/C2 colonies, which in contrast to the flat monolayers formed by HCT116 cells grow outwards and form three dimensional structures. One caveat of the co-culture model is the proliferation of facultative anaerobes during incubation impacting cfu counts. Additionally, Figure 3.5 demonstrates that co-aggregation of *E. coli* Nissle and *E. faecalis* cells also occurs at the tumour cell surface. It is unclear what effect bacterial replication and co-aggregation may have on the cfu counts for bacterial attachment; however, their proximity to the tumour cell surface means that they may remain capable of influencing tumour cell behaviour through signalling to adjacent bacteria and secretion of metabolites.

Invasion of tumour cell monolayers was assessed using the gentamicin protection assay. Invasion of HCT116 monolayers by bacteria was higher than RG/C2 monolayers for all species, with the exception of *B. fragilis*. It could be expected that higher levels of attachment to the tumour cell surface would facilitate increased

invasion into the cell (see Figure 3.4); however, it is possible that invasion of HCT116 cells occurs at a faster rate, whereas bacteria are blocked from entering RG/C2 cells. Also surprising was the apparent virulence of the probiotic *E. coli* Nissle, which displayed both adhesion and invasion to both cell lines. However, sepsis induced by *E. coli* Nissle has been previously reported (Guenther *et al.*, 2010), and its pathogenicity is thought to be heavily influenced by the microbiota and immune status of the host (Gronbach *et al.*, 2010).

Evidence of invasion was also demonstrated via CLSM. In all cases, bacteria were found within the tumour cell monolayer; however, it is unclear whether the detected bacteria were present within the cytoplasm or within the intercellular space. Nevertheless, *in vitro* this invasion protects the bacteria from antibiotics used throughout the gentamicin protection assay and atmospheric oxygen. *In vivo,* this intra-tumour environment may provide bacteria with a metabolic niche and protection from the immune system. Several bacteria cells could be seen at the basal region of the tumour monolayer, highlighting the ability of these bacteria to rapidly penetrate through multiple tumour cells within the 4-hour incubation period. CRC formation is associated with aberrant tight junction regulation and an increase in intestinal barrier permeability which may facilitate bacterial invasion of underlying tissue (Soler *et al.*, 1999). Furthermore, bacteraemia caused by CRC-associated species, including *B. fragilis* and *F. nucleatum*, has been linked to subsequent CRC diagnosis (Kwong *et al.*, 2018). This suggests that the increased prevalence of species such as *B. fragilis* and *F. nucleatum* may be attributed to changes in the tumour microenvironment which facilitate bacterial proliferation and dissemination.

When investigating invasion of tumour cells by different bacterial species, it is important to consider that methods of entry into epithelial cells may differ. The

gentamicin protection assay is unable to discriminate between active bacterial invasion and phagocytosis of bacteria by the epithelial cells themselves. Non-professional phagocytes, such as colonocytes, do not possess the receptors necessary for opsonic phagocytosis, instead bacterial pathogens trigger their endocytosis by reorganising the host cytoskeleton (Günther and Seyfert, 2018). Electron micrographs appeared to show *B. fragilis*, *E. coli* Nissle and *E. faecalis* cells being encapsulated by pits on the tumour cell surface. However, single *F. nucleatum* cells appeared to span multiple tumour cells and only partially penetrate the cell surface. This is likely attributable to the unusually large size of *F. nucleatum* cells, which are on average 5-10 µm long, but were observed at lengths up to 30 µm in this study. This perhaps necessitates a different mode of entry into host cells compared to the other test species which are 1-2 µm in diameter. It could be hypothesised that *F. nucleatum* cells only penetrate the intercellular space; however, intracellular *F. nucleatum* have been demonstrated in numerous previous studies, and it is thought that reorganisation of the tumour cell cytoskeleton is key it its invasion (Gursoy, Könönen and Uitto, 2008; Ji *et al.*, 2010).

Bacterial pathogens utilise multiple pathways in order to invade host cells. For example, *F. nucleatum* has been suggested to exclusively invade epithelial cells via the "zipping" mechanism, whereby close contact between bacteria and epithelial cells leads to interactions between bacterial surface proteins, with host proteins triggering cytoskeletal rearrangements (Han *et al.*, 2000). In contrast, species such as *E. faecalis* are able to employ multiple mechanisms to invade epithelial cells including receptor-mediated endocytosis and the induction of membrane ruffling (Bertuccini *et al.*, 2002). Differences in cell architecture and cell surface receptors between benign and malignant cells are likely to impact bacterial virulence. For example, attachment and invasion of both *B. fragilis* and *F. nucleatum* is thought to be mediated by binding to

E-cadherin (Rubinstein *et al.*, 2013; Devaux, Mezouar and Mege, 2019). Expression of E-cadherin is downregulated during CRC progression (Khoursheed *et al.*, 2003); therefore, it would be expected that these species show a greater capacity to invade benign RG/C2 cells than malignant HCT116 cells due to their higher levels of E-cadherin (Qualtrough 2021, personal communication). The gentamicin protection assay results presented in this project are contrary to this and suggest that other factors may play a pivotal role in influencing bacteria interactions with tumour cells.

In addition to intracellular invasion, the results of this study successfully demonstrated that the four species tested were able to survive, and remain viable, within tumour cells for a period of at least 24 hours. Persistence of *B. fragilis*, *E. faecalis* and *F. nucleatum* was significantly higher in HCT116 monolayers, whereas no significant difference was found in the persistence of *E. coli* Nissle between the benign and malignant tumour cells (Figure 3.11). Persistence of intracellular bacteria at high levels may be specific to tumour/epithelial cell type. A study of *F. nucleatum* degradation in gingival epithelial cells found that bacterial cells are rapidly internalised within endosome and lysosome structures before being degraded (Ji *et al.*, 2010). Genes encoding lysosomal proteins such as cathepsin B are frequently dysregulated in CRC, and therefore bacteria may benefit from improper endosome formation (Campo *et al.*, 1994). Furthermore, increased microbial tumour load is not specific to CRC, and Nejman *et al.* (2020) found that cancers including ovarian, breast, pancreas and brain all harbour unique microbiotas, with bacteria found at significantly higher abundancies than in paired normal tissues (Nejman *et al.*, 2020). It is therefore likely that the tumour cell intracellular environment is more conducive to bacterial survival than that of normal tissue; however the exact mechanism of bacterial survival is yet to be elucidated.

None of the bacteria analysed are considered to be intracellular pathogens. It is therefore not surprising that, with the exception of *B. fragilis* in HCT116 monolayers, no increase in intracellular bacterial cfu representative of bacterial replication was observed. In all species, bacterial survival decreased after 96 hours of incubation, suggesting that the bacteria cannot survive indefinitely within the tumour cells in this model. It has previously been reported that secondary tumours arising from CRC harbour unique microbiomes consistent with those found within the primary tumour (Bullman *et al.*, 2017). Thus, it is likely that bacteria are able to survive within tumour cells for much longer periods *in vivo*. This limitation is possibly a result of unidentified selection pressures present in the culture conditions. Furthermore, penicillin-streptomycin was added to the culture medium to prevent outgrowth of facultative anaerobes released from tumour cells from obfuscating counts of intracellular bacteria. Therefore, bacterial cells which persisted within tumour cells and subsequently translocated to the media (either through active movement or tumour cell lysis) would not have been counted in this assay. A further limitation is the inability to detect intracellular bacteria which are viable but non-culturable, a phenotype adopted by bacteria facing adverse environmental conditions that has been found in *E. coli* and other pathogenic species (Zhang *et al.*, 2018). To be absolutely certain of the absence of bacteria after prolonged time periods in this model, more sensitive non-culture-based techniques such as 16S rRNA sequencing or fluorescence *in situ* hybridisation should be used. However, the confirmation that bacteria persist for a minimum of 24 hours in this model highlights its suitability as a means of studying the effects of bacterial infection on tumour cell behaviour in downstream assays. To the author's knowledge, these data represent the first

demonstration of *in vitro* intracellular persistence within colon tumour cells for these species.

In order to progress this model, the addition of factors which better imitate *in vivo* conditions is needed. Colonic mucin is a key regulator of the intestinal microbiota, providing both a habitat and nutrient source for microbial colonisation (Schroeder, 2019). The formation of biofilms in the mucus layer is highly relevant to colorectal tumorigenesis, with mucosal biofilms observed almost universally in patients with proximal CRC (Kinzler *et al.*, 2014). Furthermore, mucosal biofilms from both healthy and CRC hosts were able to induce inflammation and tumour formation when transplanted into healthy mice (Tomkovich *et al.*, 2019). This suggests that the phenotype of biofilm-resident bacteria is more tumorigenic than that of non-biofilm growing cells, and that the formation of biofilms on normal colon tissue is a precursor to carcinogenesis. In addition, a key difference between the RG/C2 and HCT116 cell lines is the formation of mucin droplets by RG/C2, and so the RG/C2 cell line better represents the mucosal barrier observed *in vivo* (Paraskeva *et al.*, 1989) To determine if the presence of mucin glycoproteins could alter bacterial phenotype, the ability of bacteria to form biofilms with and without mucin was determined. The results of this assay demonstrate that the presence of mucin promotes biofilm formation in three of the species tested. Although this was not the case for *E. faecalis*, mucin had no detrimental effect on its biofilm-forming capabilities and other unobserved phenotypic changes relevant to CRC may have been present e.g. altered metabolite production. Therefore, the addition of mucin promotes a bacterial phenotype more representative of that seen *in vivo*; and should be considered for future work to more accurately investigate the role of these bacteria in CRC.

In conclusion, the results presented in this chapter clearly demonstrated the differences in the interactions of bacteria with benign and malignant tumour cells. These differences may be attributed to the differing morphology of RG/C2 and HCT116 cells as demonstrated by phase contrast microscopy and CLSM, although future investigation of tumour cell surface proteins would be required to fully explore this phenomenon. Despite high levels of bacterial attachment to cell surfaces, RG/C2 monolayers experienced lower bacterial invasion, and were less conducive to harbouring viable bacteria over prolonged time periods than HCT116 cells. This suggests that more advanced tumours may be more permissive to pathogenic bacteria *in vivo*. It is important to consider that these experiments were performed using 2D co-culture techniques, and therefore the impact of the colonic tumour microenvironment, the immune system and other microorganisms has not been addressed. Furthermore, while the specific microbiotas of CMS1, CMS2 and CMS3 tumours have been established, there is less data available concerning microbial composition of CMS4 tumour, to which the HCT116 cell line belongs (Purcell *et al.*, 2017). Despite these limitations, clear differences in bacterial interaction between tumour cell lines representing benign adenomas and malignant carcinomas were observed. The effects of these bacterial interactions and their implications for tumour progression will be explored in Chapter 4 and Chapter 5.

Chapter 4: Bacterial infection increases tumour cell yield and induces metabolic changes in colorectal tumour cells

4.1. Introduction

Of the six hallmarks of cancer proposed by Hanahan and Weinburg (2000), four are concerned with sustaining cell growth and avoiding cell death. In 2011 these hallmarks were expanded to include both immune evasion and, notably, deregulation of cellular metabolism (Hanahan and Weinberg, 2011). Increased cellular proliferation and inhibition of apoptosis, in addition to the metabolic reprogramming which facilitates these changes, are therefore essential to tumour formation. In colorectal carcinogenesis, reduced apoptosis at the mucosal surface is a key early event (Moss *et al.*, 1996). Furthermore, an increased rate of proliferation facilitates the acquisition of pro-oncogenic mutations, which are then selected for through successive cell divisions. From a clinical perspective, tumour size has a direct impact on patient prognosis, with larger tumours being associated with an increased frequency of lymph node involvement and higher mortality (Mejri *et al.*, 2017).

The tumour microenvironment (TME) plays a vital role in cancer progression, with microorganisms being a fundamental component of this TME in CRC. Numerous species of bacteria have demonstrated CRC-potentiating properties, and APC^{min} mouse models have demonstrated that tumour-associated microorganisms are able to induce cancer within disease-free colonic tissue (Kostic *et al.*, 2013). Many of the previously described mechanisms for bacterial CRC promotion involve remodelling of the microbiome, inflammatory signalling and other indirect pathways to tumour formation (Arthur *et al.*, 2014; Wong *et al.*, 2017). However, imaging of tumour biopsy samples has revealed that tumour cells are in direct contact with resident

microorganisms, with many tumours harbouring intracellular bacteria (Tomkovich *et al.*, 2019). The results presented in Chapter 3 demonstrate that these interactions may differ between benign and malignant tumour cells. It is therefore possible that this direct contact with CRC-associated microorganisms may influence the behaviour of benign and malignant tumour cells differently.

Tumour cell growth is intrinsically linked to cellular metabolism, particularly in CRC (having been extensively reviewed by Cairns, Harris and Mak, 2011; Hagland and Søreide, 2015). The Warburg effect describes the phenomenon whereby cancer cells preferentially utilise aerobic glycolysis over oxidative phosphorylation (Warburg, Wind and Negelein, 1927). This provides cells with increased quantities of small molecules for cell growth (e.g. NADPH), despite the apparent disadvantage to ATP production. This mechanism is particularly relevant to colonocytes, as this metabolic switch leads to the accumulation of butyrate in the cytosol, leading to its translocation to the nucleus and facilitating its effects as an HDACi (Donohoe *et al.*, 2012). It is now known that metabolic reprogramming occurs early on in the adenoma-carcinoma sequence, and is therefore a key driver of carcinogenesis (Satoh *et al.*, 2017).

The altered metabolism seen in tumours can be exploited for cancer diagnosis and treatment. For example, positron emission tomography scanning tracks fluorescently tagged glucose molecules to regions of high metabolism and is able to diagnose CRC in asymptomatic individuals with high sensitivity (Lin *et al.,* 2011). More recently, technologies such as gas chromatography mass spectrometry (GC/MS) and selected ion-flow tube mass spectrometry (SIFT-MS) have demonstrated that tumours display distinct metabolic profiles, which in the case of CRC can be detected from faecal samples (Batt*y et al.*, 2015). These technologies, which detect volatile

organic compounds present in the headspace above samples, have also shown great promise in identifying specific microorganisms present within clinical samples (Slade *et al.*, 2017). Given the association between shifts in microbiota composition and CRC development, identification of tumour-associated microorganisms through their metabolic profiles could potentially be exploited to provide novel biomarkers to aid in determining CRC prognosis and suggest targeted adjuvant antibiotic therapies. This chapter will explore the impact of *B. fragilis*, *E. coli* Nissle, *E. faecalis* and *F. nucleatum* on tumour cell growth and metabolism, with a focus on any differences between benign and malignant cells.

4.2. Aims

The aim of this chapter was to determine the effects of bacterial invasion on tumour cell growth. In addition, this chapter also aims to investigate the metabolic differences between colorectal adenoma and adenocarcinoma cells. Once established, changes in these metabolic profiles upon bacterial infection will be investigated to determine if bacteria are capable of triggering metabolic shifts which potentiate tumour growth.

4.3. Objectives

- To investigate how tumour cell yield is impacted after direct treatment with bacteria. Apoptosis and changes to cell cycle will also be investigated to attempt to explain any observed changes to cell yield.

- To investigate the Warburg effect in colorectal adenoma and adenocarcinoma cells by quantifying rates of oxidative phosphorylation and glycolysis. The ability of bacterial infection to alter this metabolic phenotype will also be investigated.

- To examine the volatile profiles of adenoma and carcinoma cell lines, in addition to determining how bacterial invasion alters cellular metabolism, and whether infected tumour cells display species-specific metabolic profiles.

4.4 Results

4.4.1 Bacteria significantly increase cell yield in RG/C2 and HCT116 tumour cells

In the previous chapter, it was established that bacteria invade RG/C2 and HCT116 cells and persist within the intracellular space without significantly decreasing cell viability (see Section 3.4.2). To investigate whether this bacterial invasion altered cell growth behaviour, tumour cell yield was quantified by counting viable cells 96 hours post-invasion. Bacterial MOIs of 10:1 and 100:1 were selected to eliminate the effect of altered cell viability observed in tumour cells infected at an MOI of 1000:1 (Figure 3.3). The average yield of untreated RG/C2 cells, initially seeded at 1.6×10^5 cells/well in 12-well plates, was approximately 2.8×10^5 cells/well at 96 hours (Figure 4.1). In contrast, the cell yield of untreated HCT116 cells was considerably higher (7.8×10^5 cells/well) despite being seeded at the lower density of 8×10^4 cells/well.

Despite the relative differences in cell number between untreated RG/C2 and HCT116 wells, bacterial invasion by all four species tested significantly increased cell yield in both cell lines (Figure 4.1). MOIs of 10:1 and 100:1 were tested for each species, and under each experimental condition an increase in cell yield of at least 1.5 relative to untreated cells was observed. These data demonstrate that an MOI as low as 10:1 is sufficient to alter tumour cell behaviour; however, increasing bacterial concentration did not significantly impact the effect on cell yield in either cell line. Across all species and MOIs, the average cell yield in infected RG/C2 cells was 8.3×10^5 cells/well, a value greater than the cell yield observed in untreated HCT116 cells. This suggests that under the conditions of this experiment, bacterial invasion can promote a growth phenotype in benign tumour cells similar to that of malignant cancer cells.
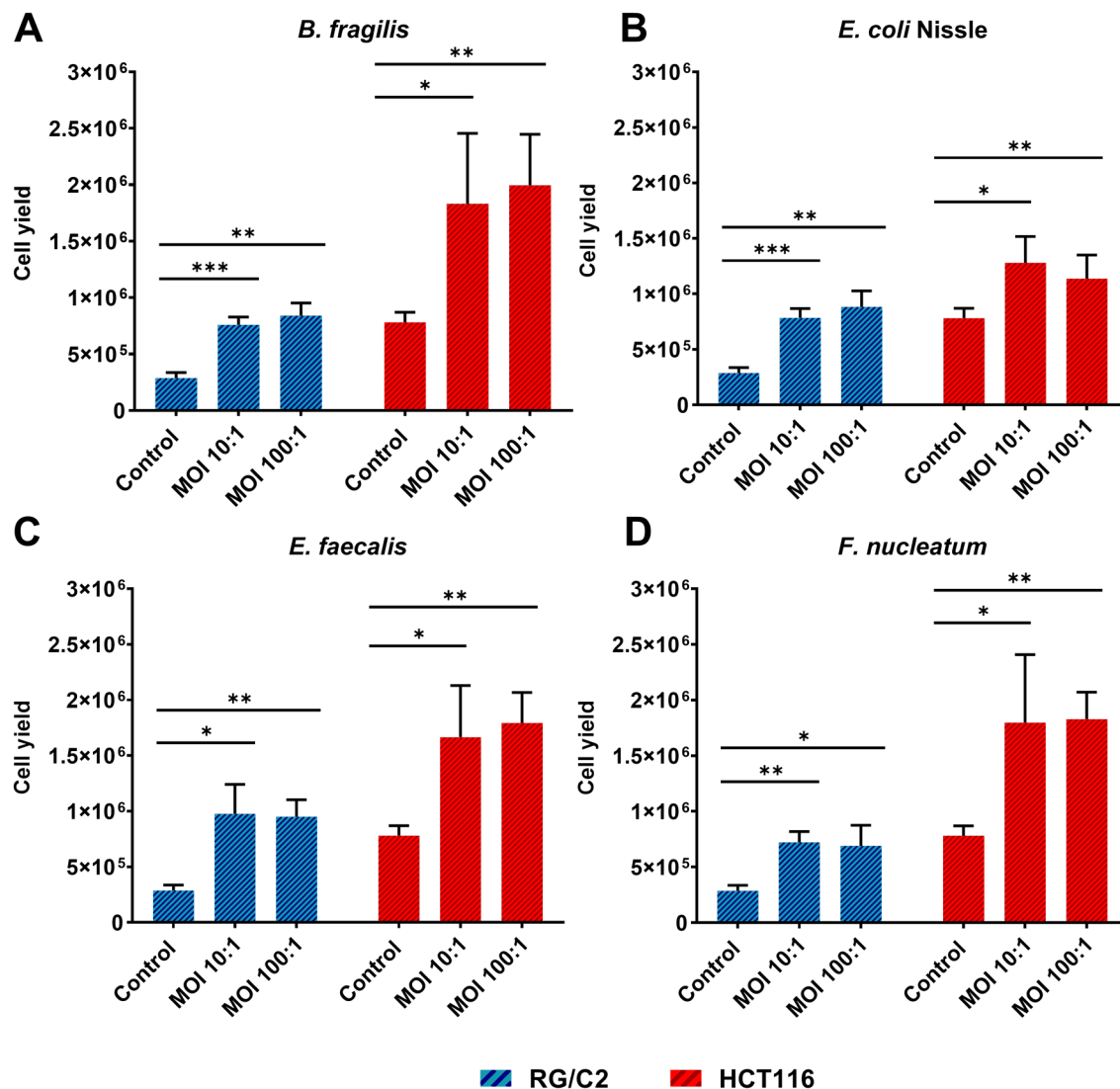
Figure 4.1: Bacterial invasion results in increased tumour cell yield in adenoma and adenocarcinoma cells. RG/C2 and HCT116 cells were seeded in 12-well plates and cultured for 72 hours before being infected with *B. fragilis* (A), *E. coli* Nissle (B), *E. faecalis* (C) or *F. nucleatum* (D) for four hours. Cell yield was measured using viable cell counts after 96 hours of incubation under standard conditions and demonstrated that bacterial infection significantly increases cell yield in all cases. Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate. T-test, $* = p < 0.05$, $** = p < 0.01$, $*** = p < 0.001$.

4.4.2 RG/C2 cells infected with colonic bacteria display lower rates of apoptosis

In order to determine whether the observed increases in cell yield could be attributed to a reduction in cell death, changes in the apoptotic cell number between untreated and infected tumour cells were measured. As previously described, apoptotic colorectal tumour cells detach from the culture surface as they undergo apoptosis (Hagu*e et al*., 1993). As such, floating cells collected from the aspirated cell culture medium before trypsinisation can be counted, and by expressing the number of floating cells as a proportion of the total cell count (attached plus floating) be used to quantify relative levels of apoptosis in a cell population. All floating cell counts were performed in parallel with the viable cell counts described in section 4.4.1.

In untreated RG/C2 cultures, apoptotic cells accounted for an average of 18.8% of the total cell count. This was significantly higher than HCT116 cultures, where the basal level of apoptosis was 0.52% ($p < 0.01$). This was visually apparent due to the large number of floating cells and other debris seen in RG/C2 cultures under the microscope. In most cases, bacterial infection lowered the proportion of apoptotic cells. Infection by *B. fragilis* reduced apoptosis at an MOI of 10:1 and 100:1 in RG/C2 and HCT116 cells; however, the reduction of apoptosis in RG/C2 cells treated at an MOI of 10:1 was not statistically significantly. RG/C2 cells treated with *E. coli* Nissle, *E. faecalis* and *F. nucleatum* at MOIs of 10:1 and 100:1 exhibited a significant reduction in apoptosis. The largest decrease in apoptosis was observed in *E. faecalis*-treated cells, which produced a decrease of 60% and 64% at MOIs of 10:1 and 100:1, respectively (Figure 4.2C). This is consistent with the observed effects on cell yield, where *E. faecalis* induced the largest increase in RG/C2 cell count (Figure 4.1C).

In contrast, bacterial invasion did not have a consistent effect on floating cell number in HCT116 cultures. A significant reduction in apoptosis was only observed in HCT116 cells infected with *B. fragilis* and with *E. faecalis* at an MOI of 100:1 ($p < 0.01$). However, non-significant reductions were seen in cells infected with *E. coli* Nissle (MOI 10:1 and 100:1), *E. faecalis* (MOI 10:1) and *F. nucleatum* (MOI 10:1). Surprisingly, apoptosis appeared to increase in HCT116 cells infected with *F. nucleatum* at an MOI of 100:1, although this increase was not statistically significant ($p = 0.09$, Figure 4.2D). The lack of statistical significance observed could perhaps be attributed to the relatively low numbers of apoptotic cells counted (less than 1% of the total cell count).
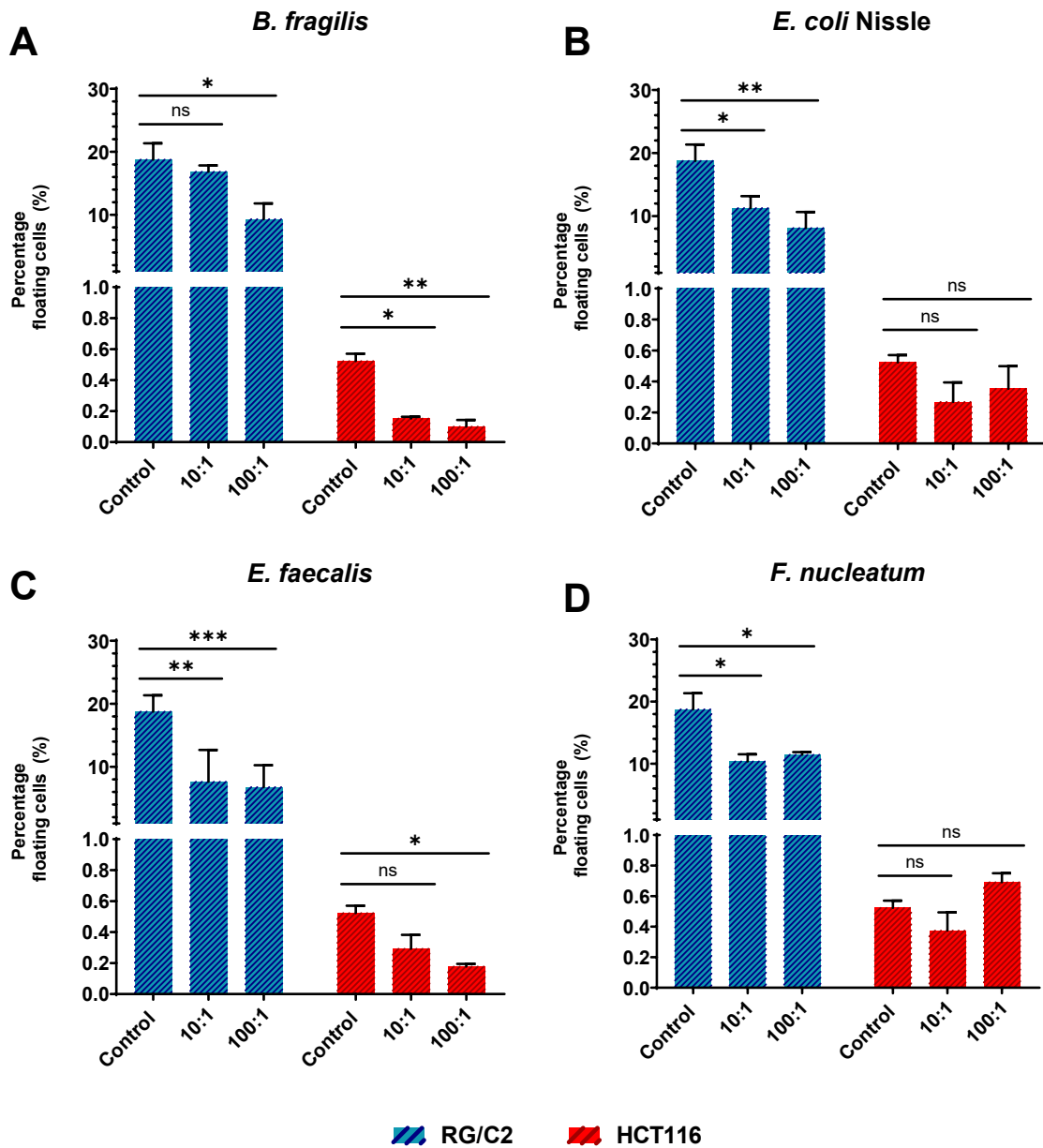
Figure 4.2: Floating cell counts in RG/C2 and HCT116 cultures 96 hours post-infection. Floating cell counts were performed in parallel with cell yield experiments. Floating cells were collected 96 hours post-infection and counted using a haemocytometer. Values are expressed as a percentage of the total cell population. Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate. T-test, $* = p < 0.05$, $** = p < 0.01$, $*** = p < 0.001$.

4.4.3 The relative impact of bacterial infection on tumour cell yield and apoptosis is consistent across RG/C2 and HCT116 cells

Due to the natural disparities in growth rate and viability of RG/C2 and HCT116 cells *in vitro*, it is beneficial to compare the fold change in cell yield and apoptosis percentage caused by bacterial infection. Towards the end of the 96-hour incubation period, bacteria-treated HCT116 cultures were beginning to reach confluency, which may have obscured the overall impact of infection on cell yield. However, Figure 4.3A demonstrates that despite the obvious difference in cell numbers seen in Figure 4.1, bacterial infection induced a minimum 1.5-fold increase in cell yield across all experimental parameters. A Two-way ANOVA was used to determine whether the differences in the fold increase in cell yield between RG/C2 and HCT116 cultures were significant. This revealed that the cell line did not statistically impact the degree of cell yield increase for each species ($p = 0.79$), highlighting that both RG/C2 and HCT116 cells exhibit similar growth responses to bacterial infection.

With the exception of HCT116 cells infected with *F. nucleatum* at an MOI of 100:1, bacterial infection reduced apoptosis in all cases with varying levels of significance. *B. fragilis* infection caused a greater reduction in apoptosis in HCT116 cells (-0.29), with the fold decrease at an MOI of 10:1 being significantly higher than that of RG/C2 cells ($p = 0.03$; Figure 4.3B). In cells infected with either *E. coli* Nissle or *E. faecalis*, the fold change in apoptosis was approximately equivalent for both RG/C2 and HCT116 cells. For example, RG/C2 apoptosis was reduced by a factor of 0.35, compared to a reduction of 0.34 in HCT116 cells. The greatest divergence was observed when cells were infected with *F. nucleatum*. Despite an MOI of 10:1 reducing apoptosis in both cell lines, an MOI of 100:1 significantly increased apoptosis in HCT116 cells, despite reducing apoptosis in RG/C2 cultures. Regardless

of the direction of change, Figure 4.3 demonstrates that bacterial infection clearly

influences tumour cell behaviour in a manner which is largely consistent between

benign and malignant cells.

A

B

RG/C2   HCT116
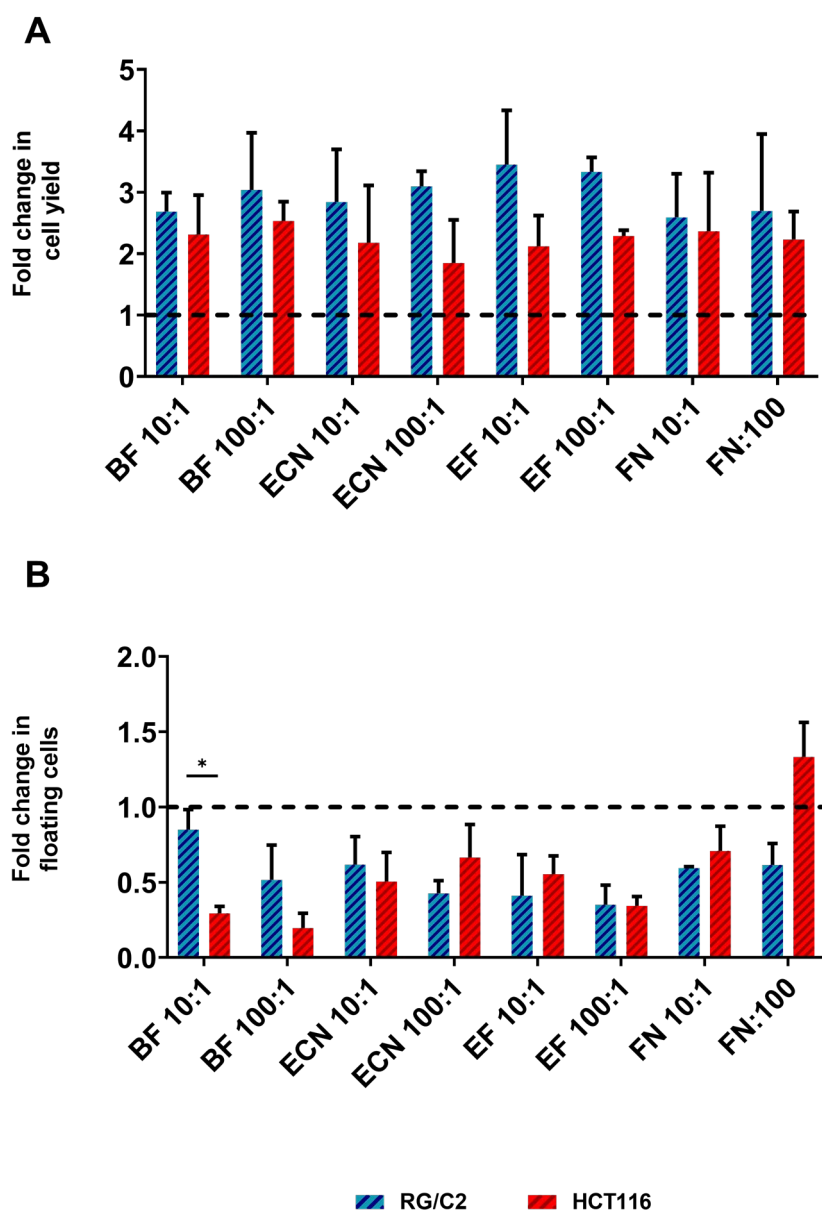
Figure 4.3: Comparison of the fold change in cell yield and floating cell proportion between RG/C2 and HCT116 cells infected with bacteria. Values for cell yield (A) and floating cell percentage (B) were normalised to the control group (untreated cells) to allow for better comparison between cell lines. BF = *B. fragilis*, EcN = *E. coli* Nissle, EF = *E. faecalis*, FN = *F. nucleatum*. Numbers on the x-axis indicate the multiplicity of infection. Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate. Two-way ANOVA $* = p < 0.05$.

4.4.4 Bacterial infection does not promote proliferation in tumour cells *in vitro*

Figure 4.1 demonstrates that bacterial infection with the four species investigated significantly increases cell yield regardless of cell type. To determine whether this was due to changes in the cell cycle, analysis of DNA content was performed via flow cytometry. Cells were fixed 24 hours post-infection and stained with propidium iodide. After staining, DNA content was quantified and the proportion of cells within G1-phase, S-phase and G2-phase was modelled. In RG/C2 cells infected with either *B. fragilis* or *E. coli* Nissle, a small non-statistically significant increase in the proportion of cells in S-phase was observed (Figure 4.4A). *E. faecalis* and *F. nucleatum* infection decreased the proportion of cells in S-phase, but this was only statistically significant in *E. faecalis*-infected cells ($p = 0.0219$). The proportion of RG/C2 cells in G2-phase was significantly increased in cells infected with *E. faecalis* ($p<0.001$) and *F. nucleatum* ($p < 0.001$). There was also a small, non-statistically significant increase in G2-phase in RG/C2 cells infected with *B. fragilis* ($p = 0.33$). No difference in G2-phase was observed in *E. coli* Nissle infected cells. In HCT116 cells the percentage of cells in G1-phase was significantly increased regardless of the bacterial species tested (Figure 4.5A). The largest increase was seen in cells infected with *F. nucleatum*, with a 15% rise in G1-phase cells. In contrast, the percentage of RG/C2 cells in G1-phase was decreased upon bacterial infection (Figure 4.5). This decrease was significant in cells infected with *B. fragilis* ($p=0.0387$), *E. faecalis* ($p = 0.0219$) and *F. nucleatum* ($p = 0.0068$), although not in cells infected with *E. coli* Nissle. Infection with all species decreased the proportion of HCT116 cells in both S-phase and G2-phase, however these decreases were not statistically significant.
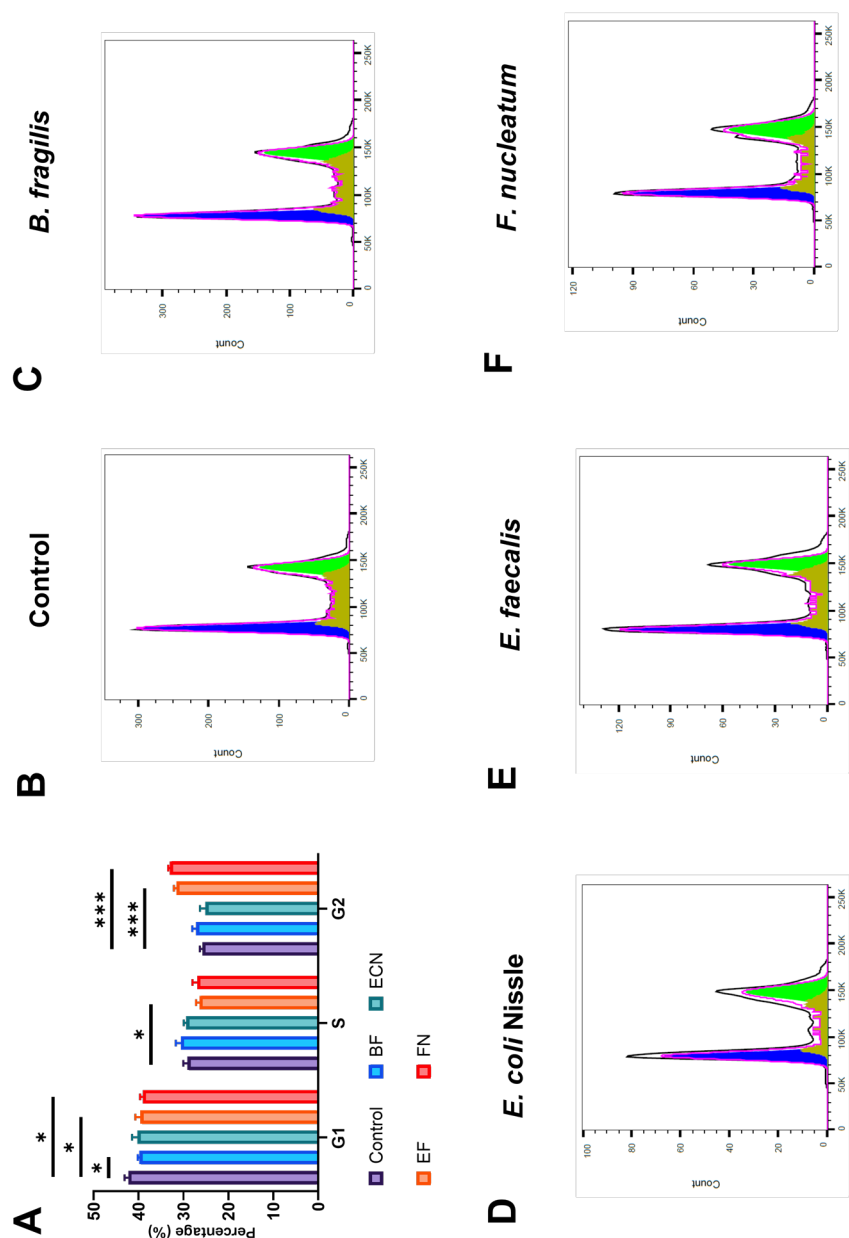
Figure 4.4: Cell cycle analysis of infected RG/C2 cells. Samples of RG/C2 cells infected at an MOI of 10:1 were collected 24 hours post-infection. DNA content was analysed via propridium iodide staining and flow cytometry. Approximately 20,000 events were recorded per sample, and events were gated to only include singlet cells and exclude cellular debris. (A) The proportion of cells in each phase of the cell cycle. Representative histograms for untreated control cells (B) and cells infected with *B.fragilis* (C), *E. coli* Nissle (D), *E.faecalis* (E) and *F. nucleatum (F)*. Colours on histograms correspond to G1-phase (blue), S-phase (gold) and G2-phase (green). Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate. T-test $* = p < 0.05$, $** = p < 0.01$.
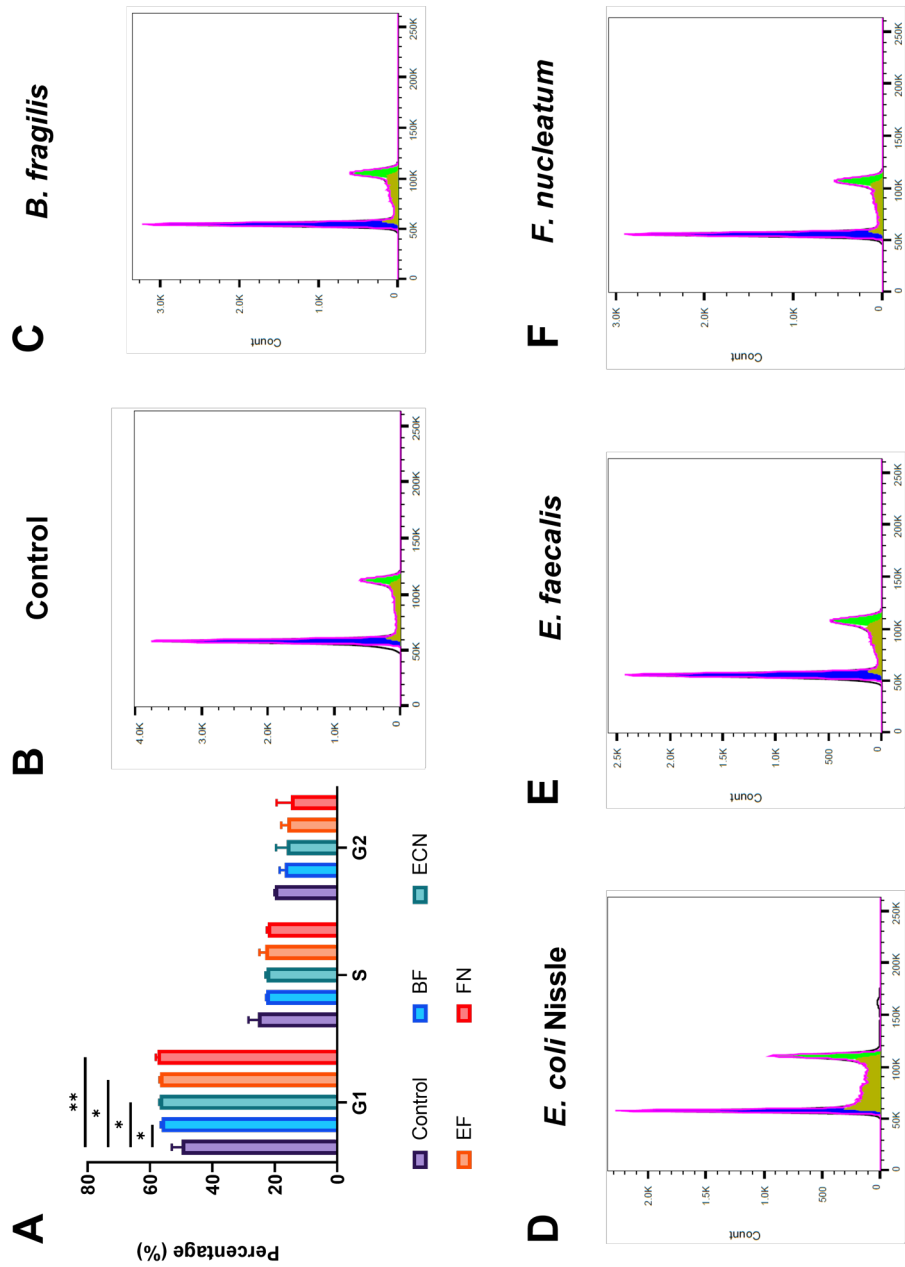
Figure 4.5: Cell cycle analysis of infected HCT116 cells. Samples of HCT116 cells infected at an MOI of 10:1 were collected 24 hours post-infection. DNA content was analysed via propridium iodide staining and flow cytometry. (A) The proportion of cells in each phase of the cell cycle. Representative histograms for gated to only include singlet cells and exclude cellular debris. (A) The proportion of cells in each phase of the cell cycle. Representative histograms for untreated control cells (B) and cells infected with *B.fragilis* (C), *E. coli* Nissle (D), *E.faecalis* (E) and *F. nucleatum* (F). Colours on histograms correspond to G1-phase (blue), S-phase (gold) and G2-phase (green). Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate. T-test * = $p < 0.05$, ** = $p < 0.01$.

4.4.5 Benign RG/C2 cells display a metabolic profile distinct from malignant adenocarcinoma cell lines

Cellular metabolism is an important limiting factor on the expansion of tumour cell populations. In CRC, as in many other cancers, the Warburg effect is thought to be a key characteristic which facilitates increased tumour growth and allows tumour cells to outcompete neighbouring cells for nutrients (Liberti and Locasale, 2016). However, there is some debate surrounding the point in the adenoma-carcinoma sequence at which the Warburg effect occurs. To investigate whether metabolic differences are apparent between adenoma and adenocarcinoma cell lines, cellular metabolism in RG/C2 cells and three adenocarcinoma cell lines, HCT116, HT29 and SW480 was analysed using the Seahorse Xf analyser.

Before commencing Seahorse experiments, the optimal seeding density for each cell line was determined (Figure 4.6). According to the manufacturer's recommendations, a seeding density should be chosen where a linear increase in cell number corresponds to a linear increase in the oxygen consumption rate (OCR). A range of seeding densities were tested for each cell line. Figure 4.6 demonstrates that for all cell lines, a linear relationship between cell number and OCR is present at low seeding densities. For example, increasing the seeding density of HT29 cells from 4×10^5 to 6×10^5 increased OCR from 195 pmol/min to 288 pmol/min, a fold increase of 1.47. This linear relationship was not present between RG/C2 and SW480 cells seeded at higher densities (Figure 4.6A and 4.6D). Therefore, lower seeding densities were selected for each cell line. The final selected seeding density for HCT116 cells was 3×10^4 cells/well. A density of 4×10^4 cells/well was selected for HT29 and SW480 cells, and 6×10^4 cells/well for RG/C2 cells. In addition, RG/C2 cells were incubated for a further 24 hours post-seeding to allow cells to fully recover before all assays.

Final cell numbers post-assay were determined using crystal violet staining and used to normalise results. The optimal concentration of FCCP was also determined by testing a range of concentrations between 0.5-2 μM (data not shown). The manufacturer's recommended concentration of 1 μM was sufficient to induce a two-fold increase in OCR for all cell lines and so was chosen for all future experiments.

The metabolic rates of the four cell lines were then analysed using the Seahorse Energy Phenotype test. OCR and extracellular acidification rate (ECAR) are quantified as indirect measures of oxidative phosphorylation and glycolysis, respectively. First, basal levels of metabolism are recorded, followed by maximal rates of glycolysis and oxidative phosphorylation which are stimulated by the real-time injection of Oligomycin and FCCP. Each cell line exhibited distinct rates of both basal and maximum oxygen consumption (Figure 4.7A and B, respectively). RG/C2 cells displayed the highest basal and maximal OCR, in addition to the highest potential for increased OCR (Figure 4.7D). OCR in SW480 cells was also significantly higher than HCT116 and HT29 cells, which were not statistically different from each other.

The basal ECAR of RG/C2 cells was lower than both HCT116 and SW480 cells (Figure 4.7D), although this difference was only statistically significant against HCT116 cells ($p = 0.0318$). Interestingly, the maximum ECAR of the four cell lines were not significantly different (Figure 4.7E), demonstrating that RG/C2 cells have the capacity when under stress to reach levels of glycolysis similar to that of the adenocarcinoma cells despite having a lower basal rate. This was reflected in the spare glycolytic capacity of RG/C2 cells, which was significantly higher than HCT116 ($p = <0.001$) and SW480 ($p = 0.0041$) cells (Figure 4.7F). Due to the similarity in basal ECAR between RG/C2 and HT29 cells, no statistically significant difference in their glycolytic capacities was observed.

Energy phenotype was visualised by dividing the OCR by the ECAR for each cell line (Figure 4.8A). The energy phenotypes of all four cell lines tested were statistically distinct, as assessed by a Two-way ANOVA. However, the benign adenoma cell line RG/C2 displayed the greatest ratio of oxidative phosphorylation to glycolysis. Figure 4.8B demonstrates that despite exhibiting significantly different metabolic profiles, the basal metabolic rates of the three adenocarcinoma cell lines form a distinct cluster away from that of the RG/C2 cell line. This difference becomes more pronounced when analysing maximum metabolic rates. The three adenocarcinoma cell lines shifted markedly towards a glycolytic phenotype when placed under stress. This is in contrast to RG/C2 cells (blue line) where the rate of increase of oxidative phosphorylation and glycolysis were approximately equal. This demonstrates that regardless of whether a Warburg effect occurs in the transition from normal epithelial to colorectal adenoma, further metabolic shifts also occur between adenoma and adenocarcinoma cells.
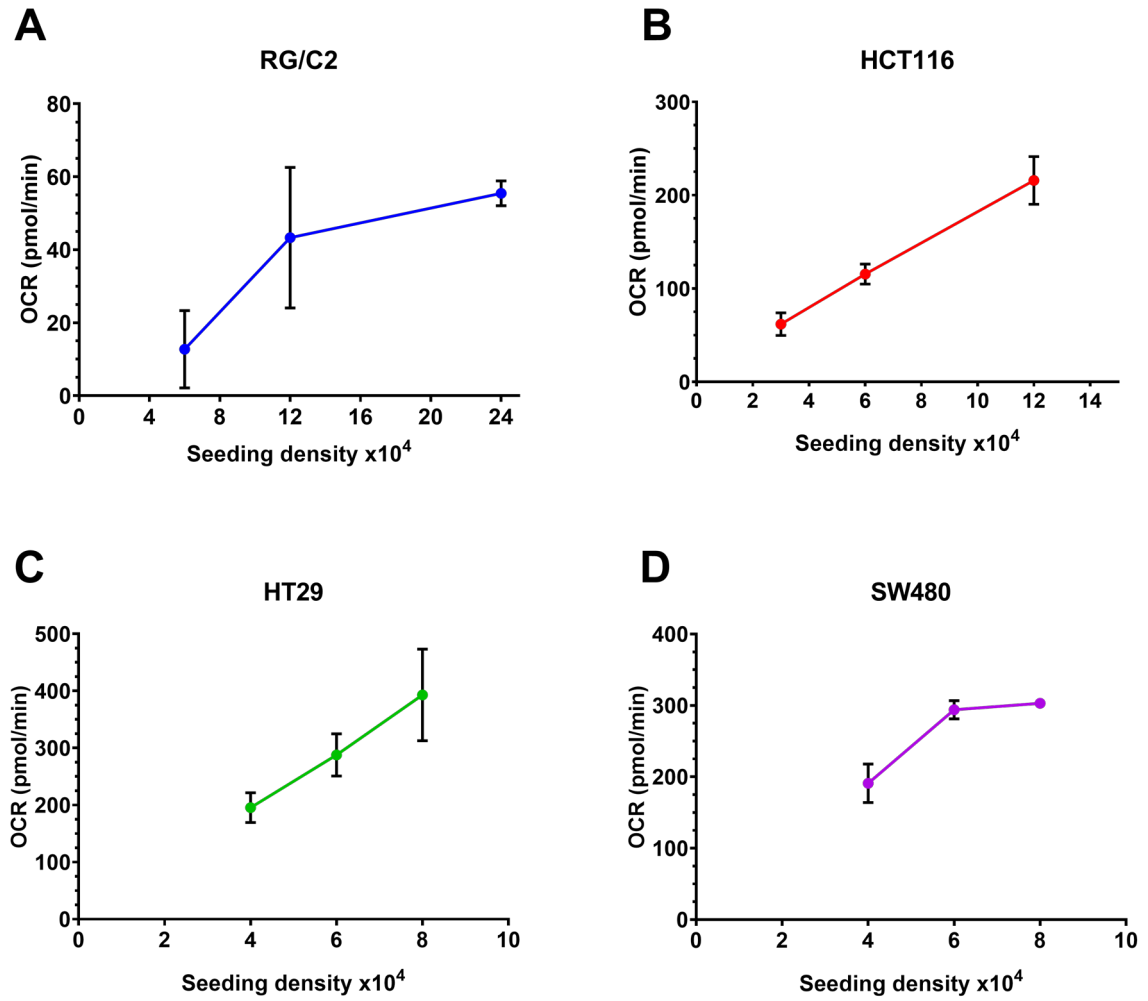
Figure 4.6: Seeding density optimisation for Seahorse assays using colorectal tumour cell lines. RG/C2 (A), HCT116 (B), HT29 (C) and SW480 (D) cells were seeded at various densities and oxygen consumption rate (OCR) was measured using the Seahorse Xf24 analyser. For each cell line, the lowest seeding density was chosen, as higher seeding densities for RG/C2 and SW480 cultures did not result in a linear increase in OCR. Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate.
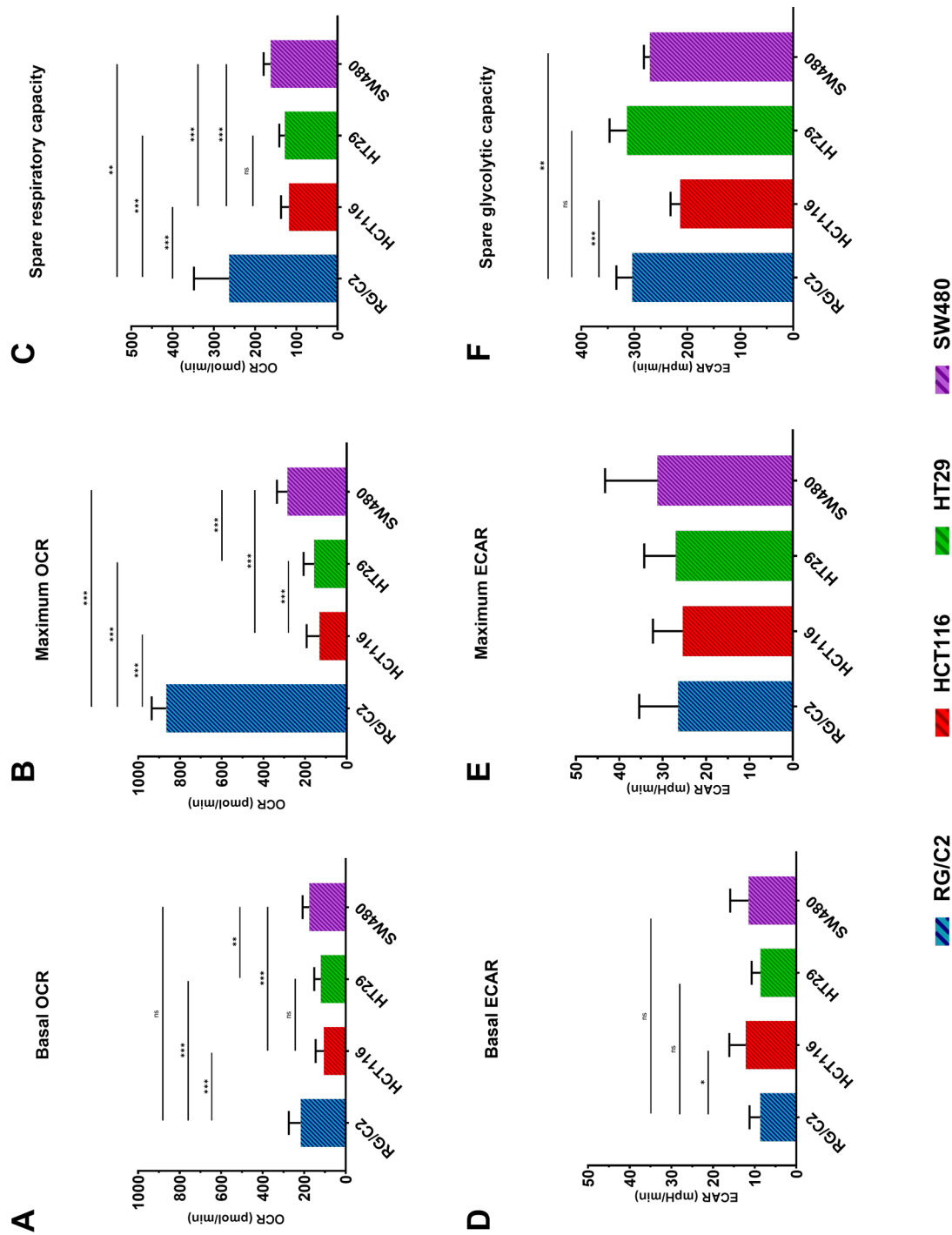
Figure 4.7: Rates of oxidative phosphorylation and glycolysis in colorectal tumour cells. Tumour cell metabolism was measured using the Seahorse Energy Phenotype Test. (A) Baseline oxygen consumption rate (OCR). (B) Maximum OCR measured after the injection of 1μM FCCP. (C) Spare respiratory capacity calculated as the difference between basal and maximum OCR. (D) Baseline extracellular acidification rate (ECAR) as a measure of glycolytic rate. (E) Maximum ECAR measured after the injection of 1μM Oligomycin. (D) Spare glycolytic capacity calculated as the difference between basal and maximum ECAR. All values were normalised to cell number using crystal violet staining. Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate. T-test $* = p < 0.05$, $** = p < 0.01$, $*** = p < 0.001$.

Figure 4.8. Energy phenotypes of colorectal tumour cells. Basal and stressed (post-injection of Oligomycin/FCCP) OCR and ECAR values were converted to an energy phenotype ratio. (A) Energy phenotype expressed as the ratio of oxygen consumption rate (OCR) to extracellular acidification rate (ECAR). A greater value represents a more oxygen-dependent metabolism. (B) Energy phenotype plot comparing OCR and ECAR to compare relative aerobic/glycolytic and energetic/quiescent axis between cell lines and showing difference between basal rates (full symbols) and maximum rates (hollow symbols). Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate. * = p < 0.05, ** = p < 0.01, *** = p < 0.001.

4.4.6 *F. nucleatum* increases glycolytic capacity in infected HCT116 cells

The ability of bacteria to modify cellular metabolism was investigated by performing the Energy Phenotype test on HCT116 tumour cells 24 hours post-infection. As the cell monolayers were incubated for an additional 24 hours, higher basal levels of OCR and ECAR were observed (Figure 4.9A and 4.9D) in comparison to the values seen in Figure 4.7. Bacterial infection did not significantly affect OCR for any of the species tested (Figure 4.9A&B). Small decreases in maximum OCR and spare respiratory capacity were witnessed, however these were not statistically significant. In contrast, infection of HCT116 cells by *B. fragilis* and *F. nucleatum* significantly increased basal ECAR (Figure 4.9D). Basal ECAR was also increased in response to *E. faecalis*, however this result was not statistically significant ($p = 0.1101$). Infection by *F. nucleatum* also significantly increased maximal ECAR ($p = 0.0283$) by approximately 30% (Figure 4.9E). As both *B. fragilis* and *F. nucleatum* infection caused the basal ECAR to increase, the spare glycolytic capacity was reduced in comparison to un-infected cells, suggesting that infection by bacteria caused cells to perform glycolysis closer to their maximum rate. When comparing energy phenotypes, both *B. fragilis* and *F. nucleatum*-treated cells displayed a lower ratio of OCR:ECAR (Figure 4.10A), and Figure 4.10B demonstrates that infection by *F. nucleatum* causes HCT116 cells to shift to the right, exhibiting a greater capacity for glycolysis.
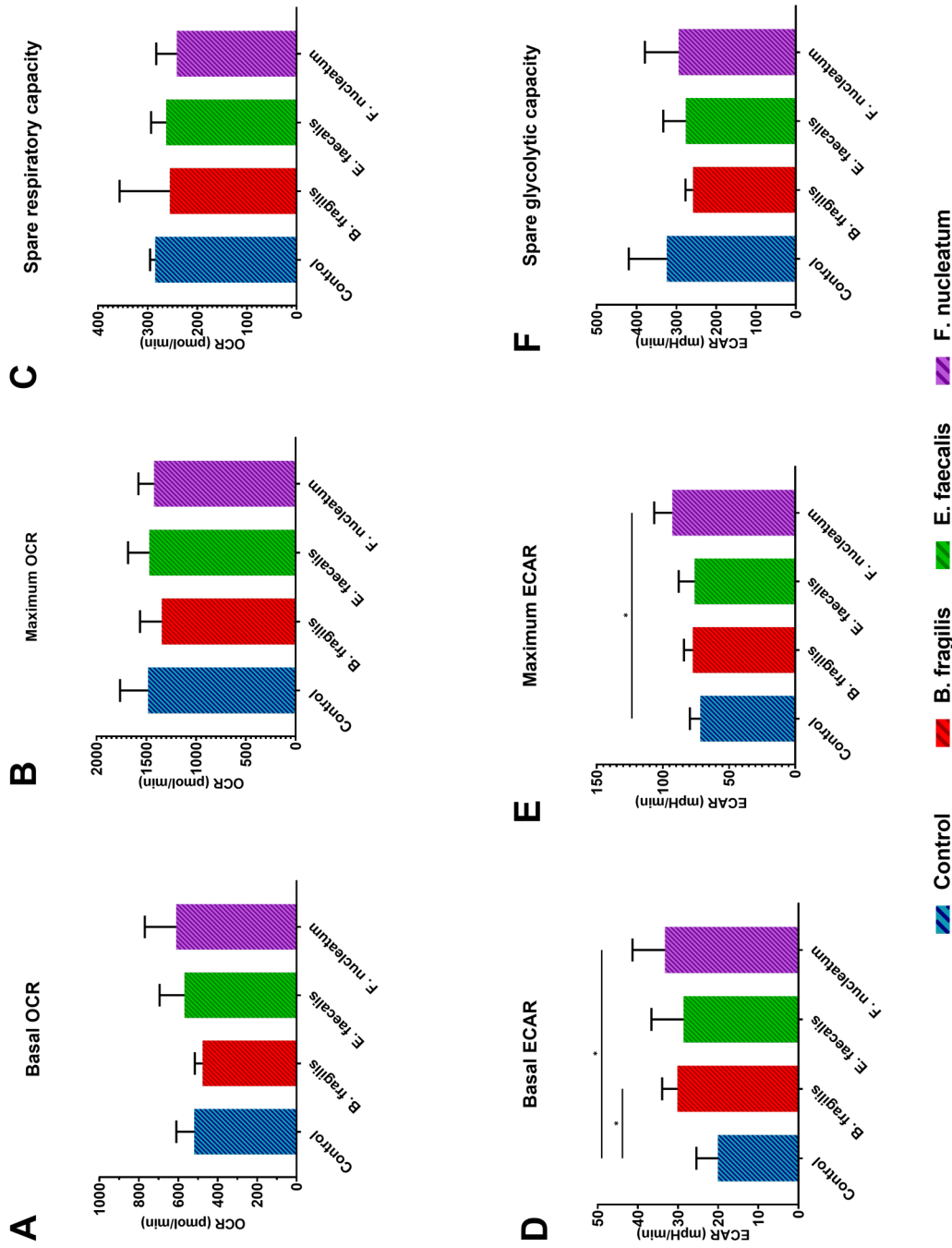
Figure 4.9. Rates of oxidative phosphorylation and glycolysis in HCT116 cells after bacterial infection. Seahorse assays were performed 24 hours post-bacterial infection Tumour cell metabolism was measured using the Seahorse Energy Phenotype Test. (A) Baseline oxygen consumption rate (OCR). (B) Maximum OCR measured after the injection of 1μM FCCP. (C) Spare respiratory capacity calculated as the difference between basal and maximum OCR. (D) Baseline extracellular acidification rate (ECAR) as a measure of glycolytic rate. (E) Maximum ECAR measured after the injection of 1μM Oligomycin. (D) Spare glycolytic capacity calculated as the difference between basal and maximum ECAR. All values were normalised to cell number using crystal violet staining. Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate. T-test $* = p < 0.05$.

Figure 4.10: The effect of bacterial infection on HCT116 energy phenotype. Seahorse assays were performed 24 hours post-infection and basal and stressed (post injection of Oligomycin/FCCP) OCR and ECAR values were converted to an energy phenotype ratio. (A) Energy phenotype expressed as the ratio of oxygen consumption rate (OCR) to extracellular acidification rate (ECAR). A greater value represents a more oxygen-dependent metabolism. (B) Energy phenotype plot of OCR and ECAR to compare relative aerobic/glycolytic and energetic/quiescent axis between cell lines and showing differences between basal rates (full symbols) and maximum rates (hollow symbols). Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate.

4.4.7 Discrimination between adenoma and adenocarcinoma cells using volatile product ion peak detection

Having shown that the metabolism of the benign colorectal tumour cell line RG/C2 differs from that of malignant cell lines, and that certain bacterial species can influence tumour metabolism, we then aimed to determine whether these changes are reflected in the volatile profiles of these tumour cells. The headspace above colorectal tumour cell cultures was analysed by SIFT-MS in full scan mode over a spectrum range of 15-200 m/z in order to detect the volatile product ion peaks produced by these cells in culture. Using the HCO_3^+ reagent ion data, a background subtraction was performed by comparing the volatile product ion peaks detected from the headspace above tumour cells with those detected in the headspace above cell culture media controls. The product ions peaks present in tumour cell samples that were significantly elevated ($p < 0.05$) compared to these controls are presented in Supplementary Table 3.

To further analyse the detected volatile profiles, hierarchical cluster analysis (HCA) was performed, and the resulting dendrogram is presented in Figure 4.11. Individual replicates of specific cell lines clustered together, demonstrating that each cell line possesses a reproducible distinct volatile metabolome. Furthermore, the three adenocarcinoma cell lines tested (Figure 4.11, red box) clustered together, with HT29 and SW480 showing the greatest degree of similarity. These three cell lines were distinctly different from the benign adenoma cell line RG/C2. The SIFT-MS data was also analysed via principal component analysis (PCA). Figure 4.12 illustrates a scatter plot of the first two principal components, which accounted for 22% and 18% of the total variation, respectively. Principle component 1 (X-axis) separated benign cells (Figure 4.12, blue circle) from malignant cells (Figure 4.12, red circle).

Adenocarcinoma samples were separated on the Y-axis by the second principal

component; however, individual replicates of each cell line did not cluster together.
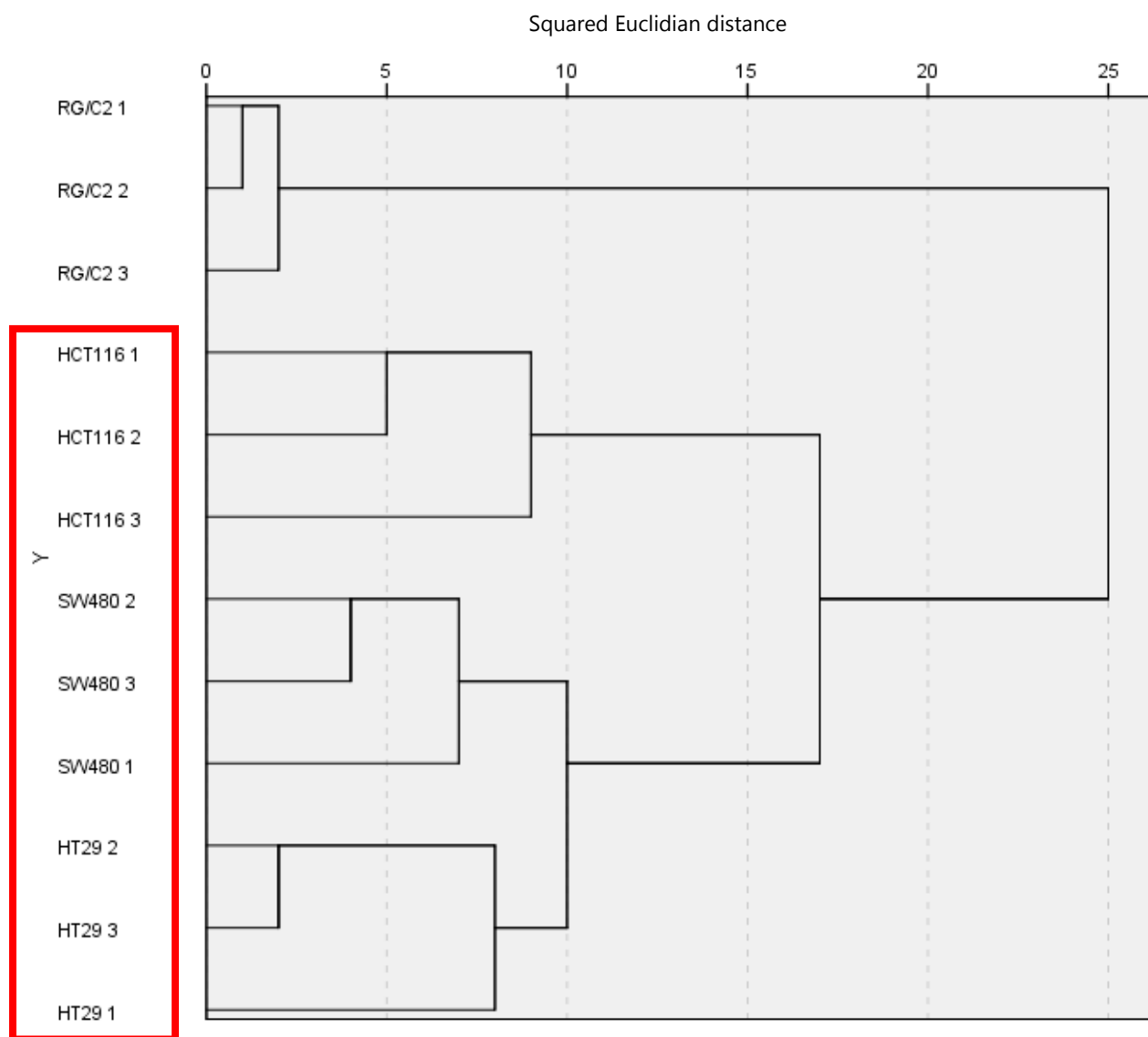
Figure 4.11: Dendrogram of volatile product ion peaks detected from cell culture headspace after hierarchical cluster analysis. HCA was performed on 82 volatile peaks detected above colorectal tumour cell cultures using SIFT-MS. Cell culture flasks were capped for 30 minutes prior to sampling. Red box designates colorectal adenocarcinoma cell lines, which clustered separately from the benign adenoma cell line RG/C2.
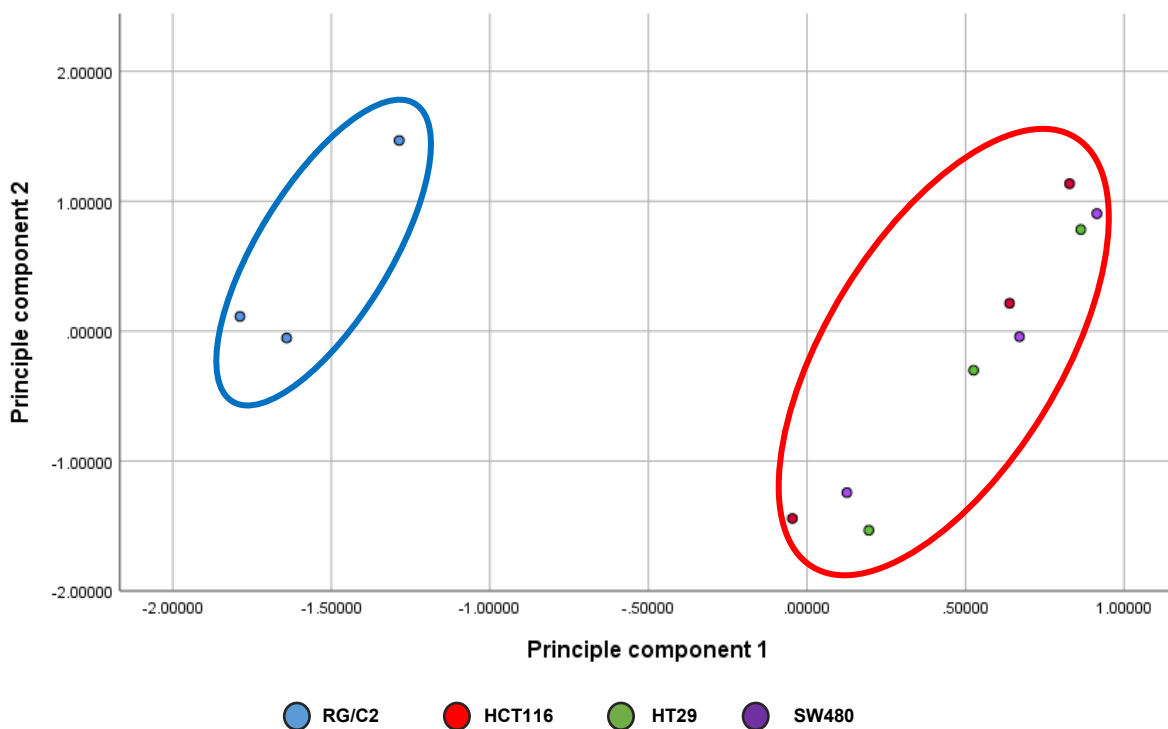
Figure 4.12: Principal component analysis of volatile product ion peaks detected from colorectal tumour cell cultures. Scatter plot produced using the first and second scores of the principal component analysis performed on 82 volatile peaks detected in the headspace above colorectal tumour cell cultures. Cell culture flasks were capped for 30 minutes before sampling. Benign RG/C2 cells clustered together (blue circle) away from malignant adenocarcinoma cell lines (red circle). However, biological replicates of adenocarcinoma cell lines did not cluster together.

4.4.8 Identification of intracellular bacteria using SIFT-MS

Analysis of tumour cell metabolism using the Seahorse Xf analyser revealed that metabolic shifts occur upon infection with specific species e.g. *F. nucleatum*. To further investigate these metabolic changes, SIFT-MS was used to analyse the headspace above bacteria-infected cells. For each species, a background subtraction was performed against non-infected cells, and the significant product ion peaks ($p <$ 0.05) for infected RG/C2 and HCT116 samples are presented in Supplementary Table 4 and 5, respectively. RG/C2 cells infected with *B. fragilis* displayed numerous peaks; however, all other species induced much smaller metabolic changes. For example, in RG/C2 cells infected with *F. nucleatum* only five significant peaks were detected across all replicates. Hierarchical cluster analyses revealed that samples of RG/C2 cells infected with *B. fragilis* displayed a high degree of similarity, with volatile profiles dissimilar to any other species tested (Figure 4.13). *F. nucleatum*-infected cells were also closely related, whereas *E. coli* Nissle and *E. faecalis*-infected cells could not be discriminated through HCA. This was attributed to a lack of significant peaks, as opposed to multiple shared peaks. Figure 4.13 displays these samples plotted against the first and second principal components following PCA; however, individual replicates of each species were not found to cluster together.

In contrast, HCT116 cells exhibited a significant metabolic shift upon infection, with numerous significant peaks detected which presented a similar profile to RG/C2 cells infected with *B. fragilis*. The majority of these peaks were significantly different from the control in cells infected with any of the four bacterial species tested, with notable large peaks detected at m/z 27 and m/z 50. HCA was also performed on this data set; however, individual replicates for each species were not found to cluster

together (Figure 4.14). Similarly, bacterial species could not be separated after PCA

(Figure 4.15).
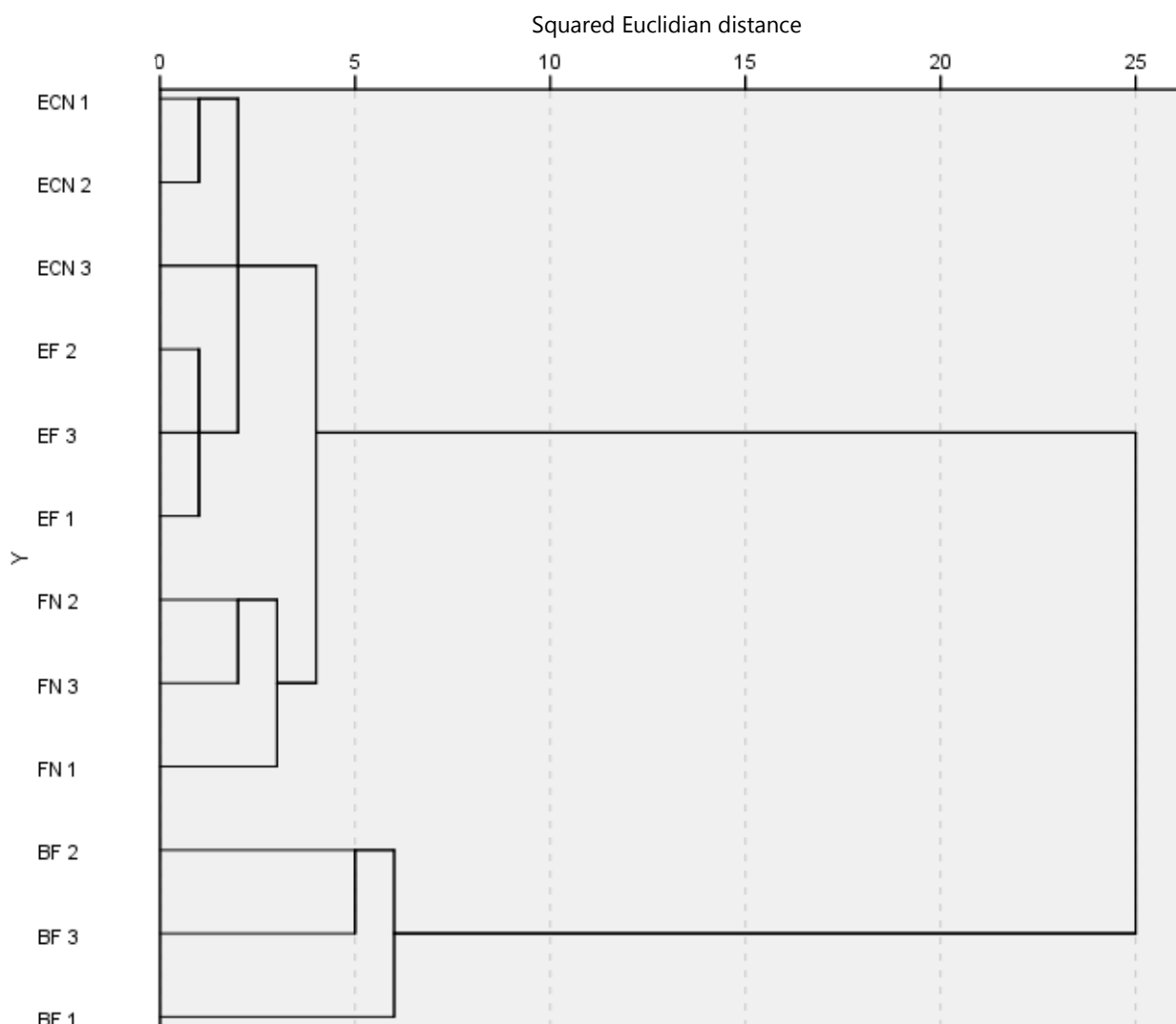
Figure 4.13: Dendrogram of volatile product ion peaks detected from the headspace above infected RG/C2 hierarchical cluster analysis. HCA was performed on 84 volatile peaks detected above infected RG/C2 cells. *B. fragilis* and *F. nucleatum*-infected cells clustered discretely, but no discrimination between *E. coli* Nissle and *E. faecalis*-infected cells could be achieved. Cell culture flasks were capped for 30 minutes prior to sampling.
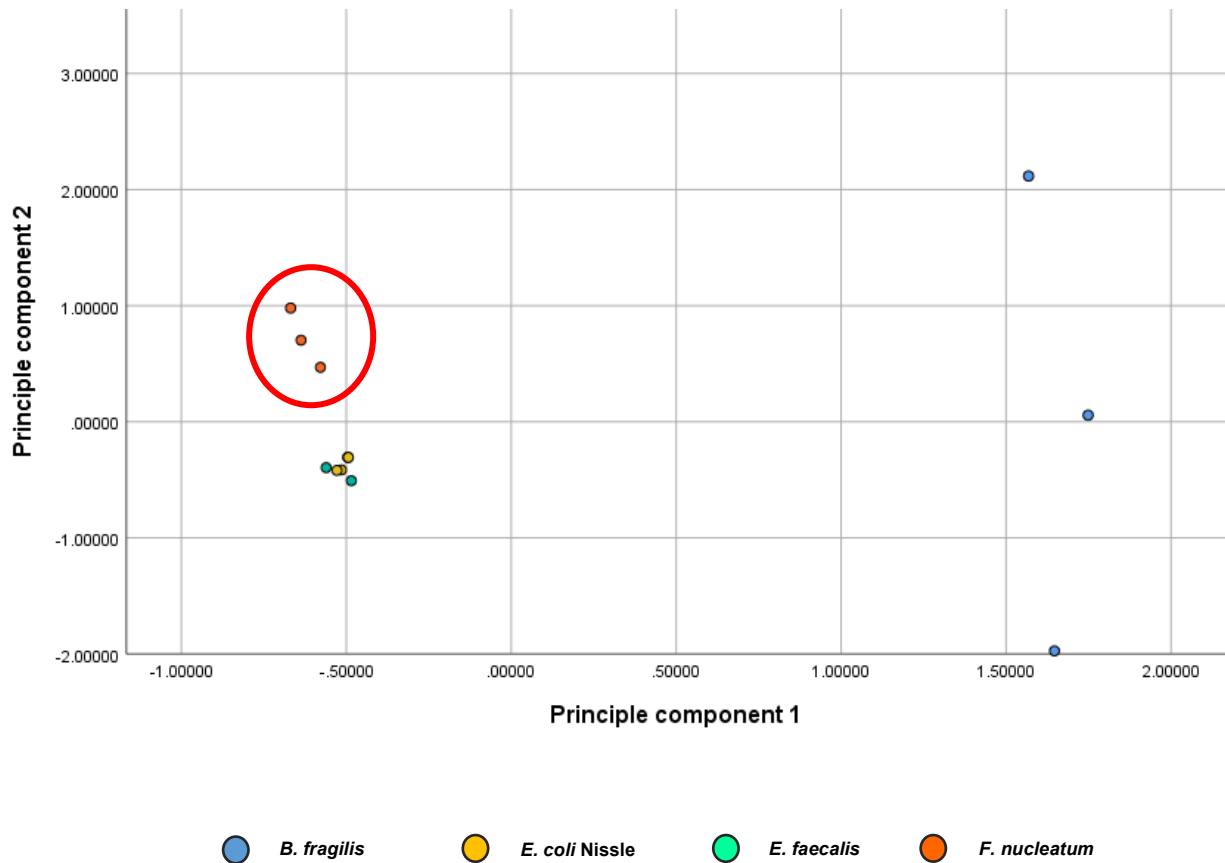
Figure 4.14: Principal component analysis of volatile product ion peaks detected from the headspace above infected RG/C2 cells. Scatter plot produced using the first and second scores of the principal component analysis performed on 84 volatile peaks detected in the headspace above infected cultures. Cell culture flasks were capped for 30 minutes before sampling. *B. fragilis* and *F. nucleatum* samples (red circle) were distinct from *E. coli* Nissle and *E. faecalis* samples.
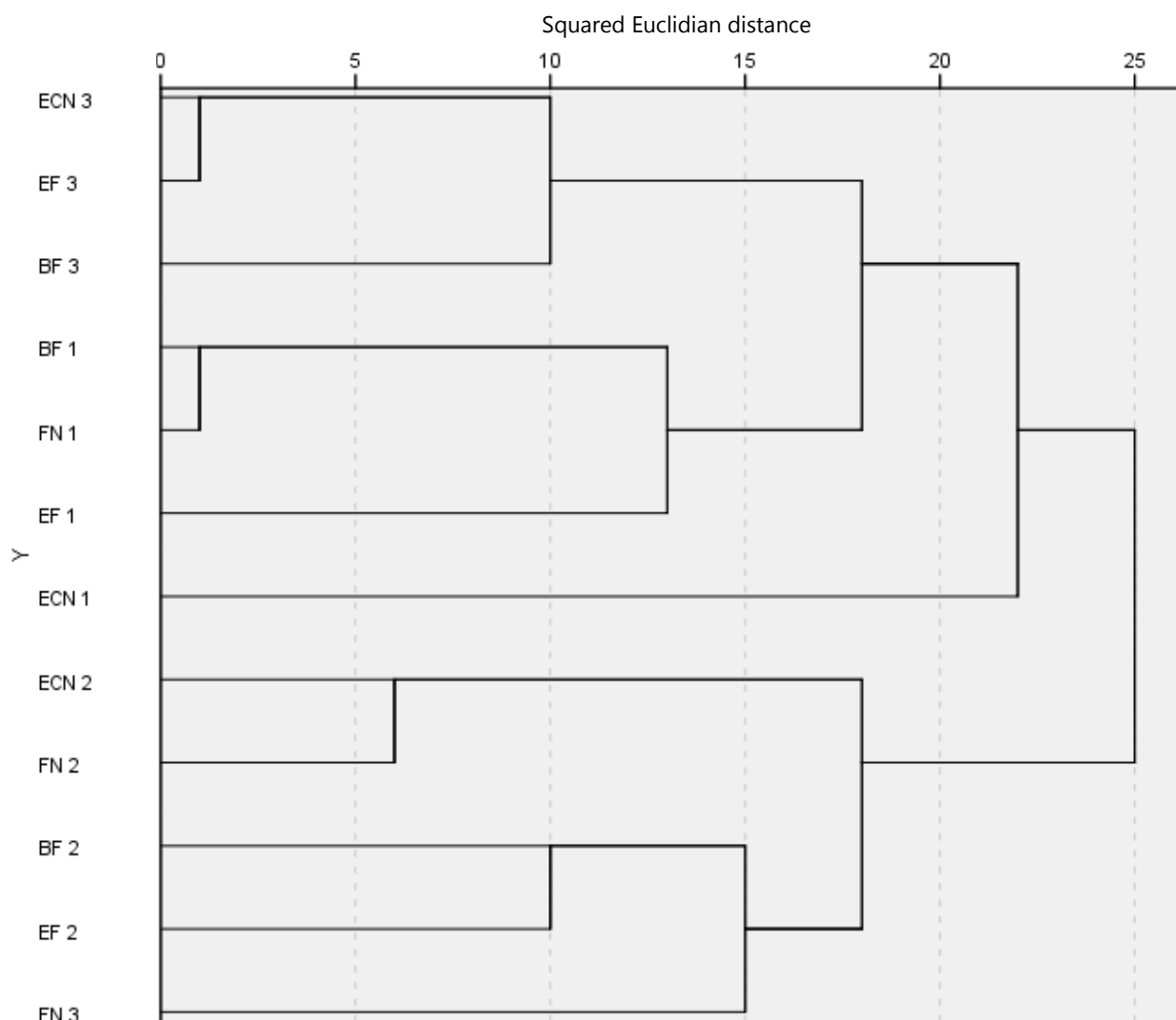
Figure 4.15: Dendrogram of volatile product ion peaks detected from the headspace above infected HCT116 cells. HCA was performed on 88 volatile peaks detected above infected HCT116 cells. No discrimination between infective bacterial species could be achieved. Cell culture flasks were capped for 30 minutes prior to sampling.
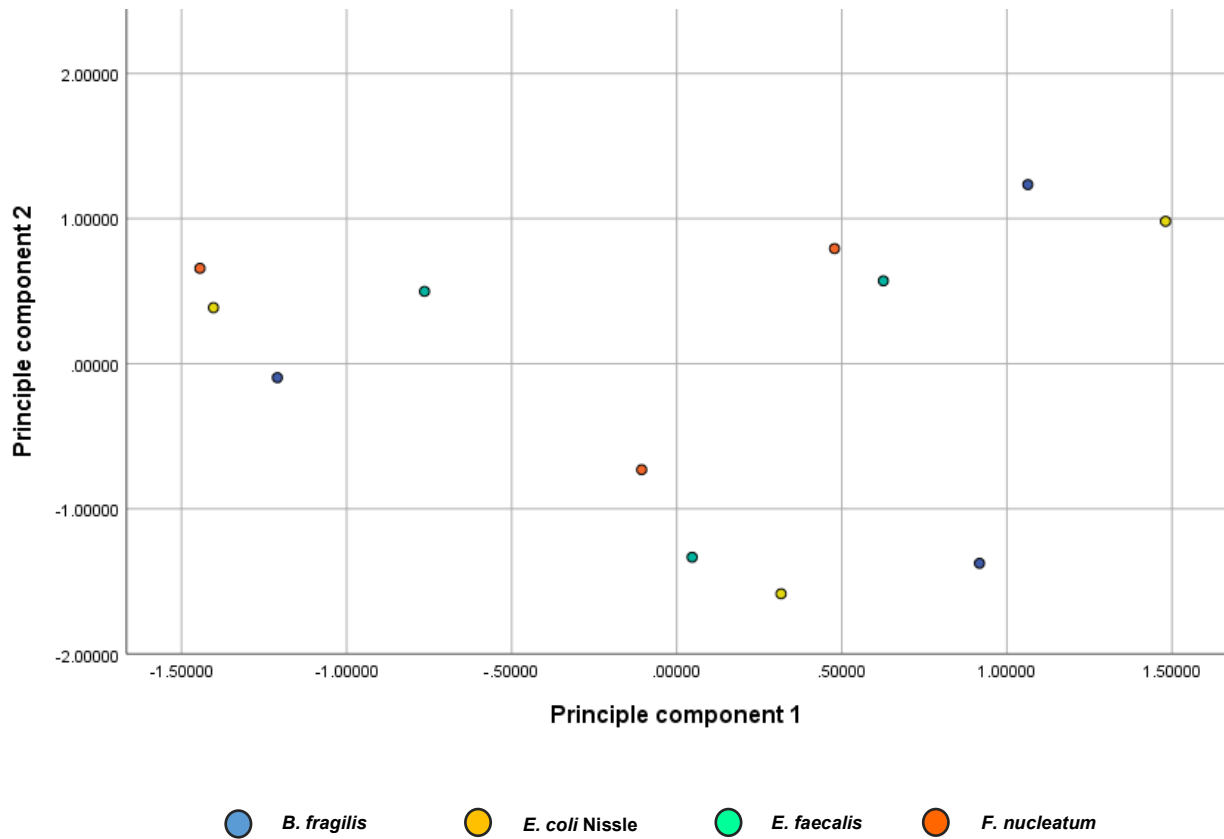
Figure 4.16: Principal component analysis of volatile product ion peaks detected from the headspace above infected HCT116 cells. Scatter plot produced using the first and second scores of the principal component analysis performed on 88 volatile peaks detected in the headspace above infected HCT116 cell cultures. Cell culture flasks were capped for 30 minutes before sampling.

4.5 Discussion

This chapter aimed to investigate how direct interactions with bacteria influenced tumour cell growth and metabolism. After previously establishing in Chapter 3 that these bacteria infiltrate and persist within tumour cells without negatively impacting viability, the potential tumour-promoting effects of this bacterial invasion was studied. Firstly, cell yield in response to bacterial infection was investigated. Bacterial infection significantly increased tumour cell yield in both cell lines, with similar fold increases observed for all four species of bacteria tested. Chapter 3 highlighted that levels of bacterial invasion are significantly different between tumour cell lines and between bacterial species. It is therefore surprising that the degree of cell yield increase was similar across species, and across two separate MOIs. This could suggest that relatively small numbers of bacteria are needed to significantly impact cell growth, which is consistent with current opinions on bacterial drivers and 'alpha-bugs' in CRC, which are thought to be present at very low concentrations (see Section 1.6; Sears and Pardoll, 2011). Induction of tumour growth by gut bacteria is not a novel concept; however, much of the current literature attributes the promotion of tumour formation by bacteria to genetic damage induced by the production of toxic metabolites or stimulation of the immune system, as opposed to direct interactions between tumour cells and bacteria (see reviews by Candela *et al.*, 2014; Louis, Hold and Flint, 2014). Furthermore, comparative analysis of the effects of bacteria on adenoma and adenocarcinoma cells is also lacking. The results of the present study demonstrate that these bacteria directly contribute to an increase in tumour cell yield in benign and malignant cells through contact-dependent mechanisms. The similar increases observed in RG/C2 and HCT116 cells suggest that the mechanism is likely to be conserved between adenomas and adenocarcinomas. Crucially, this suggests that

bacteria may contribute to the progression of CRC, as an increase in tumour cell yield allows for clonal expansion and the subsequent accumulation of further driver mutations (Greaves and Maley, 2012).

Aberrant cell growth is the most fundamental hallmark of cancer formation, and can be attributed to sustained proliferation and resistance to cell death (Hanahan and Weinberg, 2011). To investigate whether bacterial infection was increasing proliferation or reducing apoptosis in colorectal tumour cells, parallel counts of floating cell populations were performed during cell yield assays. A simple count of floating cells is a viable measure of apoptosis in *in vitro* colon tumour cell cultures, as >99% of floating cells are apoptotic, as confirmed by acridine orange staining (Hague *et al.*, 1993). With the exception of HCT116 cells treated with *F. nucleatum* at a high MOI, bacterial infection reduced apoptosis, although this was not statistically significant in all cases. Apoptosis at the mucosal surface is essential to gut homeostasis. Cells are continuously sloughed into the lumen to be replaced by cells emerging from the crypts; this removes damaged cells and prevents abnormal cell proliferation (Sträter *et al.*, 1995). In contrast, apoptosis within the colonic crypts is relatively rare in healthy epithelium. The distribution of apoptosis is reversed in colorectal tumours, with reduced apoptosis at the mucosal surface and increased apoptosis in the crypts (Moss *et al.*, 1996). This highlights the influence of apoptosis deregulation to cancer development. There is much controversy surrounding the relative levels of apoptosis in adenomas and adenocarcinomas *in vivo,* which has been extensively reviewed by Koornstra *et al.* (2003). However, *in vitro* adenomas display much higher levels of apoptosis when compared to adenocarcinoma-derived cells (Qualtrough *et al.,* 2004), as highlighted by the relative levels of floating cells

presented in Figure 4.2. Nevertheless, the overall effect of bacterial infection reducing apoptosis was consistent across both cell lines.

For each of the four species of bacteria investigated, distinct mechanisms have been proposed which could explain the observed increases in cell yield and reductions in apoptosis (see Section 1.5). For example, cleavage of E-cadherin by BFT has been shown to induce c-myc expression, resulting in sustained proliferation (W*u et al.*, 2003). *F. nucleatum* is also thought to promote aggressive tumour behaviour through the interaction of its FadA adhesin with E-cadherin (Rubinstein *et al.*, 2019). However, as previously discussed levels of E-cadherin expression differ between cell lines, suggesting this cannot be the sole mechanism responsible (see Section 1.5.5). Multiple other genetic and phenotypic differences between RG/C2 and HCT116 cells exist which make their near-identical response to bacterial infection surprising. For example, RG/C2 cells are hemizygous for *TP53*, an important tumour suppressor gene which influences apoptosis, with the remaining allele being non-functional (Williams *et al.*, 2000). In contrast, HCT116 cells possess wild-type *TP53* and therefore the mechanism of apoptosis reduction by bacteria is likely to be *TP53*-independent. Further investigation is required to elucidate this mechanism, which may differ according to each bacterial species.

It could be deemed counterintuitive that invasion by bacteria decreased apoptosis in colonic tumour cells, given that intracellular gut pathogens such as *Salmonella* serovars and invasive *E. coli* have been shown to induce apoptosis in infected epithelial cells (Kim *et al.*, 1998). However, suppressing apoptosis is an important virulence factor for many intracellular pathogens, as it provides time for intracellular replication and subsequent dissemination (reviewed by Behar and Briken, 2019). The similar results observed across all bacterial species suggest that a common

characteristic may be responsible. It was recently discovered that intracellular *Shigella flexneri* lipopolysaccharide (LPS), found in the outer membrane of Gram negative bacterial cell walls, binds to and inhibits pro-apoptosis caspases (Günther *et al.*, 2020). This activity was specific to the LPS O-antigen moiety and would not account for anti-apoptotic activity of Gram-positive species such as *E. faecalis*, but is an example of how fundamental structural components of prokaryotes may influence tumour cell behaviour.

The fold increase in cell yield across all species was far larger than the fold decrease in apoptosis. Furthermore, high infection doses of *F. nucleatum* increased the apoptotic cell count, albeit non-significantly, despite infection with these species significantly increasing cell yield. This raises questions as to whether a reduction in apoptosis could be solely responsible for the increases in cell yield. However, it should be considered that apoptotic floating cells were only sampled at the end of the experiment. Therefore, additional floating cells may have been discarded during media changes, or have degraded to debris during the incubation period and therefore not have been counted. To determine whether the observed increases in tumour cell yield could be attributed to an increase in proliferation upon bacterial infection, the tumour cell cycle was analysed. S-phase analysis was performed after staining tumour cell DNA. RG/C2 cells exhibited decreased proportions of cells in G1-phase after infection with almost all species. *E. faecalis* and *F. nucleatum* infection caused a significant increase in the number of cells in G2-phase, which is perhaps indicative of a delay in mitosis (Sorenson, Barry and Eastman, 1990).

Conversely, increased proportions of cells in G1-phase were observed in HCT116 cells infected with all species. A G1-phase cell cycle arrest would be expected in colorectal tumour cells experiencing serum starvation or density inhibition

(Tay, Bhathal and Fox, 1991); however, cell cycle arrest at various stages of the cell cycle have been observed upon bacterial infection. For example, DNA damage caused by infection with *Staphylococcus aureus* leads to G2/M-phase delay in HeLa cells (Deplanche *et al.*, 2019). Similarly, the gut pathogen *Salmonella enterica subsp. enterica* ser. Typhimurium induces a G1-phase arrest in HCT116 cells, which facilitates further bacterial invasion (Mambu *et al.*, 2020). The results of the analysis performed in this thesis appear to suggest that bacterial infection did not increase tumour cell proliferation, which would be demonstrated by an increase in the proportion of cells in S-phase. However, it must be considered for this analysis that cells were only sampled at a single time point post-infection (24 hours), and so additional effects on cell cycle as a result of bacterial infection may not have been observed. To better understand the effect of bacterial infection on tumour cell yield and apoptosis, parallel tumour cell counts and flow cytometry samples should be taken at regular intervals across the complete 96-hour period and compared.

Tumour cell growth is intrinsically linked to tumour metabolism. In cancer cells, the rate-limiting factor on cell growth is not the rate of ATP production, but the availability of small molecules such as NADPH (Vander Heiden, Cantley and Thompson, 2009). Therefore, tumour cells exhibit the Warburg effect, whereby cells preferentially undergo aerobic glycolysis, despite the availability of sufficient oxygen for oxidative phosphorylation (Warburg, Wind and Negelein, 1927). In addition to the increased small molecule availability, the Warburg effect confers numerous additional benefits to tumour cells. For example, increased glucose uptake allows tumour cells to out-compete neighbouring cells with slower metabolic rates (Pfeiffer, Schuster and Bonhoeffer, 2001). In addition, secretion of increased amounts of lactate, produced by aerobic glycolysis, into the microenvironment inhibits the activity of T-cells and

natural killer cells (Brand *et al.*, 2016). Finally, despite aerobic glycolysis being a far less efficient means of generating ATP when compared to oxidative phosphorylation, the speed at which cells are able to perform glycolysis means that the net rate of ATP production is comparable (Shestov *et al.*, 2014).

The Warburg effect has been observed in many types of cancer, but there is ongoing debate as to whether it initiates tumorigenesis or is a consequence of tumour formation (Devic, 2016). In CRC, loss of *APC* induces *PKM2* and *LDHA* expression via β-catenin-related transcription, resulting in increased glucose uptake and lactate secretion (Cha *et al.*, 2021). Due to the prevalence of *APC* mutations in CRC, the Warburg effect is therefore thought to be activated early in colorectal tumours (Satoh *et al.*, 2017). In this study, the Seahorse Xf analyser was used to investigate the metabolic profiles of selected colorectal tumour cell lines. Although oxygen consumption rate (OCR) differed between all cell lines, benign RG/C2 cells displayed a higher OCR than all three adenocarcinoma cell lines. This was best reflected by the maximal OCR of the RG/C2 cell line, which was three-fold higher than the nearest adenocarcinoma cell line (SW480). In contrast, extracellular acidification rates (ECAR) were similar across all cell lines. This would suggest that under these conditions adenocarcinoma cells have a lower metabolic rate due to a decreased rate of oxidative phosphorylation.

Many tumour cells face hypoxia as their rate of growth overtakes the rate of angiogenesis needed to maintain sufficient blood supply (Brahimi-Horn, Chiche and Pouysségur, 2007). In CRC, hypoxia has been demonstrated to increase along the adenoma-carcinoma sequence, with the expression of hypoxia-inducible factor 1 alpha (HIF-1α) occurring in early adenomas and being highest in invasive cancers (Simiantonaki *et al.*, 2008). Hypoxia-inducible factors regulate gene expression to

allow adaptation to hypoxic environments, and actively downregulate oxygen consumption by the mitochondria through pyruvate dehydrogenase kinase 1 (PDK1) while stimulating glycolysis-related genes (Papandreou *et al.*, 2006). In order to further investigate the metabolic differences between RG/C2 cells and adenocarcinoma cells, investigating the activation of HIF-1α-related genes would be appropriate. This could be achieved using commercially available luciferase-based hypoxia reporter assays (Doran *et al.*, 2011).

Finally, although the OCR and ECAR data presented have been normalised to cell number, it must be considered that each cell line may respond differently to the experimental conditions required to perform Seahorse assays. In particular, RG/C2 cells are routinely cultured in media containing insulin, hydrocortisone and increased concentrations of serum and L-glutamine, all of which are absent from Seahorse assay media. In addition, these cells are more susceptible to anoikis when seeded at lower densities (Qualtrough 2020, personal communication). Therefore, the use of non-buffered, non-supplemented Seahorse media, incubation in a CO_2-free incubator, and the necessity to seed at lower densities may influence the metabolic profiles of specific cell lines depending on their sensitivity to these factors. In an attempt to overcome issues relating to seeding density, one assay was performed comparing the metabolic profiles of confluent cell monolayers. As cells become confluent, changes in membrane structure and signalling pathways including YAP/TAZ and cAMP occur which alter cellular metabolism. Furthermore, the diffusion dynamics of molecules such as glucose are altered in confluent monolayers, and this change in nutrient accessibility may be reflected in altered OCR and ECAR rates (Pelletier *et al.* 1990). However, when attempting to investigate the metabolism of confluent cell monolayers, the increased number of cells produced OCR and ECAR values which

exceeded the limits of detection for the Seahorse (data not shown). The manufacturer now supplies Seahorse assay plates with an increased volume, and therefore increased capacity for detection of metabolic changes. Repeating these assays with these plates would allow comparison of the metabolic profiles of confluent cells, which may reduce any effects of low seeding density and be more representative of *in vivo* conditions. However, the metabolic effect of confluence *in vitro* is also significant, and could influence metabolic outputs. For example, lactate is produced at high quantities in actively dividing cultured cells, whereas there is a net consumption of lactate in stationary phase (which represents the majority of cells in a confluent monolayer) cells (Zagari *et al.*, 2013). Altered lactate metabolism not only impacts mitochondrial oxidation, but also alters media acidity, which can have further metabolic affects.

The impact of bacterial infection on the metabolism of HCT116 cells was also analysed using the Seahorse. The OCR did not increase upon infection with any of the species studied. However, both *B. fragilis* and *F. nucleatum* increased the basal ECAR, with *F. nucleatum* also increasing the maximal ECAR. This suggests that while both species are capable of triggering a metabolic shift towards aerobic glycolysis, the action of *F. nucleatum* increases tumour cells' capacity for glycolysis and therefore may promote aggressive behaviours as discussed above. Metabolic changes are part of the 'core host response' to bacterial infection and are used as a means to suppress bacterial survival and proliferation (Boldrick *et al.*, 2002). These are often triggered by non-specific bacterial factors (e.g. LPS) and result in the production of antimicrobial reactive oxygen species (Hsu and Wen, 2002). The lack of metabolic response to the Gram positive (and therefore LPS-negative) *E. faecalis* suggests that this may be the case in infected HCT116 cells. It has been reported that

immune cells, such as M1 and M2 macrophages, shift to a Warburg-like metabolism upon infection with bacteria (Escoll and Buchrieser, 2018). Despite being an innate response of the immune system, there is evidence to suggest that intracellular pathogens may benefit from these metabolic changes. For example, *Salmonella* enterica sv. Typhimurium requires abundant intracellular glucose during infection, in addition to the associated glucose transport to invade and replicate within macrophages (Bowden *et al.*, 2009). Furthermore, intracellular pathogens such as *Mycobacterium tuberculosis* benefit from increased glycolysis as they are able to metabolise the resulting lactate (Billig *et al.*, 2017). It is therefore plausible that species such as *B. fragilis* and *F. nucleatum* trigger glycolysis in colonocytes to potentiate their own survival and replication.

To complement the metabolic analysis performed using the Seahorse, the volatile product ion peaks detected from the headspace above tumour cell cultures were analysed using SIFT-MS. The detection of volatile compounds from a range of biological materials has shown potential as a novel diagnostic tool. Benefits of this approach are the speed of analysis and the non-invasive nature of sample collection; and there is a wealth of research investigating the application of volatile analysis to a range of pathologies including bacterial infection, gastrointestinal disorders and cancer (Altomare *et al.*, 2013; Slade *et al.*, 2017; Rondanelli *et al.*, 2019). These studies have demonstrated discrimination between bacterial species, and between diseased and healthy individuals; however, the underlying metabolic pathways responsible for the production of specific volatile peaks remains less understood.

Analysis of the statistically significant volatile product ion peaks produced by RG/C2, HCT116, HT29 and SW480 cell lines revealed that colorectal adenomas can potentially be discriminated from colorectal adenocarcinomas based upon their

volatile profiles. Numerous peaks were present across all cell lines, representing conserved metabolic pathways. This is consistent with previous studies using gas chromatography mass spectrometry (GC/MS) which demonstrated that SW480 cells shared many peaks with NCM460 cells, which are considered a model of normal colon (Zimmermann *et al.*, 2007). More recently, analysis of urine volatile product ion peaks via gas chromatography with ion-mobility spectroscopy was shown to be able to discriminate between healthy patients, patients with adenomas and patients with adenocarcinomas by analysing urine samples from individuals diagnosed with colorectal adenomas, CRC patients and healthy controls (Mozdiak *et al.*, 2019). Similar studies have shown distinct biomarkers present in the blood of lung cancer patients (Deng, Zhang and Li, 2004). This research highlights that the differences in tumour cell metabolism may be reflected in systemic metabolite levels and be detectable from relatively less-invasive samples.

The bulk of research into the use of volatile analysis for cancer detection has focussed on the use of GC/MS and liquid chromatography mass spectrometry (LC/MS). However, SIFT-MS offers the advantages of being simple to use, having high sensitivity and specificity, and the ability to simultaneously analyse a wide range of volatile compounds. Regardless of the method of mass spectrometry used, metabolite-based screening methods of patient samples could complement or even replace current CRC screening methods. Bond *et al.* (2019) demonstrated that analysis of propan-2-ol detected from the headspace above faecal samples using solid-phase microextraction (SPME) GC/MS displayed greater diagnostic capabilities than the currently employed faecal immunochemical test (FIT) programme in a UK cohort (Bond *et al.*, 2019). However, it must be considered that uptake of faecal-based screening programmes worldwide is low. A recent study in a Dutch population

investigating FIT participation failed to achieve an uptake of greater than 63% despite multiple rounds of invitation (Van Der Vlugt *et al.*, 2017). There are various reasons for this low uptake rate, although stigma surrounding the handling of faeces and a general lack of medical knowledge are thought to contribute. It has been demonstrated that simplifying the screening procedure (by reducing the number of samples required) significantly improves participation (Cole *et al.*, 2003). Volatile screening could represent a further simplification of this process, particularly as breath analysis via SIFT-MS has shown promise in diagnosing other cancers including head and neck squamous cell carcinoma (Chandran *et al.*, 2019). The results of the principal component analysis (PCA) presented in this study demonstrate that colorectal adenoma cells can be discriminated from adenocarcinoma cells when cultured *in vitro*. The next goal of this research would be to identify the specific compounds responsible for these differences. Once identified, these products could be analysed from *in vivo* samples to determine whether they represent an appropriate biomarker for advanced disease.

Numerous studies have shown the potential for detection and discrimination of infectious agents by volatile analysis alone (Slad*e et al*., 2017; Traxler *et al*., 2019). Furthermore, it has been reported that CRC can be identified via volatile markers detected from the headspace above faecal samples using SIFT-MS with similar results to the faecal occult blood test (FOBT; Batty *et al.*, 2015). In addition, CRC patients have been found to possess unique breath biomarkers which can diagnose with greater accuracy than current screening techniques (Altomare *et al.*, 2013; Woodfield *et al.*, 2018). The majority of the detected compounds are thought to be attributed to CRC-specific microorganisms. Therefore, detecting bacteria within patient tumours could be greatly beneficial to CRC diagnosis and treatment. One study demonstrated that

mice with *F. nucleatum*-positive colorectal tumours treated with metronidazole exhibited reduced *F. nucleatum* load and decreased tumour growth (Bullman *et al.*, 2017). However, it has also been found that indiscriminate use of antibiotics can promote colorectal adenomas and CRC *(*Cao *et al.*, 2018; Zhang *et al.*, 2019). Therefore, the use of SIFT-MS to detect biomarkers for tumour cells infected with specific bacteria could lead to diagnostic approaches that help guide the administration of adjuvant antibiotic therapy in CRC patients.

Bacterial infection did produce a number of statistically significant volatile product ion peaks in both HCT116 and RG/C2 cells. However, RG/C2 cells infected with *E. coli* Nissle, *E. faecalis* and *F. nucleatum* only produced a small amount of significant m/z peaks. Hierarchical cluster analysis was able to separate *B. fragilis* and *F. nucleatum* in RG/C2 cells, but no discrimination between species was possible in infected HCT116 cells. To the author's knowledge discrimination between these species via volatile analysis has not been previously demonstrated. However, mass spectrometry has been shown to accurately discriminate between other infectious agents in other laboratory co-culture models. For example, Detroit cells infected with *Streptococcus pyogenes* and influenza A virus exhibited distinct volatile profiles which allowed for discrimination between both infectious agents and between co-infection using needle-trap microextraction coupled with GC/MS (Traxler *et al.*, 2019). The level of specificity, and the abundance of compounds detected in HCT116 cells suggests that discrimination between infecting bacterial species may be possible using a more refined method. Metabolites detected from cultured CRC cells can vary significantly in both intensity and distribution depending upon the solvent used for extraction (Ibáñez *et al.*, 2017). Therefore, the use of a solvent (e.g. formic acid) may allow for more in-depth analysis of these samples. Finally the length of the incubation

period post-infection, in addition to the period for which the filter caps were sealed to allow volatile compound accumulation should be further optimised to ensure any changes in volatile profile are adequately detected.

The results presented in this chapter demonstrate that infection of colorectal tumour cells by gut bacteria promotes an increase in tumour cell yield. Crucially, this effect was observed in both benign and malignant cells, indicating that these bacteria can contribute towards multiple stages of CRC development and progression. There are numerous mechanisms through which tumour cell yield can be upregulated. Quantification of floating cells suggests that these bacteria inhibit apoptosis in tumour cells, supported by the results of S-phase analysis. The metabolic differences between benign and malignant colorectal tumour cell lines were highlighted using Seahorse assays and SIFT-MS, with *B. fragilis* and *F. nucleatum* inducing higher rates of glycolysis in HCT116 cells. In summary, these results demonstrate a bacterial contribution to increased tumour growth, which is a key hallmark of cancer and an important prognostic indicator. The following chapter will aim to investigate the effects of these species on tumour cell migration and invasion, an equally important hallmark of cancer which governs the progression of late-stage disease and has the largest impact on patient prognosis.

Chapter 5: Gut bacteria influence tumour cell motility in a tumour stage-dependent manner

5.1 Introduction

In the UK, approximately 25% of CRC patients present with synchronous metastases upon diagnosis, and just 44% of these individuals survive for longer than one year (Cancer Research UK, 2021). This is in stark contrast to pre-metastatic patients, where one year survival rates exceed 90%. The most common sites for CRC metastasis are the liver and lungs, and secondary tumours at these sites are found in more than 50% of metastatic-CRC patients (Riihimaki *et al.*, 2016). Tumour spread to the peritoneum is also a common feature of advanced CRC, although there is some debate as to whether this represents metastasis or local invasion (Pretzsch *et al.*, 2019). It is often reported that cancer metastasis is responsible for 90% of all cancer-related deaths (Mehlen and Puisieux, 2006). A recent study investigating deaths caused by cancer across a range of diseases found that metastasis was present in 80.2% of CRC deaths, the second highest proportion behind ovarian cancer (90.4%) (Dillekås, Rogers and Straume, 2019). The contribution of metastatic spread to cancer mortality can be attributed to two factors: (1) secondary tumours are difficult to detect (Menezes *et al.*, 2016), and can cause disease recurrence at secondary sites decades after initial treatment (Baird *et al.*, 2020). (2) despite the identification of numerous potential therapeutic targets, there is currently a paucity of treatment options for metastatic disease (Ganesh and Massagué, 2021).

Tumour metastasis is a multistep process which involves local invasion of surrounding tissue, intravasion into the circulatory system, followed by transportation to a secondary site, extravasation from the vasculature and eventual tumour colonisation

(Pantel and Brakenhoff, 2004). Tumour cells can migrate through a number of mechanisms including as single mesenchymal-like cells or as multicellular clusters (Lintz, Muñoz and Reinhart-King, 2017). A fundamental prerequisite for the formation of metastatic tumour cells is the gaining of motility and the ability to remodel the surrounding extracellular matrix (ECM). In epithelial tumours such as those seen in CRC, this involves reorganisation of the cytoskeleton, a loss of adherens junctions and resistance to anoikis (Friedl and Wolf, 2003). Upregulation of several pathways has been associated with colorectal tumour motility, however, this process is thought to be primarily driven by the TME (Gout and Huot, 2008).

The gut microbiome is both a key component and regulator of the TME and has been suggested as a modulator of tumour motility in CRC. For example, virulence by gut microorganisms can lead to increased tight junction permeability, which is also a hallmark of CRC (Soler *et al.*, 1999). Disruption of adherens junctions in the colon is also a virulence factor of many bacterial species, with the increased epithelial barrier permeability facilitating bacterial access to underlying tissue (Desai *et al.*, 2016). Furthermore, bacterial effectors including BFT and FadA, produced by *B. fragilis* and *F. nucleatum*, respectively, have been demonstrated to upregulate β-catenin signalling, an important regulator of both migration and epithelial-mesenchymal transition (EMT; Wu *et al.*, 2003; Rubinstein *et al.*, 2019). This corresponds with nuclear accumulation of β-catenin which is a common feature of the invasive front in metastatic CRC (Suzuki *et al.,* 2008). In the previous chapter, it was established that direct contact with selected species may contribute to disease progression by promoting an increase in cell yield. The acquisition of motile characteristics in benign cells is a crucial step in malignant transformation, and as previously discussed, the migration and invasion of malignant tumour cells is the largest contributor to patient mortality. Therefore, any

contribution to migratory behaviour by bacteria would represent a key role for bacteria in the progression of CRC.

5.2 Aims

The aim of this chapter was to investigate whether direct treatment with CRC-associated bacteria can induce motile and invasive behaviour in either benign and/or malignant colorectal tumour cells. Furthermore, where an effect on tumour migration or invasion was observed, this chapter also aimed to elucidate the underlying mechanisms responsible, with a view to identifying novel targets for further investigation.

5.3 Objectives

The objectives of this chapter were as follows:

- To investigate how infection with bacteria effects single and multiple cell migration in RG/C2 and HCT116 cells.

- To determine whether species which demonstrate an ability to promote migration are also capable of promoting tumour invasion *in vitro*.

- To investigate how bacteria influence β-catenin signalling activity using luciferase reporter assays.

- To use qPCR microarrays to quantify the expression a wide range of cell motility and EMT-related genes and determine how this expression alters upon bacterial infection.

5.4 Results

5.4.1 *E. coli* Nissle and *F. nucleatum* promote migration in HCT116 cells

The migration of infected colorectal tumour cells was assessed by seeding cells 24 hours post-infection into transwell filter migration assays (see Section 2.10.2). During this assay, tumour cells seeded in the apical compartment of the filter are attracted to the serum present in the lower compartment via chemotaxis and migrate through the filter towards it. As cells migrate, they become embedded in the collagen coating on the basal layer of filter, and can subsequently be fixed, stained and counted. The average number of migratory RG/C2 cells counted was 17 per field of view (Figure 5.1), which was significantly lower than the average number of HCT116 cells (91; Figure 5.2) despite being initially seeded at twice the density to protect the benign cells from the negative effects associated with a lack of cell-cell contact. Infection with bacteria had no effect on the migration of RG/C2 cells, although *E. coli* Nissle, *E. faecalis* and *F. nucleatum* caused small, non-statistically significant reductions in the number of migratory cells counted. In contrast, *E. coli* Nissle and *F. nucleatum* significantly increased the migration of HCT116 cells. *E. coli* Nissle-treated HCT116 cells had an average of 129 migratory cells ($p = 0.0169$), but the largest effect on migration was seen in *F. nucleatum*-treated cells, with an average of 173 cells counted ($p = 0.0257$).
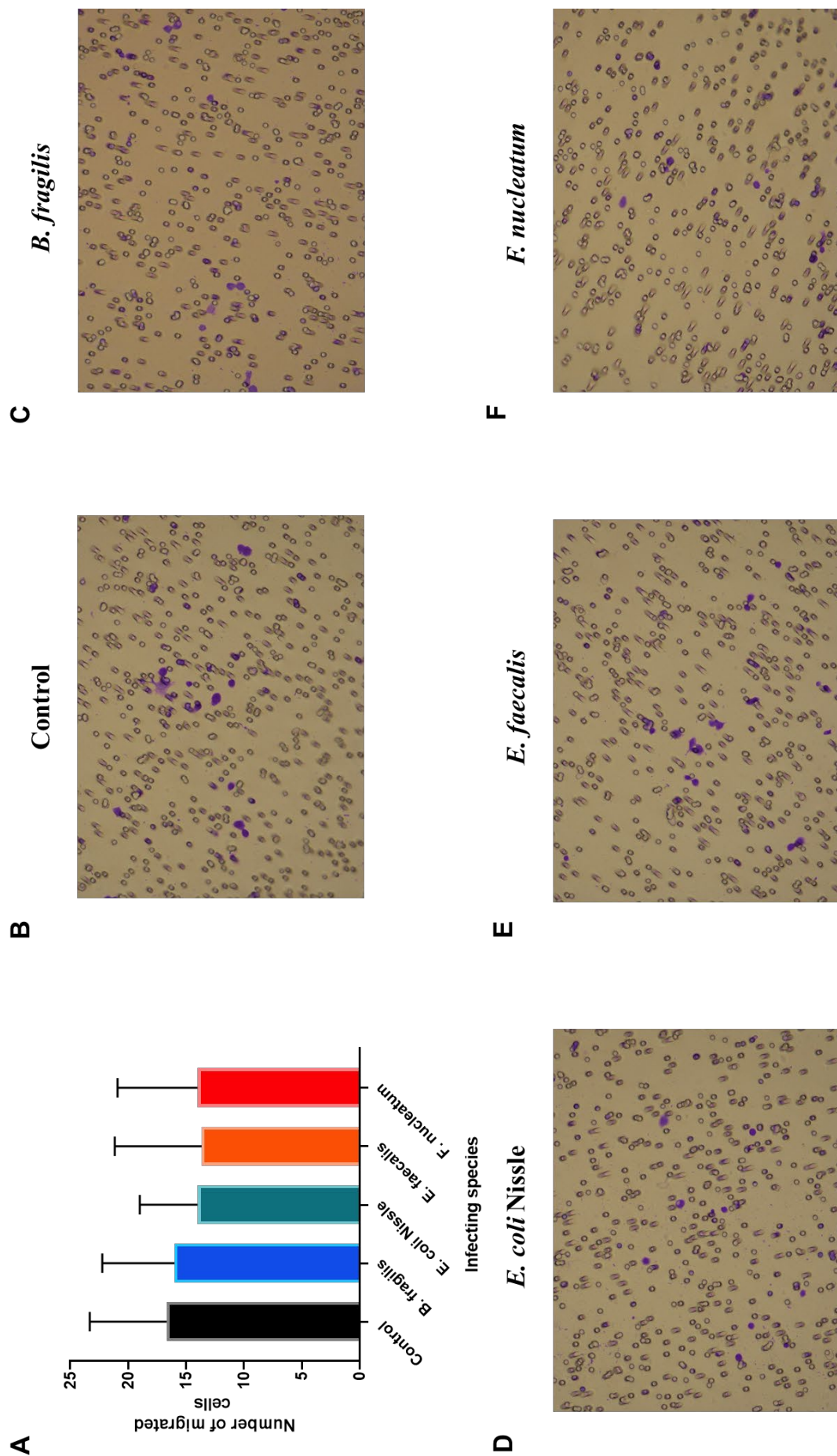
Figure 5.1: Migration of RG/C2 cells after bacterial infection. (A) Average number of migrated RG/C2 cells counted following bacterial infection. Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate. (B to F) Representative microscope images of bacteria-infected migratory cells (purple) after methanol fixation and staining with crystal violet taken at x200 magnification.
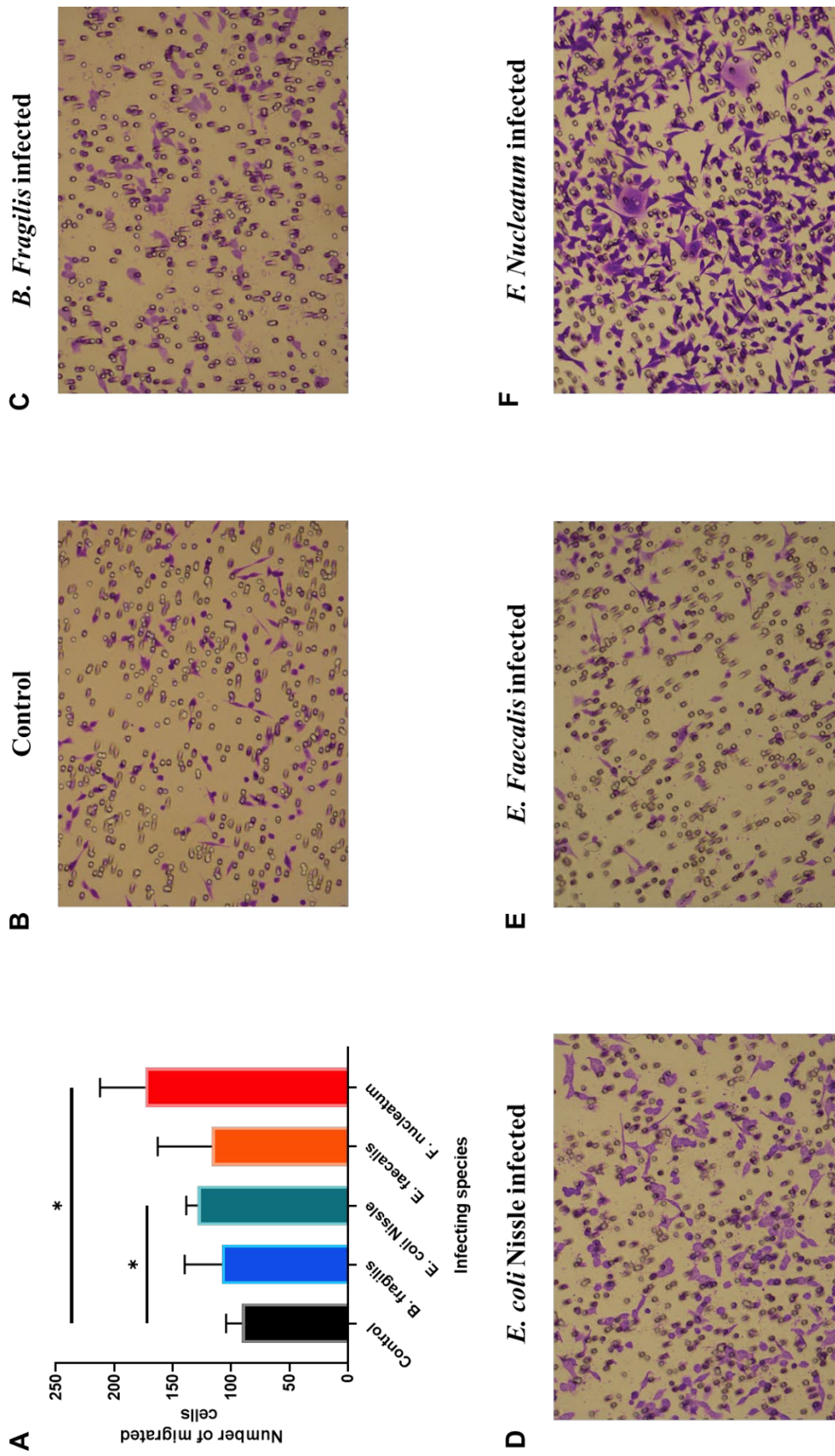
Figure 5.2: Migration of HCT116 cells after bacterial infection. (A) Average number of migrated HCT116 cells counted following bacterial infection. Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate. T test, $* = p < 0.05$. (B to F) Representative microscope images of bacteria-infected migratory cells (purple) after methanol fixation and staining with crystal violet taken at x200 magnification.

5.4.2 Wound healing in HCT116 cells is increased by *E. coli* Nissle and *F. nucleatum* infection

The results presented in Section 5.4.1 established that both *E. coli* Nissle and *F. nucleatum* increased single-cell migration in HCT116 cells. As the mechanisms of single-cell migration can differ from those of collective cell migration, the ability of bacteria to promote collective cell migration was also investigated using the wound healing assay. Furthermore, to the author's knowledge the experiments performed in Section 5.4.1 are the first to use previously infected cells in a transwell filter migration assay. Therefore, performing a secondary migration assay had the added benefit of providing validity to these results.

Confluent monolayers of tumour cells were infected with bacteria, before a sterile pipette tip (P200) was used to create a wound. Phase contrast images were taken of the wound at three separate locations per well, and the reduction in wound surface area was calculated after 24 hours. Untreated RG/C2 cells exhibited an average reduction in wound area of 11.53% (Figure 5.3). Infection with all species of bacteria tested inhibited wound healing, although not to a statistically significant degree. The largest decrease was seen with cells infected with *F. nucleatum*, where a reduction of 5.71% was observed ($p = 0.0736$). HCT116 cells displayed a greater capacity for wound healing than RG/C2 cells, with an average reduction of 15.01% in untreated cells (Figure 5.4). In contrast to RG/C2 cells, infection with bacteria increased wound healing, with the exception of *E. faecalis*-treated cells where the average wound area reduction was 14.89%. Infection with *B. fragilis* increased wound area reduction to 17.87%, although this was not statistically significant ($p = 0.0574$). However, both *E. coli* Nissle ($p = 0.0380$) and *F. nucleatum* ($p = 0.0399$) infection significantly increased wound healing in HCT116 cells to 27.90% and 24.53%, respectively.

Finally, RG/C2 cultures exhibited a large number of cellular debris during wound healing, despite cells being washed in PBS after wound creation to remove debris. In contrast, minimal debris was observed in HCT116 cultures (Figure 5.3B and 5.4B), suggesting that scratching RG/C2 monolayers using a pipette tip may reduce viability in the remaining cells.
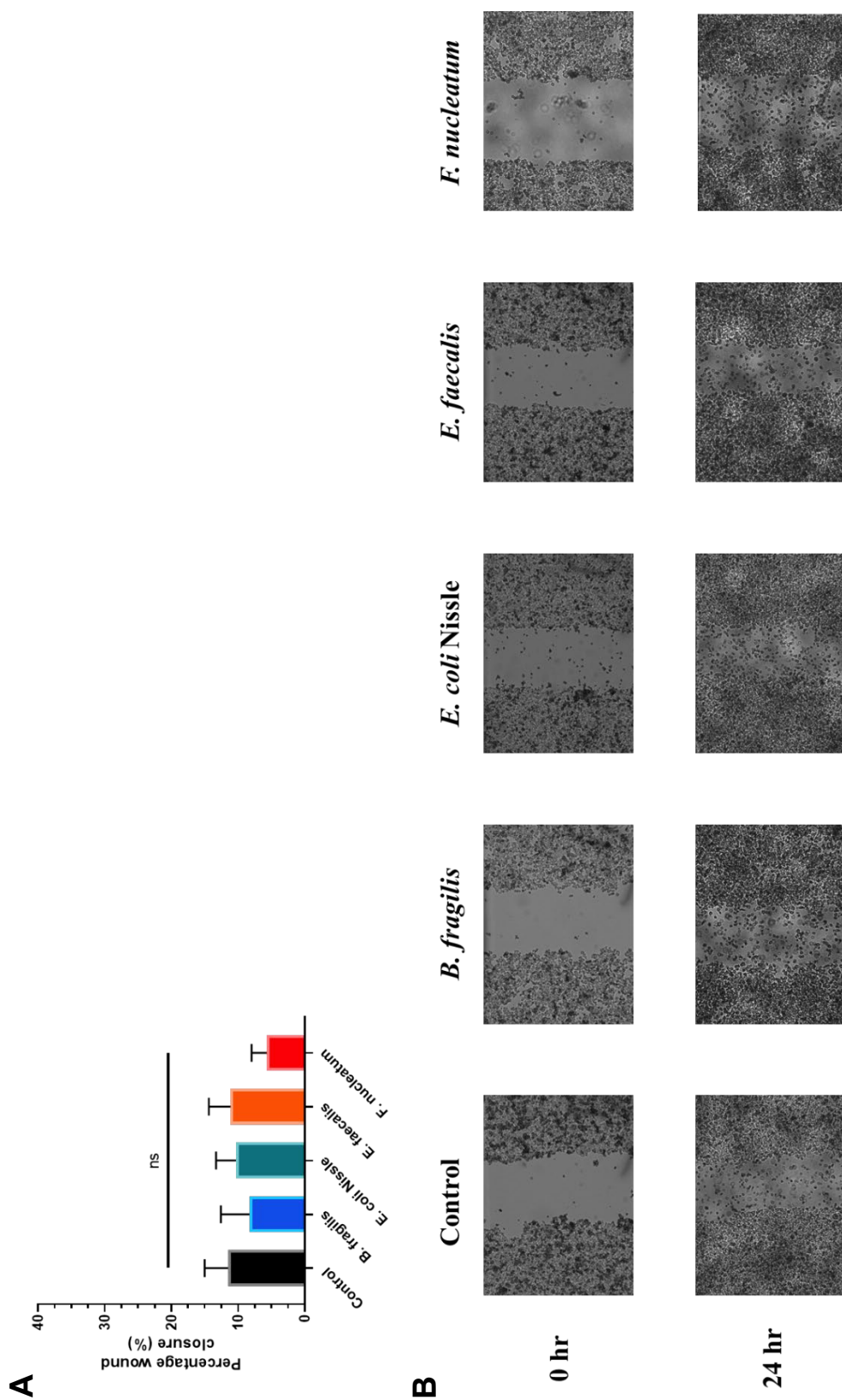
Figure 5.3: Wound healing in RG/C2 cells after bacterial infection. (A) Average reduction in wound surface area in RG/C2 cells, measured using ImageJ Wound healing analysis tool. Error bars represent standard deviation and data represents the results of three independent experiments performed in triplicate (B). Representative images of RG/C2 wounds at 0hrs and 24hrs taken at 40x magnification using a phase contrast microscope. Significant cellular debris was observed at 24hrs; however, this did not interfere with image analysis.
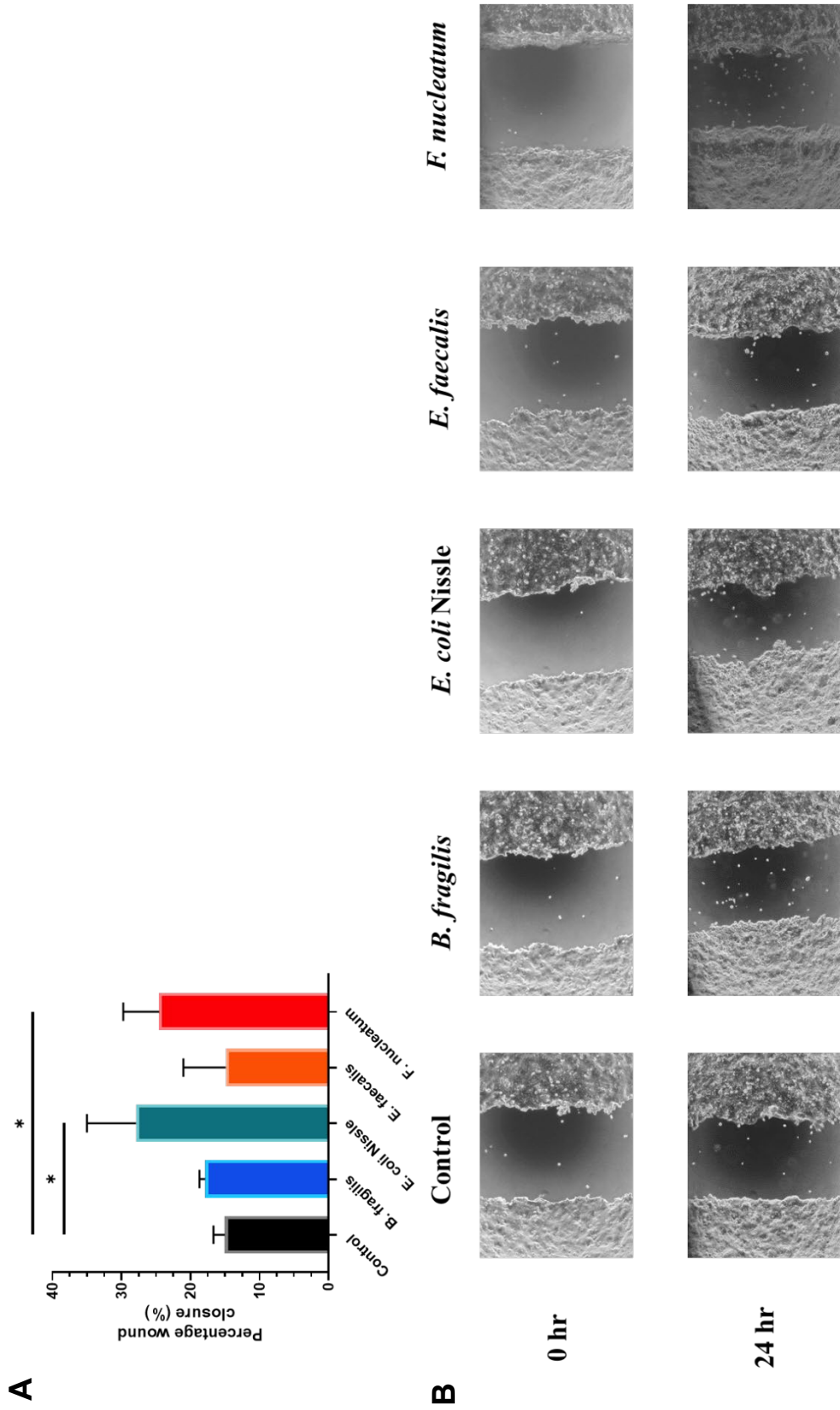
Figure 5.4: Wound healing in HCT116 cells after bacterial infection. (A) Average reduction in wound surface area in HCT116 cells, measured using ImageJ Wound healing analysis tool. Error bars represent standard deviation and results of three independent experiments performed in triplicate. (B) Representative images of HCT116 wounds at 0hrs and 24hrs taken at 40x magnification using a phase contrast microscope.

*= $p < 0.05$.

5.4.3 *F. nucleatum* promotes invasion in HCT116 cells

Having established that *E. coli* Nissle and *F. nucleatum* infection is able to promote both single and collective cell migration in HCT116 cells, the effect of these species on HCT116 invasion was investigated. In addition to migration, cellular invasion requires the remodelling of the extracellular matrix, which is facilitated by the secretion of enzymes such as matrix metalloproteinases which degrade ECM components e.g. gelatin (Murphy *et al.*, 1985). Invasion was quantified using transwell filters coated in Matrigel®, a gelatinous protein mixture produced by mouse sarcoma cells which resembles complete mammalian ECM (Hughes, Postovit and Lajoie, 2010). Therefore, unlike transwell filter migration assays where cells are able to freely pass through the filter pores, these invasion assays require tumour cells to produce the necessary enzymes to digest and subsequently invade the Matrigel® substrate.

On average, 37 invasive HCT116 cells were counted per field of view in the control group (Figure 5.5). Infection with *F. nucleatum* significantly increased HCT116 invasion, with an average of 61 cells counted ($p = 0.0326$). Infection with *E. coli* Nissle had no effect on HCT116 invasion. The effect of infection with by *B. fragilis* or *E. faecalis* was not assessed due to the cost of performing Matrigel® invasion assays, and the assumption that there was unlikely to be an effect on tumour cell invasion when no effect on migration was observed. A single invasion assay was also performed using RG/C2 cells infected with both *E. coli* Nissle and *F. nucleatum*; however, no invasive cells were detected in the control or either of the treated cell groups, and so no further replicates were performed (data not shown). This was expected, as RG/C2 cells were derived from a benign (and therefore non-invasive) tumour.
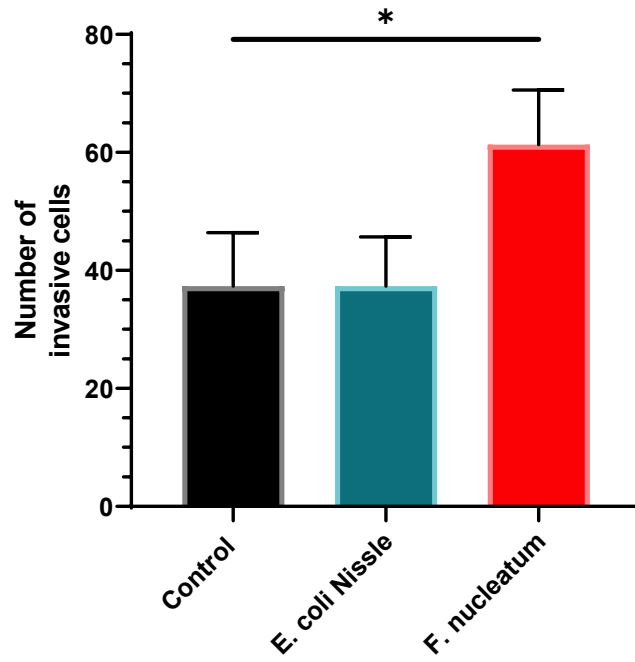
Figure 5.5: Invasion of HCT116 cells after bacterial infection. Average number of invasive HCT116 cells counted after migration through Matrigel® coated transwell filter inserts. Error bars represent standard deviation, and results of three independent experiments repeated in triplicate. *= $p < 0.05$.

5.4.4 Promotion of migration and invasion of HCT116 cells by

F. nucleatum is not accompanied by a change in catenin-related

transcription

It has been reported in the literature that *F. nucleatum* promotes catenin-related transcription (CRT) in colorectal tumour cells (Rubinstei*n et al*., 2019). Furthermore, the acquisition of an invasive phenotype normally involves epithelial-mesenchymal transition (EMT), which is accompanied by accumulation and nuclear translocation of β-catenin (Schmalhofer, Brabletz and Brabletz, 2009). To determine whether the observed effects on tumour cell migration and invasion seen in *F. nucleatum*-infected cells could be attributed to this signalling pathway, SDS-PAGE followed by Western blotting was performed on protein samples collected from infected cells. Figure 5.6B illustrates a representative image of a Western blot for β-catenin performed on RG/C2 and HCT116 cells infected with *F. nucleatum*. Infection with *F. nucleatum* had no significant effect on the level β-catenin recovered in either cell line. This was confirmed using densitometry to quantify relative levels of protein concentration in each band (Figure 5.6A), which after normalisation using beta-tubulin revealed that RG/C2 cells possessed approximately five times the amount of β-catenin protein when compared to HCT116 cells.

As nuclear localisation of β-catenin is essential for CRT, a change in CRT is possible without a change in overall levels of protein due to changes in sub-cellular localisation. To accurately quantify CRT in infected tumour cells, the TOPFLASH assay was performed (see Section 2.11; Molenaar *et al.*, 1996). The ratio of TOPFLASH/FOPFLASH luminescence accurately quantifies CRT, after normalisation to cell number using luminescence produced by the Renilla plasmid. TOP/FOP ratios were expressed as a fold change of the control for both RG/C2 (Figure

5.7A) and HCT116 (Figure 5.7B) cells. The high-throughput nature of this assay also allowed investigation of the effects of the other bacterial species used throughout this project on CRT. RG/C2 cells exhibited increased TOP/FOP upon bacterial infection, with increased CRT observed relative to the control in cells infected with *B. fragilis* and *F. nucleatum* (+20.6% and +39.55 respectively). CRT was almost doubled in cells infected with *E. coli* Nissle (+98.2%) and *E. faecalis* (+93.2%), however due to the large variation seen in TOP/FOP ratios these results were not statistically significant. In contrast, bacterial infection of HCT116 cells resulted in a non-significant decrease in TOP/FOP, which ranged between -10% and -13% for all species.
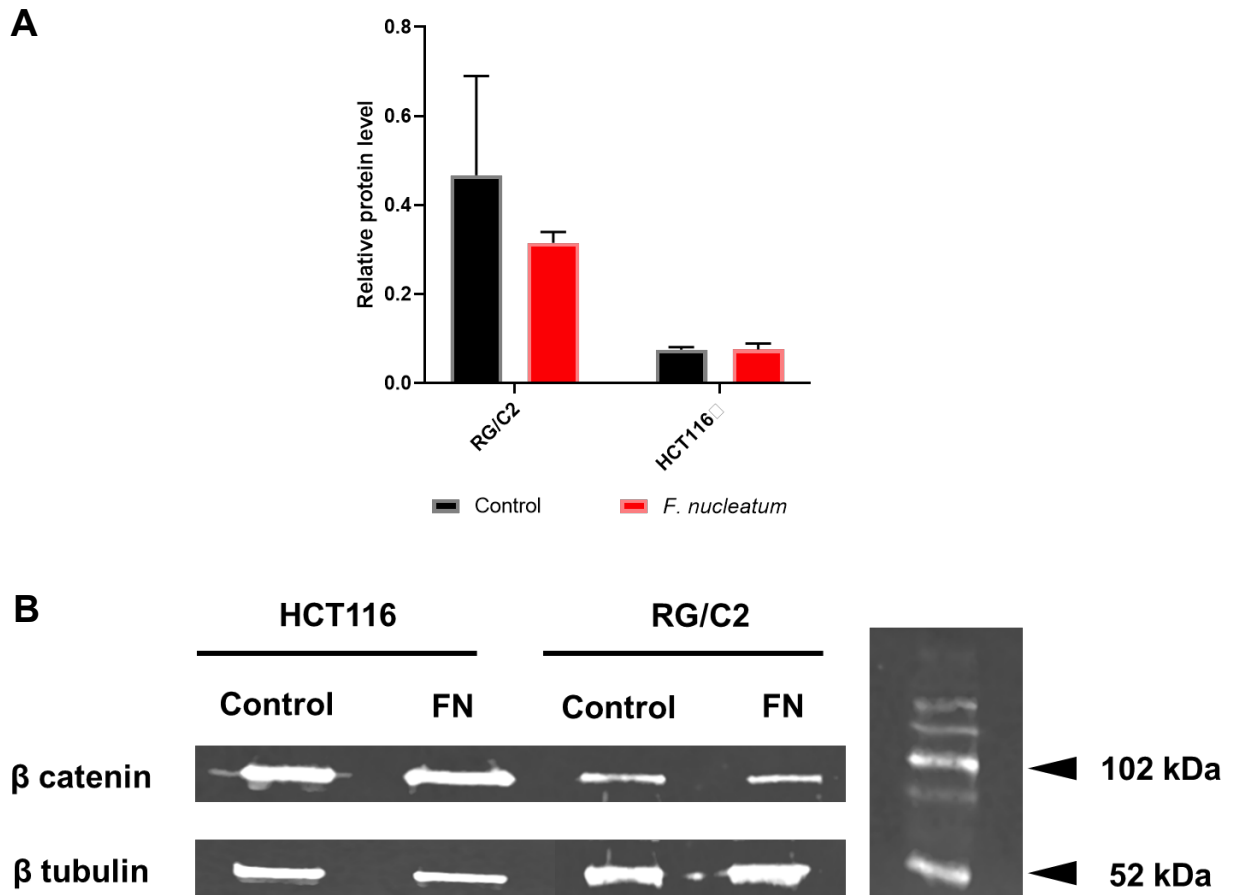
Figure 5.6: Analysis of protein expression in RG/C2 and HCT116 cells after *F. nucleatum* infection. (A) After Western blotting and SDS PAGE, densitometry was performed to quantify relative levels of β-catenin protein present in infected cells. Error bars represent standard deviation and results of three independent experiments performed in triplicate. (B) Representative image of β-catenin and β-tubulin bands following SDS PAGE used to quantify protein expression.
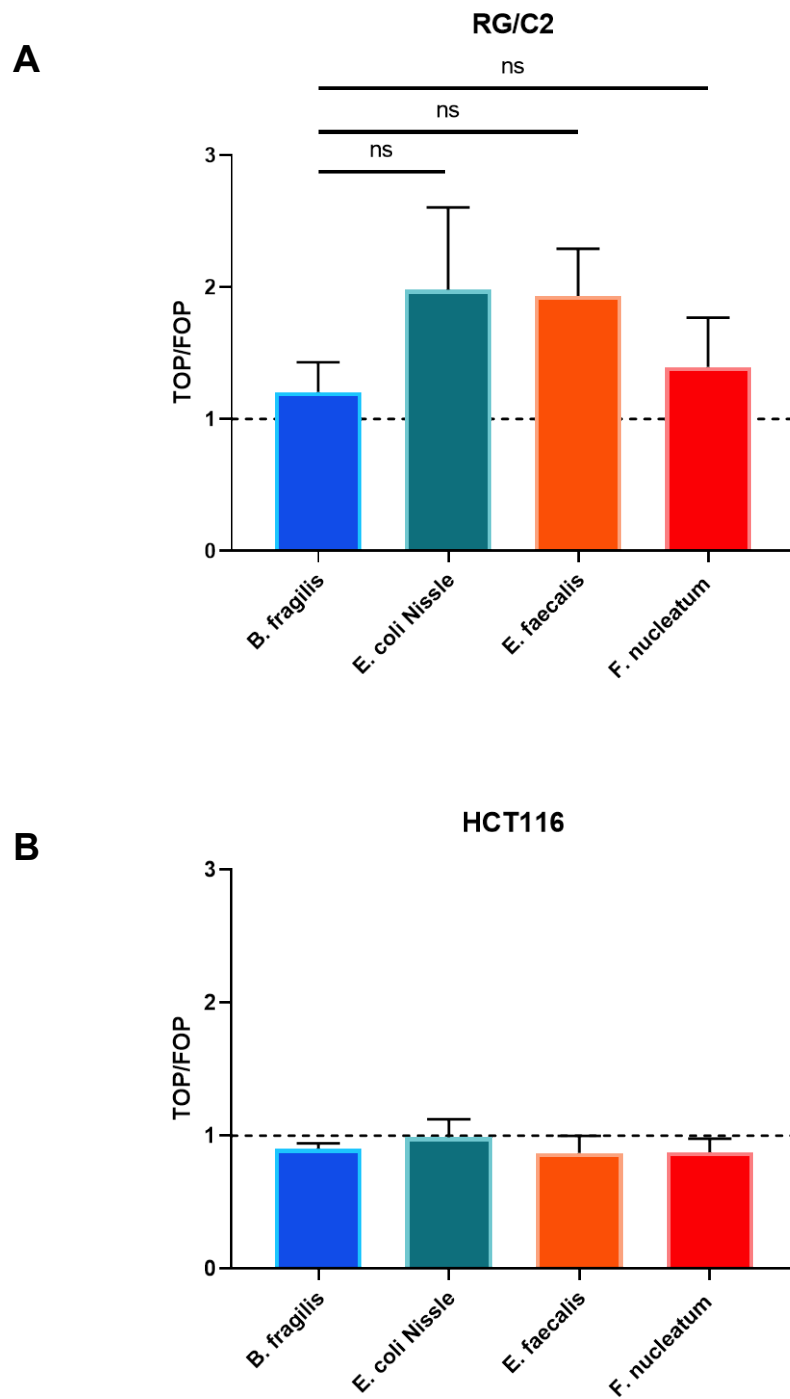
Figure 5.7: TOPFLASH luciferase activity in RG/C2 and HCT116 cells after bacterial infection. Fold change in TOPFLASH/FOPFLASH ratio after normalisation using Renilla luminescence relative to non-infected tumour cells. Despite large increases in TOP/FOP in RG/C2 cells, the effect was not statistically significant. Error bars represent standard deviation and results of three independent experiments performed in triplicate.

5.4.5 Regulation of cell motility genes upon infection of HCT116 cells with *F. nucleatum*

To investigate the underlying mechanism responsible for the induction of tumour cell migration and invasion by *F. nucleatum* infection, qPCR microarrays were performed analysing the expression of 84 cell motility-related genes (Supplementary Table 1). The difference in C_T (the point at which fluorescence for a particular gene crosses the threshold value) between untreated and *F. nucleatum*-infected cells was compared, and the fold regulation of each gene is presented in Figure 5.8. A total of 19 genes were upregulated by a factor of 1.3 or more, indicating an increased concentration of mRNA in infected cells. Three genes were downregulated to the same degree. The largest increase was seen in the expression of *Supervillin* (+21.11), with the largest decrease seen in *Integrin beta 2* (-8.13).

In order to further investigate the impact of the modulation of these genes on cell motility, the 10 genes for which the greatest change in expression were detected were cross-referenced with current literature on CRC motility (Table 5.1). This included the endogenous control gene *ACTB*, for which a fold regulation of +8.31 was recorded. The increased expression of this control gene did not interfere with data analysis, as the PCR microarrays used in this study contain a panel of 6 housekeeping genes allowing for normalisation to be performed on a stably expressed gene.

Where no available literature on the expression of these genes in CRC was available, research investigating the role of these genes in other cancers was considered. Seven of these changes were associated with a tumour cell motility-promoting effect. In contrast, alterations in the expression of two genes, *ITGB2* (-8.13) and *WIPF1* (3.23), were considered to have negative effects on cell motility. Parallel assays were performed in RG/C2 cells, with *F. nucleatum* regulating gene expression

184

differently in RG/C2 and HCT116 cell lines (Supplementary Figure 4). The impact of *F. nucleatum* infection on the 10 genes described in Table 5.1 was compared in both cell lines (Figure 5.9). *ACTB*, *ITGB3*, *MMP2*, *RAC2* and *SVIL* were all negatively regulated by *F. nucleatum* infection in RG/C2 cells, despite being upregulated in HCT116 cells, highlighting a difference in response by these cell lines. *ACTN3*, *FAP*, *ITGB2* and *WIPF1* were regulated in the same manner in both cell lines, however a greater degree of change was seen in infected HCT116 cells when compared with RG/C2 cells, suggesting that the gene expression profile of this cell line is more sensitive to influence by bacterial infection.

Gene	Value	Gene	Value	Gene	Value	Gene	Value	Gene	Value	Gene	Value
ACTN1	1.15	ACTN3	2.17	ACTN4	1.15	ACTR2	1.01	ACTR3	1.11	AKT1	1.05
ARF6	1.13	ARHGDIA	-1.01	ARHGEF7	1.36	BAIAP2	1.27	BCAR1	1.25	CAPN1	1.35
CAPN2	1.19	CAV1	1.14	CDC42	-1.0	CFL1	1.12	CRK	-1.22	CSF1	-1.21
CTTN	1.35	DIAPH1	1.07	DPP4	1.39	EGF	1.17	EGFR	1.26	ENAH	1.21
EZR	1.21	FAP	5.1	FGF2	1.46	HGF	1.0	IGF1	2.74	IGF1R	1.0
ILK	1.12	ITGA4	1.0	ITGB1	1.34	ITGB2	-8.13	ITGB3	2.61	LIMK1	1.1
MAPK1	1.24	MET	1.27	MMP14	1.19	MMP2	4.76	MMP9	1.09	MSN	1.11
MYH10	1.13	MYH9	1.45	MYL9	1.07	MYLK	1.5	PAK1	-1.01	PAK4	1.06
PFN1	1.27	PIK3CA	1.07	PLAUR	1.35	PLCG1	1.18	PLD1	1.03	PRKCA	1.18
PTEN	1.04	PTK2	1.15	PTK2B	1.12	PTPN1	1.27	PXN	1.03	RAC1	-1.03
RAC2	2.27	RASA1	1.08	RDX	-1.3	RHO	1.0	RHOA	-1.05	RHOB	1.23
RHOC	1.16	RND3	1.53	ROCK1	1.25	SH3PXD2	1.31	SRC	-1.01	STAT3	1.24
SVIL	21.11	TGFB1	1.06	TIMP2	-1.01	TLN1	1.15	VASP	1.16	VCL	1.07
VEGFA	1.09	VIM	1.49	WASF1	1.26	WASF2	1.23	WASL	1.49	WIPF1	-3.23

Fold regulation

| -10 | -5 | 0 | 5 | 10 | 10+ |

Figure 5.8: Impact of *F. nucleatum* infection on cell motility gene expression in HCT116 cells. Threshold values for each gene obtained from qPCR microarrays performed on *F. nucleatum*-infected cells were normalised to un-infected cells. Colour intensity indicates the relative fold regulation for each gene. Positive regulations are denoted in green, whereas negative regulations are displayed in red.

Table 5.1: Literature review of cell motility genes modified by infection of HCT116 cells with *F. nucleatum.*

Gene	Description	Fold change	Net effect on motility	Literature
ACTB – Beta actin	Actin cytoskeletal protein which regulates cell structure, integrity, and motility.	+8.31	Positive	Increased expression correlates with cell motility in LS180 colorectal adenocarcinoma cell line variants (Nowak *et al.*, 2005). Essential for bleb formation in colon tumour cells (Simiczyjew *et al.*, 2017).
ACTN3 – Alpha-actinin 3	Actin-binding protein exclusively expressed in type-II muscle fibres.	+2.17	Unclear	Limited literature on *ACTN3* in cancer. Increased expression of ubiquitously expressed ACTN proteins e.g. *ACTN4* enhances colorectal tumour invasion by suppressing focal adhesions (Fukumoto *et al.*, 2015).
FAP – Fibroblast activation protein	Cell surface serine protease.	+5.10	Positive	High stromal levels of *FAP* in CRC linked to advanced disease and poor prognosis (Henry *et al.*, 2007). Promotes angiogenesis via ERK signalling *(Cao et al.*, 2018a).
IGF1 – Insulin-like growth factor 1	Anabolic growth hormone produced in liver tissue.	+2.74	Positive	Increased levels of circulating IGF1 linked to future CRC development (Ma *et al.*, 1999). Promotes invasion and apoptosis resistance in CRC through Akt/Bcl-x(L) pathway (Sekharam *et al.*, 2003).
ITGB2 – Integrin beta-2	Integral cell surface protein involved in ECM binding.	-8.13	Negative	Reduction in *ITGB2* is associated with reduced migration, and specifically reduced hepatic metastasis (Benedicto *et al.*, 2017).
ITGB3 – Integrin beta-3	Integral cell surface protein involved in ECM binding.	+2.61	Positive	Expression induced by ROS increases migration and invasion in SW480 and SW620 cells (Lei *et al.*, 2011). Increased expression in CRC leading to enhanced invasion after loss of MiR-30a-5p (Wei *et al.*, 2016).
MMP2 – Matrix metalloproteinase 2	Secreted enzyme responsible for ECM degradation.	+4.76	Positive	Increased activity in CRC, which is associated with invasion depth, lymph node involvement and Duke's stage (Li *et al.*, 2005). Activated by Twist1/2 binding to *MMP2* promoter (Lu, Dong and Fan, 2018). Effective prognostic marker for overall survival, recurrence and distant metastasis in breast cancer patients (Jiang and Li, 2021).
RAC2 – Rac family small GTPase 2	Cell membrane GTPase.	+2.27	Positive	Expression of Rho GTPases is associated with a reduction in cell-cell adhesion, and an increase in

				migration and invasion (Leve and Morgado-Díaz, 2012).
SVIL – **Supervillin**	Cell membrane-associated F-actin binding protein associated with actin punctae.	**+21.11**	**Positive**	Reorganises actin cytoskeleton to facilitate invadopodia formation in COS-7 cells (Crowley *et al.*, 2009). Promotes EMT and metastasis in hepatocellular carcinoma under hypoxic conditions (Chen *et al.*, 2018) and promotes angiogenesis in malignant disease *(C. Zhao et al>*, 2020).
WIPF1 – **WAS/WASL-interacting protein family 1**	Regulates cytoskeletal organisation and actin polymerisation.	**-3.23**	**Negative**	Sequesters the β-catenin destruction complex in glioblastoma and promotes tumour proliferation and stemness(Gargini *et al.*, 2016). Promotes disease aggressiveness in BRAF V600E thyroid cancer (Zhang *et al.*, 2017).
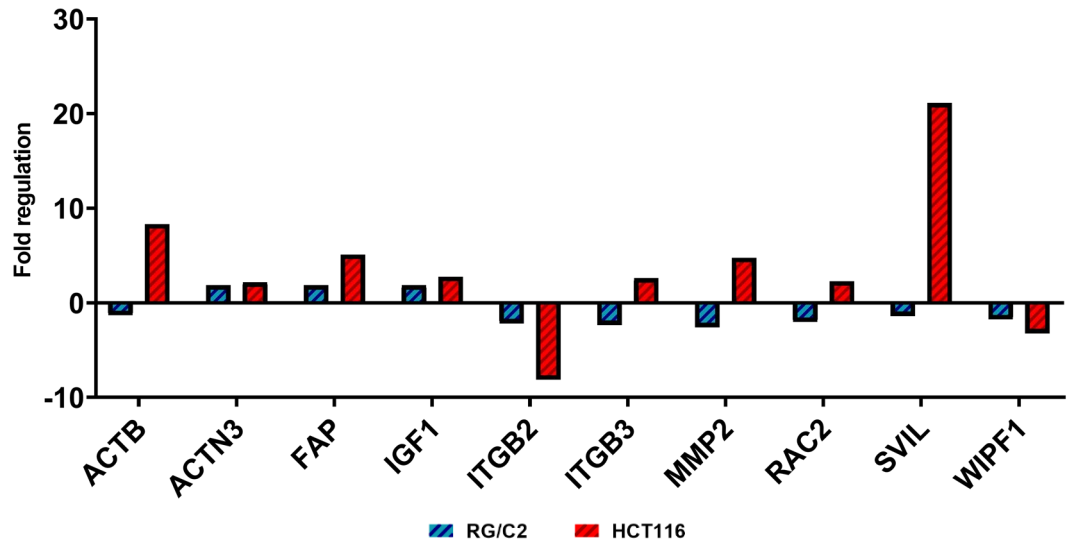
Figure 5.9: Comparison of regulation of cell motility genes by *F. nucleatum* in RG/C2 and HCT116 cells. Parallel qPCR arrays were carried out in RG/C2 and HCT116 cells infected with *F. nucleatum*. The genes which were most affected by *F. nucleatum* infection in HCT116 cells were compared with RG/C2 experiments.

5.4.6 Regulation of EMT-associated genes upon infection of HCT116 cells by *F. nucleatum*

A qPCR microarray was also performed to investigate the regulation of genes associated with EMT (Supplementary Table 2). In total, 9 genes were upregulated by a factor of 1.3 or greater, compared to 8 genes which were similarly downregulated (Figure 5.10). The largest increase in expression was observed in *IGFBP4* (+19.51), compared to *COL1A2* which exhibited the largest decrease in expression (-6.08). Five of the observed genetic changes were associated with a net increase in EMT, compared with 4 changes associated with a downregulation of EMT (Table 5.2). The literature surrounding the role of *SOX10* in CRC seems to implicate it as a negative regulator of EMT, despite being positively associated with EMT in other diseases. Once again, parallel assays were performed using RG/C2 cells (Supplementary Figure 5). When comparing these changes in gene expression with *F. nucleatum*-infected RG/C2 cells, *BMP2*, *COLIA2*, *GSC*, *TFPI2* and *WNT5A* all displayed similar patterns of regulation by *F. nucleatum* infection in both cell lines (Figure 5.11). However, the intensity of the regulation differed between cell lines. In comparison, *GNG11*, *IGFBP4* and *SOX10* displayed increased expression in HCT116 cells and decreased expression in RG/C2 cells. *WNT11* was increased upon infection of RG/C2 cells but was decreased in infected HCT116 cells. Finally, the fold regulation of *BMP2*, *MMP3*, *TFPI2* and *WNT5A* was greater in infected RG/C2 cells.

Figure 5.10: Impact of *F. nucleatum* infection on EMT gene expression in HCT116 cells. Threshold values for each gene obtained from qPCR microarrays performed on *F. nucleatum*-infected cells were normalised to un-infected cells. Colour intensity indicates the relative fold regulation for each gene. Positive regulations are denoted in green, whereas negative regulations are displayed in red.

191

Table 5.2: Literature review of EMT genes modified by infection of HCT116 cells with *F. nucleatum*.

Gene	Description	Fold change	Net effect on EMT	Literature
BMP2 – Bone morphogenic protein 2	Ligand for TGF-β and hedgehog signalling pathways.	+1.70	Positive	Induces stemness and EMT in HCT116 and SW480 cells through *STAT3* activation (Kim *et al.*, 2015). Facilitates SNAIL1 signalling in colon and mammary tumour cells (Frey *et al.*, 2020).
COL1A2 – Collagen type I alpha chain 2	Required for production of type I collagen.	-6.08	Negative	Expression increased during upregulation of *SNAIL1* in DLD-1 cells (Tanaka *et al.*, 2016). Loss of BMP1 in SW620 cells increases expression of COL1A1 and COL1A2 leading to EMT (Zhu *et al.*, 2019).
GNG11 – Guanine nucleotide binding protein subunit 11	Lipid-anchored cell membrane protein.	+2.60	Positive	Promotes survival in circulating colon and pancreatic tumour cells (Yadavalli *et al.*, 2017).
GSC – Goosecoid homeobox	Transcription factor involved in organogenesis.	+4.48	Positive	Expression in hepatocellular carcinoma promotes invasion and EMT (Xue *et al.*, 2014).
IGFBP4 – Insulin-like growth factor binding protein 4	Circulating insulin-binding protein.	+19.51	Negative	Increased apoptosis in HT29 cells through Bcl-2/Bax dependent mechanism (Durai *et al.*, 2007).
MMP3 – Matrix metalloproteinase 3	Enzyme involved in degradation of ECM.	+2.02	Positive	High expression of *MMP3* leads to formation of invasive, mesenchymal tumours in breast cancer (Sternlicht *et al.*, 1999). Stimulates EMT through Rac1b, and causes cell spreading in mouse mammary epithelial cells (Nelson *et al.*, 2008).
SOX10 – Transcription factor SOX-10	Transcription factor implicated in embryonic development.	+2.09	Unclear	*SOX10* expression has been demonstrated to inhibit Wnt/β-catenin signalling in a range of digestive cancer cell lines, including HCT116 (Tong *et al.*, 2014). However, found to promote EMT by inducing a mesenchymal phenotype in breast (Dravis *et al.*, 2015) and nasopharyngeal (He and Jin, 2018) cancers.
TFPI2 – Tissue factor pathway inhibitor 2	Serine proteinase inhibitor and tumour suppressor.	+4.19	Negative	Induces differentiation in hepatocellular carcinoma (Li *et al.*, 2017). Inhibits progression of breast cancer through *TWIST1* expression (D. Zhao *et al.*, 2020).

Gene	Function	Value	Direction	Notes
				Silenced through methylation during CRC progression (Jung *et al.*, 2020).
WNT5a – **Wnt family member 5a**	Wnt family signalling molecule, ligand for Frizzled receptors.	**-2.50**	**Positive**	Antagonises canonical Wnt signalling in CRC, inhibiting proliferation and EMT (Cheng *et al.*, 2014). Lower expression in metastatic SW620 than SW480 cells due to silencing through histone modifications (Li and Chen, 2012). Potentially different activity depending on differential expression of mRNA isoforms, with the Wnt5a-short being associated with invasion by HCT116 and tumour depth, whereas Wnt5a-Long suppressed β-catenin expression (Huang *et al.*, 2017).
WNT11 – **Wnt family member 11**	Wnt family signalling molecule.	**-1.63**	**Negative**	High *Wnt11* expression is associated with increased invasion and 5-year mortality in a CRC patient cohort (Gorroño-Etxebarria *et al.*, 2019). AGR2 signalling promotes EMT and tumour invasion through Wnt11 in CRC cell lines including HCT116. HT29 and SW480 (Tian *et al.*, 2018).
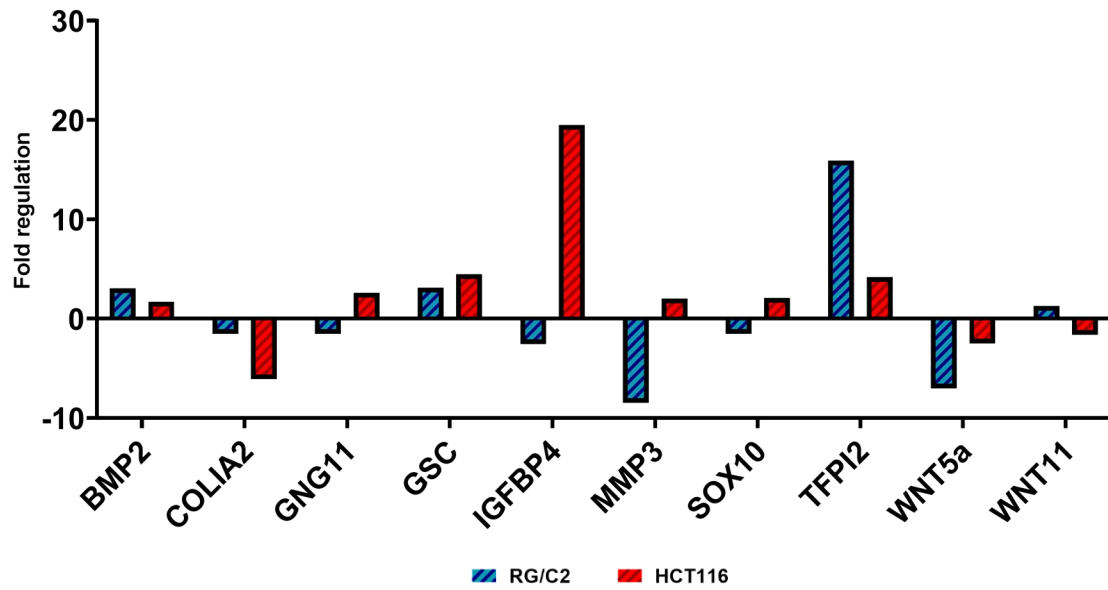
Figure 5.11: Comparison of regulation of EMT genes by *F. nucleatum* in RG/C2 and HCT116 cells. Parallel qPCR arrays were carried out in RG/C2 and HCT116 cells infected with *F. nucleatum*. The genes which were most affected by *F. nucleatum* infection in HCT116 cells were compared with RG/C2 experiments.

5.5 Discussion

The previous chapter demonstrated that infection with the four bacterial species tested, *B. fragilis*, *E. coli* Nissle, *E. faecalis* and *F. nucleatum* increased tumour cell yield after a relatively short infection period in both RG/C2 and HCT116 cell lines. This represents one clear means by which these species may contribute towards tumour progression *in vivo*. However, perhaps the most significant event in carcinogenesis is the development of migratory and invasive tumour characteristics; as patients with metastatic disease face significantly lower survival prospects and an increased likelihood of disease recurrence (Yu *et al.*, 2019). As previously discussed, development of metastatic CRC from healthy epithelium normally occurs over several decades (Vogelstein and Kinzler, 1993). However, the predicted increase in global CRC incidence, coupled with the increased rates of CRC detected in younger populations in developed countries, suggests that the burden of metastatic CRC is set to rise (Rawla, Sunkara and Barsouk, 2019). It is therefore imperative that the contribution of microorganisms to colorectal tumour invasion and metastasis is fully understood in order to allow for the development of targeted treatments and improved diagnostics.

This chapter begins by assessing the impact of bacterial infection on tumour migration using the transwell filter migration assay. This assay measures single cell migration, which in tumour cells can be classified into two types: amoeboid or mesenchymal (Friedl and Wolf, 2003). There are notable differences between these two forms of migration. Amoeboid migration is characterised by a rounded cell body shape, cell blebbing and weak adhesions to the substrate (Lintz, Muñoz and Reinhart-King, 2017). In contrast, cells adopting a mesenchymal phenotype polarise to form a leading and trailing edge, which involves the formation of stress fibres (Hecht *et al.*,

2015). Migratory RG/C2 cells primarily displayed a rounded morphology indicative of amoeboid migration (Figure 5.1). In contrast, both amoeboid and mesenchymal-like HCT116 cells were observed, and the frequency of mesenchymal-like cells appeared to increase in *F. nucleatum*-infected cells (Figure 5.2F). Carcinoma cell lines are able to switch between these two forms of motility as required, in a Rho-signalling dependent manner (Paňková *et al.*, 2010), and the results presented in this study may suggest that *F. nucleatum* infection specifically promotes mesenchymal migration. This hypothesis is supported by the cell motility array data which showed that *F. nucleatum* infection increased the expression of *Rac2*, a Rho GTPase*, by a factor of +2.23.

When analysing the number of migrated cells observed for each cell line, it is clear that RG/C2 cells are far less migratory than HCT116 cells. The average number of migrated cells in the RG/C2 control group was 17, which is consistent with previous studies which have quantified migration in this cell line using the same approach (Qualtrough *et al.*, 2007). This is in stark contrast to HCT116 cells, where an average of 91 cells were counted in the control group, despite initially being seeded at half the density. Direct comparisons between cell lines in transwell filter assays are difficult, as each cell line will respond differently to the assay parameters. These include the use of Ca^{2+}-free medium and relatively low seeding densities, reducing cell-cell contact. Normal epithelial cells undergo anoikis upon loss of ECM attachment and the breakdown of adherens junctions with neighbouring cells (Bretland, Lawry and Sharrard, 2001). Therefore, acquisition of anoikis resistance is essential for cells to become metastatic, as reviewed by Kim *et al.* (2012). As RG/C2 cells were originally isolated from a benign adenoma, they are more closely related to normal epithelium than the adenocarcinoma cell line HCT116. It could therefore be suggested that anoikis

may negatively impact RG/C2 cell migration in the transwell filter assay, which should be considered when comparing migration between these two cell lines.

Upon bacterial treatment of RG/C2 and HCT116 cells, these assays revealed that two of the four bacterial species tested, *E. coli* Nissle and *F. nucleatum*, were able to promote tumour migration. However, effect was only observed in HCT116 cells, with *E. coli* Nissle and *F. nucleatum* increasing migration by 42% and 91%, respectively. After extensive review of the literature, no previous example of the use of pre-infected cells within a transwell filter migration assay could be found. Therefore, in order to verify the observed effects on tumour migration, and examine whether the migration-promoting effects of *E. coli* Nissle and *F. nucleatum* extended to collective cell migration, wound scratch assays were also performed. It is also important to consider the impact of bacteria on collective tumour cell migration as it has been suggested that collective cell migration and the entry of 'circulating tumour microemboli' into the vasculature is linked to poorer patient prognosis when compared to individual circulating tumour cells (Hou *et al.*, 2012).

Wound scratch assays also demonstrated that *E. coli* Nissle and *F. nucleatum* infection promotes motility in HCT116 cells. In contrast to the transwell filter migration assays, *E. coli* Nissle infection had a greater effect on migration in the wound scratch assay than *F. nucleatum* (+86% compared with +63%). RG/C2 cells experienced a decrease in migratory ability when treated with bacteria in both the transwell filter migration and wound scratch assays. *F. nucleatum* reduced RG/C2 wound healing by 51%, although a large degree of variation was observed and therefore this was not statistically significant. This could possibly be attributed to varying cell depth of RG/C2 cultures which could influence cell migration. The lack of migration promotion in RG/C2 cells demonstrates that bacteria, in particular

F. nucleatum, can promote a migratory phenotype in HCT116 cells, but crucially cannot confer this phenotype to RG/C2 cells. This suggests that the mechanism of migration promotion may be specific to adenocarcinoma cells, as although RG/C2 cells display lower basal levels of migration, several studies have demonstrated that their motility can be increased using other methods e.g. prostaglandin $F_{2\alpha}$ and fascin over-expression (Qualtrough *et al.*, 2007, 2009). When considering the Driver-Passenger model of bacterial contribution to CRC (Tjalsma *et al.*, 2012), this potentially implicates *F. nucleatum* as a driver of late-stage metastatic disease, as it may require the tumour to have already acquired aggressive characteristics to exert its effects on migration.

Unlike the effect on tumour cell yield observed in tumour cells infected with bacteria, the promotion of migration in HCT116 cells was species-specific. The increase in migration seen in *E. coli* Nissle-infected cells is surprising, as this species is often administered as a probiotic for intestinal disorders such as ulcerative colitis (Kruis *et al.*, 2004). It should be considered that motility in the colonic epithelium is necessary for efficient cell turnover, with colorectal tumours initially displaying a reduction in motility after *APC* loss (Näthke *et al.*, 1996). Therefore, the combination of the anti-inflammatory and pro-motility effects of *E. coli* Nissle may be necessary to stimulate the wound healing required for remission in ulcerative colitis patients, and protect the colon from tumour formation. However, pathogenicity of this probiotic strain has previously been reported. In 2010, a case of severe sepsis was reported in an infant administered with *E. coli* Nissle, with subsequent blood cultures identifying this species as the causative agent (Guenther *et al.*, 2010). A study published in the same year found that *E. coli* Nissle pathogenicity was enhanced where microbiota dysbiosis and perturbation of the immune system occurred, as is often the case in CRC

patients (Gronbach *et al.*, 2010; Gonzalez, Hagerling and Werb, 2018). The effect of *E. coli* Nissle on adenocarcinoma motility demonstrated in this chapter, coupled with its ability to invade epithelial cells, and promote tumour growth highlighted in previous chapters suggests that greater consideration should be given to its use as a probiotic in immunocompromised individuals and those at high risk of CRC development. As *E. coli* Nissle is commonly administered to IBD patients, an at-risk group for future CRC development, longitudinal studies may be necessary to determine the long-term effects of this treatment on CRC risk.

In the case of *F. nucleatum*, a role in in colorectal tumour motility has previously been suggested. *F. nucleatum* strains can be detected in both primary CRC tumours and their distant metastases, suggesting some association between *F. nucleatum* colonisation and tumour dissemination (Bullman *et al.*, 2017). In this project, it was also demonstrated that *F. nucleatum* is able to promote invasion in HCT116 cells, a trait which was not shared with *E. coli* Nissle. In contrast to cell migration, invasion requires the degradation and remodelling of the ECM (Wolf *et al.*, 2007). The results of the PCR microarrays revealed that several genes associated with invasion were upregulated by *F. nucleatum* infection in HCT116 cells. The most upregulated gene was *Supervillin* (*SVIL*), a member of the gelsolin family of cytoskeletal organising proteins (Silacci *et al.*, 2004). Reorganisation of the cytoskeleton is essential for cell motility, and the upregulation of both *ACTN3* and *ACTB* supports the notion that *F. nucleatum* is able to influence this process. This is supported by literature demonstrating that *F. nucleatum* infection induces the formation of new actin filaments during its intracellular replication (Gursoy, Könönen and Uitto, 2008). Although no evidence for intracellular replication was observed during the investigation of bacterial persistence (see Section 3.4.6), it is possible that

bacterial replication occurred before the initial 24-hour time point, or that rates of bacterial replication and cell death were equivalent, resulting in no observable change in colony forming units.

SVIL specifically reorganises lamellipodial proteins, particularly cortactin, to co-localise with proteins required for the formation of invadopodia (Crowley *et al.*, 2009). At the tip of invadopodia, enzymes such as matrix metalloproteinases (MMPs) are expressed which digest the surrounding ECM to facilitate invasion (Jacob and Prekeris, 2015). Upon *F. nucleatum* infection of HCT116 cells, both *MMP2* and *MMP3*. Interestingly, *MMP2* is dependent upon cortactin organisation for its vesicular transport to invadopodia, potentially highlighting a synergistic effect between *SVIL* and *MMP2* expression after *F. nucleatum* infection which increases invasive potential (Clark and Weaver, 2008). It should also be considered that the gentamicin protection assay data presented in Chapter 3 illustrated that *F. nucleatum* infection does not occur in all tumour cells present in the monolayer. Therefore, it is unlikely that each cell showing aberrant migration in the migration assays presented in this chapter are infected with bacteria, and thus, a bystander effect on neighbouring cells by infected cells may also occur to promote migration.

In addition to reorganisation of the cytoskeleton, tumour migration and invasion is often characterised by epithelial-mesenchymal transition (EMT). This process occurs in normal development and wound healing responses, but is dysregulated in many cancers including CRC, as reviewed by Brabletz *et al.* (2018). *F. nucleatum* has been implicated in the promotion of EMT across a number of cancers and through a wide range mechanisms. For example, in mouse models, *F. nucleatum* was demonstrated to colonise and promote the metastasis of breast tumours through its Fap2 surface protein (Parhi *et al.*, 2020). A study investigating the effect of

F. nucleatum on oral squamous cell carcinoma cells found that *F. nucleatum* regulated the expression of multiple EMT-associated genes including *Snail-1*, *Slug*, *Vimentin* and E-cadherin (*CDH1*). In this study these genes were not found to be affected by *F. nucleatum* using PCR microarrays. However, *BMP2*, a regulator of *Snail-1* signalling was found to be increased by 1.7-fold in HCT116 cells following *F. nucleatum* infection.

The net effect on EMT was unclear, as expression of EMT promoters and antagonists was increased upon *F. nucleatum*. One of the most important regulators of EMT in both human development and cancer is the Wnt signalling pathway (Clevers, 2006). *F. nucleatum* is considered to be a regulator of this pathway through the binding of its FadA adhesin to the adherens junction protein E-cadherin during bacterial invasion (Rubinstein *et al.*, 2013). In this study, *F. nucleatum* did not have a large effect on E-cadherin at the transcription level (as evidenced by the -1.18 fold change in expression of *CDH1*), although infection did decrease the expression of both *Wnt5a* and *Wnt11* which are considered by antagonists of canonical *Wnt* signalling (Bisson *et al.*, 2015). The effect of *F. nucleatum* infection on *Wnt* activity was further investigated by analysing the activity of the *Wnt*-regulating protein β-catenin. Surprisingly, and contrary to current literature, a small non-significant reduction in β-catenin protein level was detected by SDS-PAGE. In addition, no effect on CRT was observed in *F. nucleatum*-infected HCT116 cells via the TOPFLASH assay, although a significant increase in FOPFLASH activity was observed. The FOPFLASH plasmid contains mutated TCF binding sites, and therefore only yields luciferase activity from its minimal c-fos promoter (Molenaar *et al.*, 1996). During analysis of TOPFLASH assay data, the ratio of TOP/FOP is used to calculate the relative level of CRT. Therefore, an increase in FOPFLASH attributable to upregulation of the c-fos

promoter decreases the TOP/FOP ratio. The reduction in TOP/FOP observed in *F. nucleatum*-infected HCT116 cells may therefore be caused by an upregulation of c-fos, which has emerged as a regulator of EMT in other cancers such as head and neck squamous cell carcinoma (Muhammad *et al.*, 2017). Upregulation of this promoter may represent a novel mechanism through which *F. nucleatum* regulates EMT, although further investigation is clearly required.

The absence of an effect on CRT in HCT116 cells after infection with *F. nucleatum* is in direct contrast to an aforementioned paper by Rubinstein *et al.* which demonstrated a greater than 4-fold increase in gross TOPFLASH activity, with no effect on FOPFLASH in HCT116 cells treated with *F. nucleatum* (Rubinstein *et al.*, 2013). Their study reported a similar upregulation of TOPFLASH activity in multiple CRC cell lines, which was suggested to be a result of FadA-mediated E-cadherin cleavage. However, previous studies have shown that the adenocarcinoma cell lines used in this study possess vastly different levels of E-cadherin expression and would therefore not be expected to a show a comparable response (Qualtrough *et al.*, 2015). In addition, if E-cadherin cleavage was the primary mechanism, an effect on the motility of benign cells which have higher E-cadherin expression compared to malignant cells would be expected. Furthermore, a difference in the promotion of cell yield demonstrated in Chapter 4 between RG/C2 and HCT116 cells would also be expected due to their significantly different levels of E-cadherin (Rees & Qualtrough, unpublished data).

A reduction in β-catenin expression upon infection by *F. nucleatum* has been demonstrated using PCR in oral squamous cell carcinoma, suggesting that the interaction between *F. nucleatum* and Wnt signalling may be highly variable (Abdulkareem *et al.*, 2018). This variation may be attributable to experimental

conditions. For example, the Rubinstein group treated their cells at an MOI of 1000:1 for a 2-hour period, in contrast to the MOI of 10:1 and 4-hour infection used in this study. Finally, when introducing the Driver-Passenger model, Tjalsma *et al*. posited that the ability of bacteria to drive tumorigenesis may be highly dependent upon strain-specific virulence, and therefore this must be considered when comparing results across research groups (Tjalsma *et al*., 2012).

The differences in the effect of *F. nucleatum* infection on RG/C2 and HCT116 motility are reflected in the TOPFLASH and PCR results. In RG/C2 cells, infection with all four species of bacteria tested increased TOPFLASH activity. The divergent effects of bacterial infection on TOPFLASH activity in RG/C2 and HCT116 cells could be attributed to the Wnt signalling-specific mutations in these two cell lines. Although both are wild type for *APC*, the HCT116 cell line contains an activating β-catenin mutation, whereas the RG/C2 cell line possess a heterozygous deletion in the β-catenin gene *CTNNB1* (Greenhough *et al*., 2010; Dallosso *et al*., 2012). It should be noted that increase in TOPFLASH activity upon bacterial infection in RG/C2 cells was not statistically significant, and no increases in *CTNNB1* gene expression were detected upon *F. nucleatum* infection in either cell line using PCR microarrays. However, the increase in RG/C2 TOPFLASH activity is supported by the change in *Wnt5a* expression seen in *F. nucleatum*-infected cells, where a decrease in expression greater than the decrease seen in HCT116 cells was detected (Figure 5.10).

When analysing PCR microarray data, a general trend was observed whereby the net effect of *F. nucleatum* infection on cell motility and EMT appears to be promotive and support the migration and invasion data presented in this study. In the cell motility array, only a downregulation of *ITGB2* and *WIPF1* could be associated with reduced motility. However, many of the other genes influenced by *F. nucleatum*

supported a role of bacterial infection in the promotion of migration, with plausible mechanisms for their activation by bacteria. For example, *F. nucleatum* infection increased *ITGB3* expression. This gene has also been demonstrated to be induced through radical oxygen species (ROS), which are commonly produced by immune cells during bacterial infection, and specifically produced by some microorganisms such as *E. faecalis* (Huycke, Abrams and Moore, 2002; Lei *et al.*, 2011).

The microarrays presented a less clear picture of EMT regulation by *F. nucleatum*. The change in expression of four genes, *COL1A2*, *IGFBP4*, *TFPI2* and the previously discussed *Wnt11* were associated with a reduction in EMT. Interestingly, *IGFBP4* expression, which was increased +19.51 fold upon *F. nucleatum* infection, has been demonstrated to promote apoptosis in the HT29 cells (Durai *et al.*, 2007). Microarray data shows that despite *IGFBP4* expression being significantly increased in *F. nucleatum*-infected HCT116 cells, it is decreased in infected RG/C2 cells by a factor of -2.55 (Figure 5.10). Differential regulation of this gene in these cell lines could potentially explain the divergent effect on apoptosis seen in *F. nucleatum*-infected cells presented in Chapter 4. There are conflicting reports surrounding the role of *SOX10* in EMT, which is thought to inhibit EMT in colorectal tumour cells (including HCT116) but promote EMT in other cancers (see Table 5.1). Further investigation is therefore required to elucidate the true role of *SOX10* in EMT. In the study described in Table 5.1, the authors investigated the effect of ectopic *SOX10* expression in isolation, and it is therefore possible the effect of *SOX10* expression on colorectal tumour EMT is context-dependent (Tong *et al.*, 2014).

Despite being studied in other cancers, several of these genes identified as having expression modified by *F. nucleatum* infection have not been fully studied in CRC, including *GSC*, *MMP3* and *WIPF1*. The PCR microarrays performed in this

study have therefore succeeded in identifying potentially novel pathways which may implicate *F. nucleatum* in the promotion of invasion and metastasis in CRC. Due to the high cost of performing these arrays, it was only possible to perform a single replicate, meaning that caution must be used when attempting to draw conclusions from the data. However, the Qiagen PCR microarrays performed in this study utilise internal controls for genomic DNA contamination, PCR reproducibility and reverse transcription efficiency, which improve confidence in the dataset. It should also be considered that these assays are ultimately measuring changes in mRNA levels extracted from treated cells, which do not necessarily correspond to a change in translated protein (Liu, Beyer and Aebersold, 2016). To distinctly prove that *F. nucleatum* expression produces a functional change in protein level, as suggested by the alterations in gene expression presented in this study, there are several assays which should be performed. Firstly, analysis of protein concentration via Western blotting would confirm that any change in mRNA level is translated into a change in protein expression. For any proteins whose production is found to be altered by *F. nucleatum* infection, congenic knock-down cell lines should be produced, and migration and invasion assays repeated to definitely show that these genes regulate HCT116 motility and invasion in response to bacterial infection.

In summary, the results presented in this chapter demonstrate that bacteria, specifically *F. nucleatum*, are able to promote migratory and invasive behaviour in malignant CRC cells. This aligns with other published studies which have drawn similar conclusions, but crucially, this study suggests that the bacteria tested cannot directly promote the acquisition of these characteristics in benign tumour cells. When considering the adenoma-carcinoma sequence this implicates *F. nucleatum*, and to a lesser extent *E. coli*, in the progression of late-stage disease. As the overwhelming

majority of cancer-related deaths are attributed to metastasis, this supports the growing literature which identifies *F. nucleatum* as a marker of poor prognosis in CRC patients. This research could not link the tumour-promoting effects of *F. nucleatum* infection to an increase in CRT, and therefore disagrees with the results published by Rubinstein *et al.* (2013). Although this does not represent a direct contradiction, as differences between studies are possible due to strain specific characteristics and differences in experimental procedures, it does highlight the need for thorough investigation into the interaction of microorganisms and tumours. Finally, this chapter has identified numerous cell motility and EMT-related genes which are regulated by *F. nucleatum* and may potentially be responsible for this species' invasion-promoting effects. Although these pathways require further investigation, the apparent synergy between these regulated genes, particularly the relationship between *SVIL, ACTN* and *MMP2&3,* indicates that *F. nucleatum* may promote invadopodia formation and ECM degradation in HCT116 cells. The findings presented in this chapter suggest that not only can tumour-associated bacteria be used as prognostic markers in CRC, but that they may be associated with lesser-known pathways which may represent novel treatment targets for metastatic disease.

Chapter 6: Final Discussion and Future Work

There is an increasing focus on the role of microorganisms in CRC, although studies are limited by the inherent difficulties in studying interactions between bacteria and tumours over prolonged time periods. There is a need to improve our understanding of how this relationship unfolds over the adenoma-carcinoma sequence in order to make a positive impact on patient treatment and prognosis. The main aim of this thesis was to investigate whether CRC-associated microorganisms influence the behaviour of benign and malignant tumour cells and whether they do so differently. The presence of specific microorganisms in CRC patients has been extensively investigated, leading to several species being proposed as CRC drivers (Feng *et al.*, 2015; Mori *et al.*, 2018; Saffarian *et al.*, 2019). Giving the example of Enterotoxigenic strains of *B. fragilis*, which are found in increased abundance in CRC patients, the 'Alpha-bug' hypothesis proposes that specific bacteria are able to drive CRC development at all stages of its progression, beginning with healthy epithelium (Ulger Toprak *et al.*, 2006; Sears and Pardoll, 2011). More recently, the Driver-Passenger model was proposed which suggests that bacterial drivers, which directly promote CRC development, are transiently associated with specific stages of CRC, and are gradually outcompeted by passenger species which are better adapted to the tumour microenvironment (Tjalsma *et al.*, 2012). Therefore, some species reported as being abundant in CRC patients may simply be passengers and possess no pro-tumorigenic effects.

Developing our understanding of microbial involvement in CRC is crucial, as it remains a leading cause of cancer-related death (Cancer Research UK, 2021). Although survival rates in developed countries have improved, increasing incidence in developing nations and an earlier onset in Western populations suggests that CRC

will remain a global public health challenge (Sung *et al.*, 2021). The Western diet has been suggested as a key risk factor for CRC development in high income countries (Adlercreutz, 1990). Such dietary patterns alter the composition of the colonic microbiota, allowing the colonisation and outgrowth of bacterial species which have been linked to CRC development and progression. This project investigated the tumour-promoting characteristics of three species commonly associated with CRC: *B. fragilis* (Enterotoxigenic), *E. faecalis* and *F. nucleatum*. In addition, the influence of *E. coli* Nissle on tumour behaviour was also investigated, as it is currently administered as a probiotic to treat inflammatory bowel disease, which is itself a risk factor for CRC development (Lutgens *et al.*, 2008; Kim and Chang, 2014).

To investigate how bacteria may influence tumour behaviour at different stages of CRC development, the RG/C2 and HCT116 cell lines were selected. RG/C2 cells were derived from a tubular adenoma and have retained many of the characteristics present in normal colon epithelium, including microvilli and mucin droplet formation (Paraskeva *et al.*, 1989). Furthermore, this cell line is non-invasive and non-tumorigenic in athymic mice. This is in contrast to the HCT116 cell line which was isolated from an adenocarcinoma, and is therefore highly tumorigenic and represents an advanced stage of CRC (Brattain *et al.*, 1981). In Chapter 3, the gentamicin protection assay was used to demonstrate that all four of the species tested readily associate with colorectal tumour cells, both attaching to the surface and invading the monolayer. This assay revealed the first difference in the interaction of bacteria with benign and malignant tumour cells, with levels of bacterial attachment being higher in RG/C2 co-cultures, in contrast to bacterial invasion which was higher in HCT116 co-cultures. Although the levels of attachment and invasion differed between cell lines, the total colony forming units isolated from RG/C2 and HCT116 co-cultures were

similar, suggesting that bacteria may be blocked from invading benign cells, or that malignant cells are more permissive to bacterial entry. These differing levels of bacterial association may be key in regulating the effects of bacteria on CRC progression, as many proposed mechanisms for promotion of tumorigenesis by bacteria involve microbial binding to cell surface proteins such as E-cadherin, expression of which changes during the adenoma-carcinoma sequence (Steck *et al.*, 2011; Guo *et al.*, 2020).

The mechanisms underlying the differences in attachment and invasion may be species specific and warrant further investigation. One disadvantage of the gentamicin protection assay is that it does not distinguish between active and passive entry of bacteria in mammalian cells. This could be elucidated by repeating the assay and including heat-killed bacterial cells and testing for their internalisation using culture-independent methods e.g. antibody staining. If heat-killed bacteria were not internalised, this would indicate that the bacteria are actively invading colorectal tumour cells. This would be supported by previous research, which has demonstrated that increased numbers of invasive bacteria are present in direct contact with the epithelium in CRC patients when compared to healthy controls (Strauss *et al.*, 2011). Bacteria have been demonstrated to bind to several epithelial cell surface molecules to facilitate their invasion. For example, *F. nucleatum* was found to attach to the cell surface carbohydrate Gal-GalNAC, which is overexpressed in adenocarcinoma and specific adenoma subgroups (Abed *et al.*, 2016). The use of the RG/C2 cell line, and other adenoma-derived cell lines such as the AA/C1 series also isolated by Paraskeva *et al.* (Paraskeva *et al.*, 1984*)*, could therefore provide a unique opportunity to investigate how levels of these surface molecules alter during the adenoma-carcinoma

sequence, which coupled with further gentamicin protection assays would provide insight into the changing levels of bacteria association.

To the author's knowledge, Chapter 3 also demonstrates for the first time the intracellular persistence of the four species in a simple 2D culture of colorectal tumour cells under standard conditions (37°C and 5% CO_2). This enabled the use of 'bacteria-infected' cells for further downstream assays. Studying the interactions between tumour cells and CRC-associated bacteria *in vitro* has proved challenging as many of these species are fastidious, strict anaerobes. This has led to innovation and the development of numerous models for gut-microbe simulation including the 'Simulator of Human Intestinal Microbiome' (SHIME) and Human-oxygen Bacteria-anaerobic (HoxBan) systems (Molly *et al.*, 1994; Sadabad *et al.*, 2015). However, such models are not suitable for all applications to study mammalian cell behaviour and can be associated with high cost and requisite expertise which may preclude their use in smaller laboratories. The use of 3D co-culture systems has also proved effective in investigating bacteria-tumour interactions. In particular, a recent study by Kasper *et al.* (2020) demonstrated that tumour spheroids of colorectal origin (including HCT116-derived spheroids), when grown to a sufficient size are able to provide an anaerobic environment in their centre which facilitates the survival and proliferation of anaerobic bacteria such as *F. nucleatum* (Kasper *et al.*, 2020)*.* Finally, microfluidics-based approaches have also shown promise in modelling the interface between the colonic mucosa and microorganisms. These include the HuMiX system, which has the benefit of incorporating immune cells to better mimic the colonic microenvironment, as cross-talk between the immune system and multiple cell types is a key mechanism in CRC development and progression (Shah *et al.*, 2016). An important characteristic of these models is their modularity, allowing the incorporation

of multiple cell types to target specific research questions. The results presented in this thesis show that simple 2D co-culture systems are appropriate for rapid investigation of bacteria-tumour interactions and their effects, which can then be progressed to more complex models. In addition, simple modifications such as the addition of mucin to the bacterial culture medium were demonstrated to produce phenotypic changes in bacterial behaviour, indicating that this model could be further developed to more closely represent *in vivo* conditions.

After establishing that the four species of bacteria tested interact with colorectal tumour cells *in vitro*, Chapter 4 and Chapter 5 report the investigation of the effects of this interaction on tumour cell behaviour. By monitoring tumour cell yield of bacteria-infected cells, it was found that all four species significantly promoted cell yield in both RG/C2 and HCT116 cells. This was coupled with a decrease in the number of apoptotic cells counted, with the exception of *F. nucleatum* infection of HCT116 cells. This represents a clear means of bacterial contribution to tumour progression, with resisting apoptosis being a central hallmark of cancer (Hanahan and Weinberg, 2011). Apoptosis resistance facilitates clonal expansion in tumour cells, allowing the accumulation of cancer driving mutations; although, paradoxically some level of apoptosis within tumours may be beneficial, as it favours the outgrowth of more aggressive sub-clones (Labi and Erlacher, 2015). Nevertheless increasing tumour cell yield, and subsequent increases in tumour size could represent a mechanism through which tumour colonisation by these organisms worsens patient prognosis (Kornprat *et al.*, 2011).

When investigating the promotion of cell yield in colorectal tumours, it is important to consider cellular metabolism as a key moderating factor. To generate ATP *in vivo*, colonocytes primarily metabolise the bacterial fermentation product

butyrate, the action of which also influences proliferation, differentiation, motility and stemness (Leschelle *et al.*, 2000). Through the Warburg effect, tumours preferentially undergo aerobic glycolysis, and the resulting accumulation of butyrate in the cytosol allows it to translocate to the nucleus and perform its functions as an HDACi (Warburg, Wind and Negelein, 1927; Donohoe *et al.*, 2012). This effect has been linked to mutations in *APC*, and so is thought to be an early event in CRC development (Satoh *et al.*, 2017; Cha *et al.*, 2021). The metabolic phenotype assays performed in this study confirmed that metabolic differences are present between colorectal adenoma and adenocarcinoma cells. Specifically, RG/C2 cells exhibited far higher rates of oxygen consumption compared to the three adenocarcinoma cell lines tested, confirming that a downregulation of oxidative phosphorylation may occur during the transition from adenoma to adenocarcinoma.

In addition to differences in the ratio of oxidative phosphorylation to glycolysis, cancer formation is associated with changes in numerous other metabolic pathways. There is currently a substantial research effort to develop diagnostic tools which take advantage of these altered metabolic profiles, with the use of gas sensors showing promise in being able to accurately diagnose numerous cancers (Omar *et al.*, 2016; Bond *et al.*, 2019). In this study, SIFT-MS was used to investigate the volatile product ion peaks detected from the headspace above colorectal tumour cells. This data demonstrated that benign cells could be differentiated from malignant cells based on their volatile profiles, which is supported by the results of the Seahorse metabolic assays. A recent study has successfully demonstrated that biomarkers present in urine are able to diagnose colorectal adenomas or CRC using GC/MS (Mozdiak *et al.*, 2019). However, it should be considered that these metabolites are likely to be blood-borne and may not be reflected in the metabolites detected in the headspace above

tumour cell cultures *in vitro*. Despite this, the results presented in this study highlight the suitability of SIFT-MS analysis of RG/C2 and adenocarcinoma cells for identifying potential volatile biomarkers, which could then be quantified under different conditions (e.g. different nutrient supplements to the medium, hypoxic conditions), improving our understanding of the mechanisms which underlie these metabolic changes.

Having established that metabolic differences are apparent between benign and malignant colorectal tumour cells *in vitro*, Seahorse and SIFT-MS techniques were employed to determine if bacterial infection produced a detectable metabolic change. SIFT-MS was able to discriminate *B. fragilis* and *F. nucleatum*-infected RG/C2 cells from those infected with *E. coli* Nissle or *E. faecalis*; however, this result was not reproduced in HCT116 cells. Other research groups have succeeded in accurately discriminating between infectious agents, intracellular bacteria and viruses, *in vitro* (Slade *et al.*, 2017; Traxler *et al.*, 2019). Therefore, identification of specific bacterial agents based on the volatile profiles of the infected mammalian cells may be possible with altered assay parameters, such as increasing the infection period or using different volatile extraction techniques (e.g. solid phase microextraction or solvents).

Interestingly, Seahorse metabolic rate assays demonstrated that both *B. fragilis* and *F. nucleatum* produced a metabolic shift in infected HCT116 cells. Both species increased basal levels of glycolysis, with *F. nucleatum* also increasing the maximum glycolytic rate. A Warburg-like metabolism has previously been described in mammalian cells upon bacterial infection with species such as *Brucella abortus*; however, it is currently unclear whether this phenomenon is induced by the infecting bacteria, or is a response of the human cell to this infection (Czyz, Willett and Crosson, 2017; Escoll and Buchrieser, 2018). Regardless of the underlying mechanism, a

Warburg-like metabolic phenotype induced by *B. fragilis* and *F. nucleatum* would confer the previously discussed advantages to tumour cells (see Section 1.4.3), which include increased growth rate, increased glucose uptake and suppression of immune function through lactate secretion. In order to elucidate the contribution of this metabolic shift to the observed increases in tumour cell yield, glycolysis could be targeted to determine if this effect occurs independent of glucose metabolism. Shiratori and colleagues (2019) have found that substituting glucose present in cell culture medium for galactose yields no net ATP from glycolysis production, and this would represent a suitable model for studying the contribution of bacteria-induced glycolysis to tumour cell growth (Shiratori *et al.*, 2019).

Ultimately, the results of Chapter 4 implicate infection of colorectal tumours by bacteria in the progression of CRC through increasing cell yield, thereby facilitating clonal evolution. Chapter 5 focussed specifically on the contribution of these microorganisms to tumour invasion and metastasis. Metastasis is the leading cause of cancer-related death across all cancerous diseases (Dillekås, Rogers and Straume, 2019). The tumour microenvironment has been implicated as a key regulator of metastatic development, particularly in CRC where this microenvironment also includes the microbiota and its associated metabolites. The results presented in Chapter 5 demonstrated that two species, *E. coli* Nissle and *F. nucleatum*, are able to significantly promote HCT116 migration *in vivo*. The migration-promoting abilities of *E. coli* Nissle could be a by-product of its probiotic activity, and may be beneficial to the healing of ulcerative lesions in inflammatory bowel disease (Kruis *et al.*, 2004). Crucially, *E. coli* Nissle did not contribute towards HCT116 invasion, in contrast to *F. nucleatum* which produced a significant increase in invasive behaviour which is consistent with the findings of other research groups (Abdulkareem *et al.*, 2018; Ma

et al., 2018). This may be a pivotal means through which *F. nucleatum* colonisation of tumour worsens patient prognosis, as metastasis accounts for the overwhelming majority of CRC deaths. The lack of impact on tumour cell migration by *B. fragilis* was surprising, as this species has previously been linked to enhanced CRC progression through the activation of β-catenin signalling (Wu *et al.*, 2003). However, it is important to consider that transwell filter and wound healing assays do not fully encompass the range of mechanisms through which bacteria may contribute to tumour migration. For example, the increased rate of glycolysis observed in *B. fragilis*-infected HCT116 cells (see Chapter 4.4.6) could facilitate tumour migration through increased lactate production; as this has been demonstrated to promote tumour cell motility through acidification of the tumour microenvironment (Goetze *et al.*, 2011).

One of the most substantial findings of this thesis is that despite the significant migrating-promoting effects of certain species in the malignant HCT116 cell line, no effect on the migration or invasion of benign RG/C2 cells was reported. This would suggest that the action of these bacteria (e.g. *F. nucleatum* and *E. coli)* is more relevant to late-stage CRC patients than those harbouring benign tumours, and implicates these species as late drivers of CRC progression (Tjalsma *et al.*, 2012). Recent studies have investigated the efficacy of targeted bacterial treatment in patient-derived mouse xenograft models, and demonstrated that eradication of *F. nucleatum* in these tumours reduced overall tumour growth (Bullman *et al.*, 2017). The results of this thesis indicate that adjuvant antibiotic therapy in CRC patients could improve patient prognosis by eliminating the promotion of tumour cell yield and migration/invasion by species such as *F. nucleatum*.

Although this thesis progresses our understanding of the temporal association between tumour-promoting bacteria and disease stage, further research is required to

identify the mechanisms responsible for the observed effects on tumour growth and migration. This is particularly important if microorganisms are to be targeted in CRC treatment, as any therapy should be highly specific to avoid the negative consequences of broad antibiotic use, which can include increased risk of CRC (Cao *et al.*, 2018b). Tumour cell yield was increased after infection by all species studied, at relatively low MOI. This may suggest that the mechanism is not species-specific and could be due to tumour cell interactions with basic bacterial antigens e.g. LPS. The ability of some species to promote migration and invasion is likely to be attributable to specific virulence factors. For example, the *F. nucleatum* adhesin protein FadA has been linked to promotion of CRC and other cancers, with FadA$^-$ mutants possessing less invasive capabilities and lower tumorigenic potential (Guo *et al.*, 2020). Therefore, performing similar experiments with knockout strains for a number of common virulence factors would provide valuable insight into the mechanisms through which bacteria promote CRC progression.

The difference in response between RG/C2 and HCT116 cells, particularly in the migration assays presented in Chapter 5, is an exciting result, and requires further research. Interestingly, these changes could not be aligned with changes in catenin-related transcription of bacteria-infected cells, which has been commonly proposed as a mechanism of *F. nucleatum*-associated tumour promotion (Rubinstein *et al.*, 2013; Chen *et al.*, 2017). The PCR microarrays performed using cells infected with *F. nucleatum* in this thesis have highlighted numerous pathways which would be beneficial to explore. Of particular interest is the apparent upregulation of *SVIL*, *MMP2* and *MMP3*, which according to the literature may act synergistically with one another to promote tumour cell invasion (Clark and Weaver, 2008; Crowley *et al.*, 2009). Such investigations should begin by determining whether these changes in

mRNA levels detected by the PCR microarrays are reflected in the level of transcribed protein. Finally, knockout studies using CRISPR/Cas9 or short-interfering RNAs to eliminate expression of these genes would determine whether the migration and invasion-promoting effects of *F. nucleatum* infection persists in their absence. This would allow us to elucidate the mechanism of bacterially stimulated motility and invasion.

In summary, this research was performed with the aim of determining whether CRC-associated microorganisms, or the probiotic species *E. coli* Nissle, differentially influence the behaviour of benign and malignant tumour cells. It appears that all species are capable of indirectly promoting colorectal tumour progression through increasing cell yield, whereas *F. nucleatum* emerged as the one species investigated capable of driving the migration and invasion of adenocarcinoma cells. The data presented was collected using a simplistic 2D co-culture model under specific micro-environmental conditions, and therefore further experiments using more representative models are required. However, this thesis adds support to the growing appreciation of the key role microorganisms play in CRC, particularly in the development of migratory and invasive characteristics. The difference in both the interaction of benign and malignant cells with bacteria, and the subsequent effect of this interaction on their behaviour, suggests that the roles of these microorganisms may shift during disease progression. Therefore, integrating our knowledge of the involvement of microorganisms in CRC at specific stages is essential to improving patient prognosis, and may hold the key to providing novel treatment and diagnostic techniques

Chapter 7: Appendix
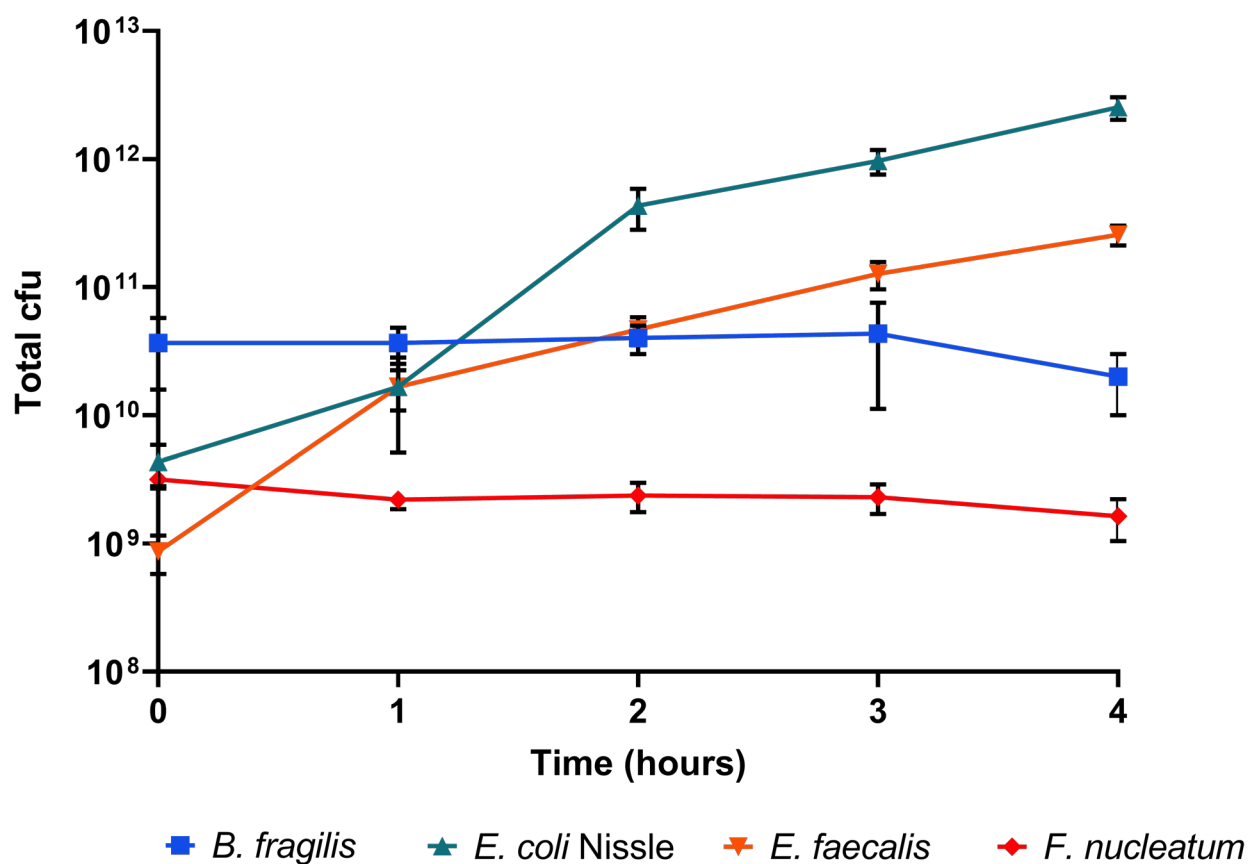
Supplementary Figure 1: Survival of bacteria in serum-free DMEM under standard culture conditions. Overnight bacterial cultures were adjusted to an optical density of 0.05 in serum-free DMEM (high glucose) in a total volume of 10 ml. Bacteria were then incubated under standard cell culture conditions (37°C and 5% CO_2 for 4 hours, corresponding to the total infection time used for all other assays. At hourly intervals, bacteria were plated onto blood agar plates, incubated under anaerobic conditions, and the total colony forming units (cfu) was calculated for each time point. The facultative anaerobes *E. coli* Nissle and *E. faecalis* were able to proliferate under these conditions. Both *B. fragilis* and *F. nucleatum* were not able to grow in DMEM under standard culture conditions, but no reduction in the total viable cfu was observed, indicating that these conditions did not negatively impact their viability.
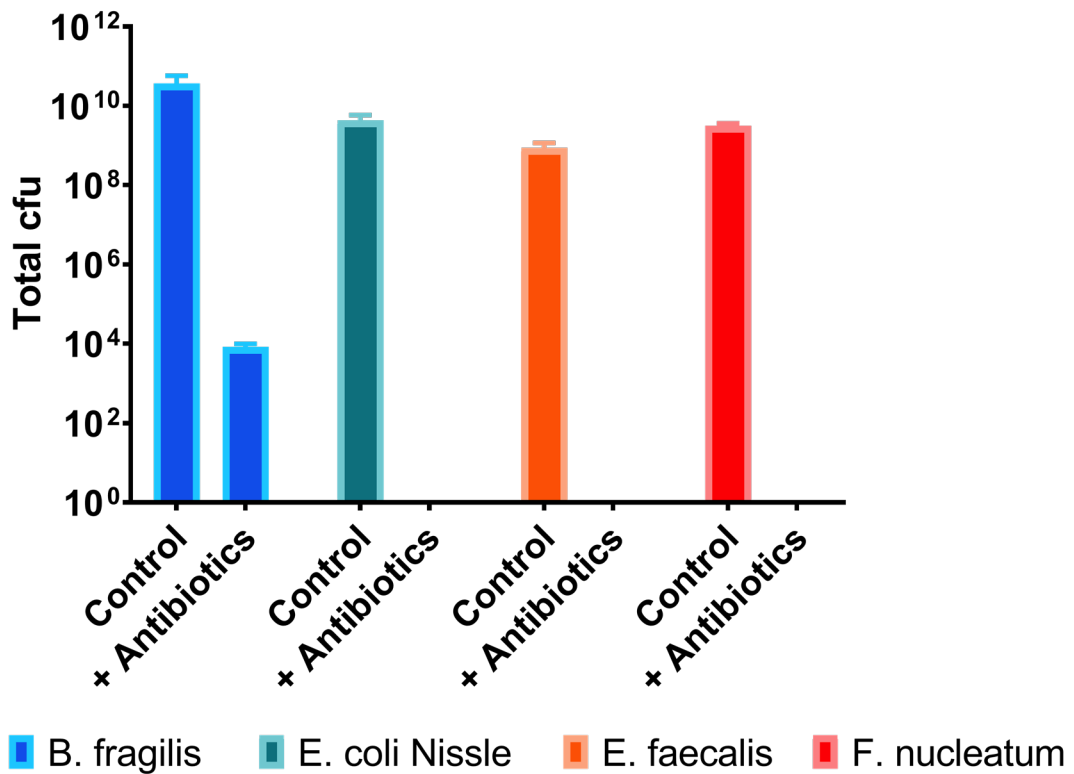
Supplementary Figure 2: Efficacy of gentamicin and metronidazole against target bacterial species. Overnight bacterial cultures were adjusted to an optical density of 0.05, before being pelleted by centrifugation and resuspended in serum-free DMEM containing 300 μg/ml gentamicin and 200 μg/ml metronidazole (antibiotics). Cultures were incubated under standard culture conditions for one hour, reflecting the antibiotic treatment time used in all gentamicin protection assays, before being enumerated onto blood agar plates. At this concentration, the antibiotics completely eliminated all viable bacteria for all species, with the exception of *B. fragilis* where a 6 log reduction in viable cfu was recorded. As increasing the concentration of antibiotics any further may have caused entry into tumour cells, negatively impacting the validity of the gentamicin protection assays, this concentration was deemed suitable for all further experiments.
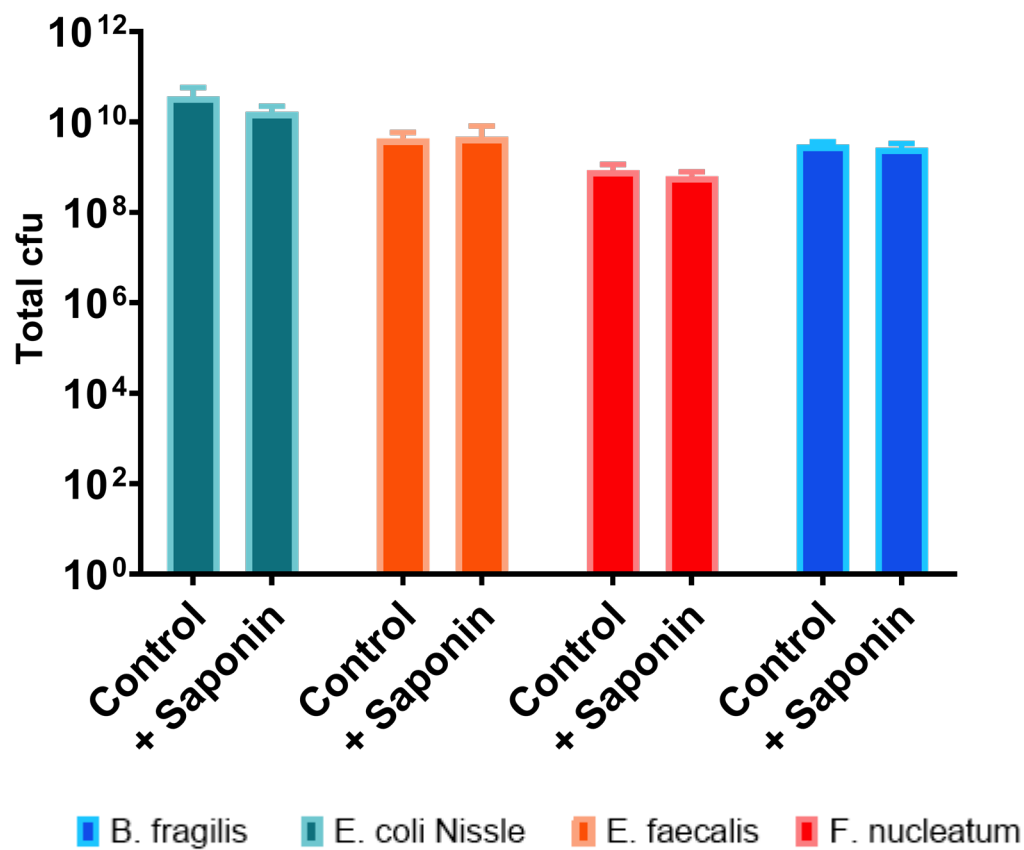
Supplementary Figure 3: Saponin does not impact bacterial viability. Overnight bacterial cultures were adjusted to an optical density of 0.05, before being pelleted by centrifugation and resuspended in serum-free DMEM containing 1% w/v saponin. Cultures were then incubated for one hour under standard cell culture conditions, before being enumerated onto blood agar plates. Incubation with saponin did not alter bacterial viability in any of the four species tested (paired T-test). This indicates that this method of tumour cell lysis does not affect the measurement of attached/intracellular bacteria.

Supplementary Table 1: Complete list of target genes in Qiagen Cell Motility PCR Microarrays (Catalogue number: PAHS-128Z)

Symbol	Description	Symbol	Description
ACTN1	Actinin, alpha 1	MYH10	Myosin, heavy chain 10, non-muscle
ACTN3	Actinin, alpha 3	MYH9	Myosin, heavy chain 9, non-muscle
ACTN4	Actinin, alpha 4	MYL9	Myosin, light chain 9, regulatory
ACTR2	ARP2 actin-related protein 2 homolog (yeast)	MYLK	Myosin light chain kinase
ACTR3	ARP3 actin-related protein 3 homolog (yeast)	PAK1	P21 protein (Cdc42/Rac)-activated kinase 1
AKT1	V-akt murine thymoma viral oncogene homolog 1	PAK4	P21 protein (Cdc42/Rac)-activated kinase 4
ARF6	ADP-ribosylation factor 6	PFN1	Profilin 1
ARHGDIA	Rho GDP dissociation inhibitor (GDI) alpha	PIK3CA	Phosphoinositide-3-kinase, catalytic, alpha polypeptide
ARHGEF7	Rho guanine nucleotide exchange factor (GEF) 7	PLAUR	Plasminogen activator, urokinase receptor
BAIAP2	BAI1-associated protein 2	PLCG1	Phospholipase C, gamma 1
BCAR1	Breast cancer anti-estrogen resistance 1	PLD1	Phospholipase D1, phosphatidylcholine-specific
CAPN1	Calpain 1, (mu/I) large subunit	PRKCA	Protein kinase C, alpha
CAPN2	Calpain 2, (m/II) large subunit	PTEN	Phosphatase and tensin homolog
CAV1	Caveolin 1, caveolae protein, 22kDa	PTK2	PTK2 protein tyrosine kinase 2
CDC42	Cell division cycle 42 (GTP binding protein, 25kDa)	PTK2B	PTK2B protein tyrosine kinase 2 beta
CFL1	Cofilin 1 (non-muscle)	PTPN1	Protein tyrosine phosphatase, non-receptor type 1
CRK	V-crk sarcoma virus CT10 oncogene homolog (avian)	PXN	Paxillin
CSF1	Colony stimulating factor 1 (macrophage)	RAC1	Ras-related C3 botulinum toxin substrate 1
CTTN	Cortactin	RAC2	Ras-related C3 botulinum toxin substrate 2
DIAPH1	Diaphanous homolog 1 (Drosophila)	RASA1	RAS p21 protein activator 1
DPP4	Dipeptidyl-peptidase 4	RDX	Radixin
EGF	Epidermal growth factor	RHO	Rhodopsin
EGFR	Epidermal growth factor receptor	RHOA	Ras homolog gene family, member A
ENAH	Enabled homolog (Drosophila)	RHOB	Ras homolog gene family, member B
EZR	Ezrin	RHOC	Ras homolog gene family, member C
FAP	Fibroblast activation protein, alpha	RND3	Rho family GTPase 3
FGF2	Fibroblast growth factor 2 (basic)	ROCK1	Rho-associated, coiled-coil containing protein kinase 1
HGF	Hepatocyte growth factor	SH3PXD2A	SH3 and PX domains 2A
IGF1	Insulin-like growth factor 1 (somatomedin C)	SRC	V-src sarcoma (Schmidt-Ruppin A-2) viral oncogene homolog (avian)
IGF1R	Insulin-like growth factor 1 receptor	STAT3	Signal transducer and activator of transcription 3
ILK	Integrin-linked kinase	SVIL	Supervillin
ITGA4	Integrin, alpha 4	TGFB1	Transforming growth factor, beta 1
ITGB1	Integrin, beta 1	TIMP2	TIMP metallopeptidase inhibitor 2
ITGB2	Integrin, beta 2	TLN1	Talin 1
ITGB3	Integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)	VASP	Vasodilator-stimulated phosphoprotein
LIMK1	LIM domain kinase 1	VCL	Vinculin
MAPK1	Mitogen-activated protein kinase 1	VEGFA	Vascular endothelial growth factor A
MET	Met proto-oncogene (hepatocyte growth factor receptor)	VIM	Vimentin
MMP14	Matrix metallopeptidase 14	WASF1	WAS protein family, member 1
MMP2	Matrix metallopeptidase 2	WASF2	WAS protein family, member 2
MMP9	Matrix metallopeptidase 9	WASL	Wiskott-Aldrich syndrome-like
MSN	Moesin	WIPF1	WAS/WASL interacting protein family, member 1

Supplementary Table 2: Complete list of target genes in Qiagen EMT PCR Microarrays (Catalogue number: PAHS-090Z)

Symbol	Description	Symbol	Description
AHNAK	AHNAK nucleoprotein	MST1R	Macrophage stimulating 1 receptor
AKT1	V-akt murine thymoma viral oncogene homolog 1	NODAL	Nodal homolog (mouse)
BMP1	Bone morphogenetic protein 1	NOTCH1	Notch 1
BMP2	Bone morphogenetic protein 2	NUDT13	Nudix (nucleoside diphosphate linked moiety X)-type motif 13
BMP7	Bone morphogenetic protein 7	OCLN	Occludin
CALD1	Caldesmon 1	PDGFRB	Platelet-derived growth factor receptor, beta polypeptide
CAMK2N1	Calcium/calmodulin-dependent protein kinase II inhibitor 1	PLEK2	Pleckstrin 2
CAV2	Caveolin 2	DESI1	PPPDE peptidase domain containing 2
CDH1	Cadherin 1, type 1, E-cadherin (epithelial)	PTK2	PTK2 protein tyrosine kinase 2
CDH2	Cadherin 2, type 1, N-cadherin (neuronal)	PTP4A1	Protein tyrosine phosphatase type IVA, member 1
COL1A2	Collagen, type I, alpha 2	RAC1	Ras-related C3 botulinum toxin substrate 1
COL3A1	Collagen, type III, alpha 1	RGS2	Regulator of G-protein signalling 2, 24kDa
COL5A2	Collagen, type V, alpha 2	SERPINE1	Serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1
CTNNB1	Catenin (cadherin-associated protein), beta 1, 88kDa	GEMIN2	Survival of motor neuron protein interacting protein 1
DSC2	Desmocollin 2	SMAD2	SMAD family member 2
DSP	Desmoplakin	SNAI1	Snail homolog 1 (Drosophila)
EGFR	Epidermal growth factor receptor	SNAI2	Snail homolog 2 (Drosophila)
ERBB3	V-erb-b2 erythroblastic leukaemia viral oncogene homolog 3 (avian)	SNAI3	Snail homolog 3 (Drosophila)
ESR1	Estrogen receptor 1	SOX10	SRY (sex determining region Y)-box 10
F11R	F11 receptor	SPARC	Secreted protein, acidic, cysteine-rich (osteonectin)
FGFBP1	Fibroblast growth factor binding protein 1	SPP1	Secreted phosphoprotein 1
FN1	Fibronectin 1	STAT3	Signal transducer and activator of transcription 3
FOXC2	Forkhead box C2	STEAP1	Six transmembrane epithelial antigen of the prostate 1
FZD7	Frizzled family receptor 7	TCF3	Transcription factor 3
GNG11	Guanine nucleotide binding protein gamma 11	TCF4	Transcription factor 4
GSC	Goosecoid homeobox	TFPI2	Tissue factor pathway inhibitor 2
GSK3B	Glycogen synthase kinase 3 beta	TGFB1	Transforming growth factor, beta 1
IGFBP4	Insulin-like growth factor binding protein 4	TGFB2	Transforming growth factor, beta 2
IL1RN	Interleukin 1 receptor antagonist	TGFB3	Transforming growth factor, beta 3
ILK	Integrin-linked kinase	TIMP1	TIMP metallopeptidase inhibitor 1
ITGA5	Integrin, alpha 5	TMEFF1	Transmembrane protein with EGF-like and two follistatin-like domains 1
ITGAV	Integrin, alpha V	TMEM132A	Transmembrane protein 132A
ITGB1	Integrin, beta 1	TSPAN13	Tetraspanin 13
JAG1	Jagged 1	TWIST1	Twist homolog 1 (Drosophila)
KRT14	Keratin 14	VCAN	Versican
KRT19	Keratin 19	VIM	Vimentin
KRT7	Keratin 7	VPS13A	Vacuolar protein sorting 13 homolog A
MAP1B	Microtubule-associated protein 1B	WNT11	Wingless-type MMTV integration site family, member 11
MMP2	Matrix metallopeptidase 2	WNT5A	Wingless-type MMTV integration site family, member 5A
MMP3	Matrix metallopeptidase 3	WNT5B	Wingless-type MMTV integration site family, member 5B
MMP9	Matrix metallopeptidase 9	ZEB1	Zinc finger E-box binding homeobox 1
MSN	Moesin	ZEB2	Zinc finger E-box binding homeobox 2

Supplementary Figure 4: Impact of *F. nucleatum* infection on Cell Motility gene expression in RG/C2 cells. Threshold values for each gene obtained from qPCR microarrays performed on *F. nucleatum*-infected cells were normalised to un-infected cells. Colour intensity indicates the relative fold regulation for each gene. Positive regulations are denoted in green, whereas negative regulations are displayed in red.

ACTN1	ACTN3	ACTN4	ACTR2	ACTR3	AKT1	ARF6	ARHGDIA	ARHGEF7	BAIAP2	BCAR1	CAPN1
-1.52	1.86	-1.5	-1.47	-1.44	-1.58	-1.26	-1.5	-1.5	-1.45	-1.27	-1.73
CAPN2	CAV1	CDC42	CFL1	CRK	CSF1	CTTN	DIAPH1	DPP4	EGF	EGFR	ENAH
-1.36	-12.55	-1.65	-1.6	-1.3	-6.35	-1.75	-1.35	-1.44	-1.64	-1.36	-1.07
EZR	FAP	FGF2	HGF	IGF1	IGF1R	ILK	ITGA4	ITGB1	ITGB2	ITGB3	LIMK1
-1.39	1.86	-1.11	1.86	1.86	-1.52	-1.65	1.86	-1.41	-2.18	-2.35	-1.49
MAPK1	MET	MMP14	MMP2	MMP9	MSN	MYH10	MYH9	MYL9	MYLK	PAK1	PAK4
-1.39	-1.4	-1.11	-2.59	1.24	-1.11	-1.23	-1.51	-1.81	-1.54	-1.5	-1.56
PFN1	PIK3CA	PLAUR	PLGC1	PLD1	PRKCA	PTEN	PTK2	PTK2B	PTPN1	PXN	RAC1
-1.52	-1.3	-1.42	-1.56	-1.41	-1.5	-1.6	-1.86	-1.47	-1.79	-1.44	-1.75
RAC2	RASA1	RDX	RHO	RHOA	RHOB	RHOC	RND3	ROCK1	SH3PXD2A	SRC	STAT3
-2.03	-1.54	-1.23	1.86	-1.63	-1.42	-1.51	-1.64	-1.08	-1.25	-1.63	-1.4
SVIL	TGFB1	TIMP2	TLN1	VASP	VCL	VEGFA	VIM	WASF1	WASF2	WASL	WIPF1
-1.41	-2.35	1.08	-1.25	-1.19	-1.78	-1.57	-1.9	-1.78	-1.64	-1.47	-1.75

Key

| <-10 | -10 | -5 | 0 | 5 | 10 |

Supplementary Figure 5: Impact of *F. nucleatum* infection on EMT gene expression in RG/C2 cells. Threshold values for each gene obtained from qPCR microarrays performed on *F. nucleatum*-infected cells were normalised to un-infected cells. Colour intensity indicates the relative fold regulation for each gene. Positive regulations are denoted in green, whereas negative regulations are displayed in red. Despite numerous genes being regulated by *F. nucleatum* infection in RG/C2 cells no effect on the migratory or invasive phenotype was observed (see Section 5.4.1-5.4.4).

AHNAK 4.75	AKT1 4.29	BMP1 3.02	BMP2 3.05	BMP7 1.46	CALD1 -19.6	CAMK2N1 -1.39	CAV2 1.5	CDH1 1.34	CDH2 -1.51	COL1A2 -1.51	COL3A1 -1.51
COL5A2 1.13	CTNNB1 1.2	DSC2 2.11	DSP 2.59	EGFR 3.66	ERBB3 4.11	ESR1 -6.79	F11R 1.83	FGFBP1 -1.6	FN1 -17.21	FOXC2 5.13	FZD7 2.3
GNG11 -1.51	GSC 3.11	GSK3B 1.25	IGFBP4 -2.55	IL1RN -1.71	ILK 2.06	ITGA5 3.89	ITGAV 2.09	ITGB1 -1.07	JAG1 1.81	KRT14 -1.19	KRT19 -1.32
KRT7 -1.25	MAP1B 1.99	MMP2 -1.51	MMP3 -8.48	MMP9 -1.09	MSN 2.6	MST1R 6.56	NODAL 1.08	NOTCH1 -1.49	NUDT13 -1.52	OCLN 1.98	PDGFRB 8.95
PLEK2 1.16	DESI1 1.67	PTK2 1.56	PTP4A1 1.44	RAC1 -1.07	RGS2 -1.23	SERPINE1 -1.57	GEMIN2 -1.11	SMAD2 1.64	SNAI1 1.16	SNAI2 -1.51	SNAI3 2.41
SOX10 -1.51	SPARC -2.1	SPP1 3.89	STAT3 2.12	STEAP1 -3.34	TCF3 -1.35	TCF4 -1.51	TFPI2 15.91	TGFB1 2.44	TGFB2 6.25	TGFB3 -2.77	TIMP1 -1.79
TMEFF1 1.08	TMEM123A 7.83	TSPAN13 1.15	TWIST1 -1.51	VCAN -5.31	VIM 1.35	VSP13A 1.79	WNT11 1.28	WNT5A -7.02	WNT5B -2.24	ZEB1 6.79	ZEB2 8.19

Key

<-10	-10	-5	0	5	10

Supplementary Table 3: Product ion array for colorectal tumour cell cultures. The headspace above tumour cell cultures was analysed via SIFT-MS. Tables shows the product ions detected from each cell line (n=3) with mean intensity log(cps) for each product ion included in analysis after background subtraction using uninoculated cell culture medium as a control. Table is continued on the following page.

m/z	RG/C2			HCT116			HT29			SW480		
15	3.66	3.60	3.46	3.63	3.59	3.55	3.64	3.58	3.58	3.64	3.57	3.59
16	4.03	3.99	3.88	4.04	3.98	3.94	4.05	3.96	3.96	4.03	3.98	3.97
23	1.89	1.83	1.65			1.45	1.87	1.79		1.96	1.58	1.65
24	2.70	2.68	2.40	2.77	2.63	2.57	2.69	2.59	2.59	2.72	2.60	2.67
25	3.11	2.98	3.04	3.10	2.99	3.02	3.08	3.01	3.06	3.09	2.95	3.02
26	2.26	1.88	2.27	2.26	1.72	2.21		2.02	2.22	2.19	2.10	2.09
27	4.96	4.85	4.91	4.93	4.86	4.88	4.92	4.85	4.89	4.93	4.85	4.91
28	3.12	3.03	3.09	3.28	3.11	3.17	3.21	3.21	3.28	3.26	3.15	3.24
32	6.60	6.50	6.49	6.68	6.52	6.60	6.66	6.55	6.61	6.67	6.54	6.61
33	3.54	3.35	3.36	3.64	3.36	3.54	3.57	3.44	3.59	3.60	3.42	3.58
34	4.09	3.99	3.98	4.18	4.02	4.11	4.15	4.02	4.14	4.17	4.03	4.15
40	3.29	3.27	3.10	3.26	3.25	3.17	3.25	3.23	3.15	3.34	3.24	3.22
41		3.01	3.07	2.96	2.84	3.10		2.82	2.96		2.87	2.96
42	3.80	3.59	3.90	3.93	3.60	4.02	3.82	3.63	4.04	3.93	3.64	4.08
44				2.99	3.27	2.81						
46				2.64								
50	5.77	5.65	5.70	5.94	5.66	5.87	5.88	5.72	5.87	5.92	5.70	5.85
51												
52	3.44	3.27	3.42	3.50	3.24	3.54	3.45	3.34	3.51	3.48	3.33	3.45
53	2.07		2.17		1.75	2.00	1.81	2.01	1.98	1.81	1.78	2.00
54	2.81	3.30	2.85	2.79	3.53	2.87	2.79	3.39	2.95	2.87	3.50	2.96
56	2.81	2.67	2.34							2.73	2.32	
58	2.95	3.01	2.63	2.96	3.00	2.89	2.78	2.77	2.77	2.89	2.90	2.78
60	3.58	3.40	3.76	3.74	3.45	3.97	3.68	3.46	3.94	3.78	3.52	3.98
64		2.40			2.59							2.45
68	2.65	2.64	2.64	2.68	2.67	2.88	2.80	2.81	2.82	2.69	2.68	2.83
69		2.76	2.88	2.61	2.86	2.93	2.41	2.83	2.99		2.79	2.86
70	2.46	2.58			2.31	2.16	2.07	1.88	2.22	2.30	2.11	2.14
72				2.36				2.43	2.07		2.28	
76		1.93			2.10		2.20	1.82	2.06	2.14	2.13	2.11
78	2.27	2.14	2.32	2.31		2.36		2.29	2.39			2.29
80	2.70	2.62	2.54	2.80			2.78					
82	2.24	2.17		2.00	1.94		2.23			2.02	2.23	1.56
84				2.22	1.92	2.17					1.99	1.96
86	2.46	2.72	2.25	2.31	2.27		2.13	2.23	2.26	2.07		1.93
88		2.29	2.34	2.07		2.21	2.10	2.31	2.15	2.12	2.25	2.24
90			1.85							1.70	1.47	
92	2.03	1.90	2.05				2.02	1.72	1.90	1.86		

94						2.70						
96		2.20	2.00	1.96	1.94		2.09			2.17	2.10	2.08
97		2.37	2.46	1.96	1.98	2.20	1.94		2.29			
98		1.77				1.75			2.11	2.24		2.05
100	2.82	2.87	2.73	3.07	3.06	3.02	3.11	3.07	3.11	3.14	3.02	3.09
104				2.01	2.08	1.99	1.88	1.80	1.74			1.97
110			2.14	2.00		2.00	2.24	1.97	1.98			1.98
112	2.07	2.30	2.06	2.31	1.89	2.36		2.12	2.30		2.04	2.02
116	2.92	3.14	3.16	2.76	3.09	3.10	2.89	3.04	3.11	2.90	2.96	3.06
118	1.75	1.98	2.08	1.75	1.85	1.88			1.80			
120	1.77			2.03		2.10						
122	2.85	2.77	2.60	2.76	2.51	2.93	2.84	2.37	2.62	2.89	2.59	2.43
124	2.10	1.94	1.94			2.05		1.94	2.13	2.11	1.98	
126	1.72	1.85	1.96									
130							1.46					
132				1.79	1.70				1.65			
134	2.46	2.57	2.75	2.55	2.71	2.77	2.43	2.51	2.72	2.70	2.52	2.68
136	2.66	2.49	2.56	2.58	2.71	2.82	2.64	2.48	2.60	2.56	2.49	2.48
137	1.95	2.20	2.40	2.28	2.42			2.21	2.35		2.08	2.26
138	2.41	2.33	2.16	2.23	2.10	2.38	2.24	1.98	2.17	2.41	2.18	2.26
140	2.03	1.87				1.94						
144	1.66	1.52										
146							1.64	1.70			1.37	
148	1.71		1.53									
150	2.62	2.53		2.62		2.64	2.34			2.61	2.30	
152	2.10	1.85		2.10		2.18	2.10				2.07	
154		1.92		1.89		1.96				1.85	1.92	
156										1.65		
158	1.74	2.05	2.19				1.68					
160	1.76											
162			1.87	2.26	1.81		1.77				1.65	1.51
164	3.05	2.85	2.69	2.94	2.54	2.95	2.98	2.64	2.84		2.73	2.73
165	1.95			2.21						2.32		
166	2.38	1.91		2.40	2.15	2.40	2.35	2.02	2.20	2.44	2.43	2.42
168	2.16						2.07					
172					1.79	2.11		1.85		1.79	2.10	
176	1.90	2.05										
178	2.54	2.42			2.61	2.77		2.45	2.43		2.48	2.67
180										2.40		
182						1.62		1.70	1.79			
188					2.12				2.23	1.66		
190	2.11		1.81	1.92								
192	2.54	2.80	2.42	2.69	2.66	2.79	2.64		2.50		2.63	2.45
194										2.34		

Key   1   2   3   4   5   6   7

Supplementary Table 4: Product ion array for bacteria-infected RG/C2 cells. The headspace above infected RG/C2 cultures was analysed via SIFT-MS after a 4-hour bacterial co-culture. Tables shows the product ions detected from each cell line (n=3) with mean intensity log(cps) for each product ion included in analysis after background subtraction using non-infected RG/C2 cells as a control. Table is continued on the following page.

m/z	B. fragilis			E. coli Nissle			E. faecalis			F. nucleatum		
15	3.67	3.61	3.51									
16	4.06	3.99	3.88									
23		1.68										
24	2.74	2.60	2.57									
25	3.13	3.03	2.97									
26	2.34	1.92	1.92									
27	4.96	4.86	4.91									
28	3.11	3.03	3.16									
32	6.59	6.49	6.52									
33	3.43	3.29	3.39	2.24	2.44	2.56						
34	4.07	3.95	4.02									
35										1.16	1.13	1.18
40	3.17	3.14	3.00				1.18				1.41	
41	2.97	3.01	2.99									
42	3.73	3.53	3.84									
49												2.16
50	5.72	5.57	5.72									
52	3.29	3.22	3.45									
54	2.83	3.22	2.81									
56	2.41	2.60										
58	2.73	2.87	2.56									
60	3.53	3.37	3.76									
64	2.13											
66	2.56	2.58	2.77									
67			2.30									
68	2.48	2.59	2.63									
69	2.79	2.93	3.02									
70	2.09	2.41	2.15									
72		2.17										
74					2.76							
76			1.97									
78	1.93	1.95	2.27									
80	2.60	2.51	2.89									
82	2.09	1.97	1.86									
85										2.71	2.63	2.59
86		2.39										
88	2.12	2.00	2.09									
89							2.43	2.40	2.33			
92	1.91		2.02									

94	2.57	2.50	2.82							
96	1.95	1.89	1.99							
97			2.05							
98	2.00							1.31	1.49	1.32
99								2.39	2.17	2.08
100	2.81	2.92	2.71							
104	1.80	2.03	0.00							
106	0.00	1.99	2.27			1.84				
108	2.54	2.60	2.76							
110	2.14		2.10							
111								2.03		
112	1.97									
113								2.73	2.75	2.50
116	2.90	3.11	3.00							
120	0.00	2.02	2.10							
122	2.56	2.61	2.84							
124	2.23		1.96							
125								1.55	1.87	1.96
132			1.63							
134	2.44	2.64	2.81							
136	2.37	2.63	2.73							
137		2.36	2.24							
138	2.23	2.41	2.15							
140	1.93									
146		1.64								
150	2.25	2.33	2.63							
154		2.02	2.07			1.31				
158		1.75	2.17							
161								1.97		
162	2.01	2.01								
164	2.67	2.69	2.96							
166	2.33					1.42				
168								1.65		
170								1.97		
175										2.00
178	2.46	2.39	2.63							
180	1.85		1.92							
183										2.57
186		1.97								
188	2.15		1.99							
192	2.43	2.42	2.89							
193				2.41						
194		2.13	2.07							
197					2.70			2.76		
198					2.09	1.64	1.82			

Key 1 2 3 4 5 6 7

Supplementary Table 5: Product ion array for bacteria infected HCT116 cells. The headspace above infected HCT116 cultures was analysed via SIFT-MS after a 4-hour bacterial co-culture. Tables shows the product ions detected from each cell line (n=3) with mean intensity log(cps) for each product ion included in analysis after background subtraction using non-infected HCT116 cells as a control. Table is continued on the following page.

m/z		B. fragilis			E. coli Nissle			E. faecalis			F. nucleatum		
	15	3.62	3.56	3.57	3.67	3.60	3.57	3.64	3.61	3.60	3.64	3.59	3.61
	16	4.07	3.98	3.95	4.07	4.00	3.96	4.05	4.01	3.98	4.05	4.00	3.98
	23	1.67	1.67			1.73					1.73	1.60	
	24	2.77	2.56	2.59	2.81	2.67	2.63	2.75	2.70	2.53	2.76	2.65	2.63
	25	3.10	3.02	3.07	3.07	2.92	3.06	3.07	2.96	3.06	3.06	3.02	3.04
	26	2.22	2.08	2.27	2.24	1.87	2.04	2.21	2.20	2.19	2.17	2.20	2.01
	27	4.93	4.87	4.92	4.95	4.87	4.92	4.96	4.88	4.92	4.96	4.87	4.92
	28	3.17	3.09	3.20	3.07	3.00	3.15	3.08	3.07	3.15	3.15	3.01	3.20
	32	6.64	6.49	6.56	6.59	6.47	6.53	6.57	6.46	6.53	6.57	6.46	6.53
	33	3.51	3.25	3.47	3.39	3.23	3.39	3.33	3.12	3.41	3.40	3.15	3.42
	34	4.12	3.94	4.10	4.08	3.91	4.04	4.04	3.94	4.04	4.06	3.94	4.04
	35		1.31									1.11	
	40	3.23	3.12	3.11	3.69	3.24	3.22	3.31	3.29	3.29	3.25	3.19	3.17
	41	2.86	2.86	3.02	2.90	2.76	2.96	2.87	2.92	2.98	2.81	2.93	2.89
	42	3.78	3.58	4.02	4.14	3.72	4.01	3.67	3.58	3.93	3.68	3.56	3.96
	44	2.93	3.23			3.25	2.71	2.50	3.31	2.63	3.97	3.80	3.81
	50	5.84	5.60	5.75	5.72	5.53	5.68	5.68	5.51	5.66	5.69	5.50	5.66
m/z	52	3.43	3.23	3.47	3.36	3.21	3.37	3.34	3.30	3.37	3.31	3.20	3.44
	53	0.00	1.87	1.88	2.12		1.93	1.93			2.17		1.88
	54	2.66	3.35	2.88	2.78	3.33	2.87	2.68	3.32	2.90	2.70	3.37	2.78
	58	2.81	2.86	2.53	3.19	2.85	2.79	2.91	2.94	2.73	3.26	3.19	3.17
	60	3.56	3.35	3.90	3.94	3.48	3.85	3.44	3.29	3.76	3.45	3.35	3.78
	62											1.96	2.12
	64		2.09	2.50			2.58		2.04	2.48			2.45
	66	2.90	2.63	2.87	2.91	2.68	2.70	2.83	3.03	2.58	2.79	2.62	2.67
	67		1.92	2.16	2.15		2.15	2.21	1.96	2.17			
	68	2.78	2.56	2.69	2.56	2.53	2.54	2.59	2.48	2.44	2.57	2.49	2.71
	69	2.69	2.83	3.05	2.76	2.98	3.00	2.76	2.79	2.93	2.84	2.98	3.10
	70	2.35	2.34	2.36	2.45	2.13	2.20	2.46	2.35	2.30	2.42	2.31	2.25
	72		2.25						2.24				
	76	2.07	1.94	2.00	2.12		2.03	1.80	1.92	1.78	2.09		2.00
	78	2.28	1.96	2.44	2.37	1.85	2.33	2.22	2.36	2.38	2.23	2.07	2.21
	80	2.86	2.66	2.95	2.97	2.58	2.82	2.80		2.62	2.72	2.66	2.66
	82		1.82	2.05			2.01	2.08		2.00	2.08		2.06
	84	2.23		2.03		1.74	1.77	2.83	3.23	2.60	1.96		1.96
	86	0.00	2.01		2.24		2.09		2.40		2.53	2.35	2.25
	88	2.16	2.02	2.08	2.36		2.14			2.25	2.30		
	90			1.78	1.81		1.72						1.91
	92	2.24		2.07				2.23	2.09	1.81	1.93	1.91	
	94	2.74	2.55	2.80	2.90	2.58		2.88	2.82	2.65	2.76	2.58	2.70
	96	2.20	2.08	2.21	2.23	2.07			2.01	2.29	2.22	1.89	2.32
	97					2.10						1.96	2.10
	98								2.01				

100	2.94	2.99	2.73	2.82	2.93	2.96	2.70	2.95	2.86	2.79	2.95	2.96
102	2.27											
104	1.89		1.73	1.81	2.02			1.99		1.94		
108	2.87	2.50	2.88	2.91	2.69	2.78	2.81	2.88	2.59	2.73	2.65	2.70
110	1.97		1.80	2.07		2.08	2.15		2.06	2.39		1.86
112	2.08				2.10	1.97	2.06	2.05	2.19		1.98	2.26
114			1.87									
116	2.70	3.01	3.08	2.81	3.09	3.06	2.85	3.10	3.04	2.88	3.01	3.04
118	1.71	1.78	1.83	2.08	1.88	1.91	1.85		1.81			1.93
120	2.34		2.12		1.72		2.06		1.82		1.76	
122	2.98	2.58	2.91	2.96	2.65		2.77	3.06	2.51	2.86	2.74	2.59
123				1.67								
124	2.13	1.84	1.96	2.10	1.97	1.58		1.99	2.07	2.03	1.78	1.89
128							1.86					
130		1.79									1.37	
132	1.74	1.80			1.70		1.70				1.57	
134	2.63	2.73	2.70	2.56		2.68	2.50	2.70	2.72	2.64	2.62	2.69
136	2.74	2.52	2.60	2.79	2.46	2.65	2.71	2.73	2.35	2.62	2.46	2.56
137	2.28			2.07				2.25		2.29	1.92	2.12	2.31
138	2.36	2.27				2.18	2.39	2.23	2.09	2.44		2.00
144								1.60				
146				2.04					1.36		1.49	
148	1.91	1.50		1.63	1.84		1.99			1.50	1.84	
150	2.75	2.45	2.59	2.74	2.51	2.56	2.67	2.73	2.20	2.66	2.29	2.35
151			1.86									
152	2.26		2.12	2.17	2.19	1.91	2.25		1.74	2.18		
154	1.89	2.10			1.73			2.10		2.17	1.85	1.90
160			1.67		1.42							
162	2.19		1.96	2.12		2.22						2.05
164	3.04	2.53	2.97	3.09	2.68	2.84	3.07		2.52	2.92	2.73	2.74
165	2.23										1.94	2.02
166	2.42	2.22	2.36		2.20	2.08	2.54	2.35	2.13	2.41	2.31	2.22
168	2.20			2.12			1.75				2.26	
172						1.97			1.71			
176	2.04											
178	2.76	2.49	2.59	2.70	2.61		2.71	2.84		2.46	2.31	2.45
179				1.96								
180	2.37		2.00	2.15		2.11			1.72	2.35	1.94	
181												
182		1.75			1.75						2.23	
186	2.07			2.02						1.85		
188												2.20
190			1.94				2.27					
192	3.01	2.62	2.74	2.93	2.58	2.68	2.92	2.98	2.49	2.94	2.59	2.78
194				1.97	1.91	2.03				1.83		

Key   1   2   3   4   5   6   7

Chapter 8 – References

Aagaard, K. *et al.* (2014) 'The placenta harbors a unique microbiome', *Science Translational Medicine*. American Association for the Advancement of Science, 6(237), pp. 237ra65-237ra65. doi: 10.1126/scitranslmed.3008599.

Abdulkareem, A. A. *et al.* (2018) 'Periodontal pathogens promote epithelial-mesenchymal transition in oral squamous carcinoma cells in vitro', *Cell Adhesion and Migration*. Taylor and Francis Inc., 12(2), pp. 127–137. doi: 10.1080/19336918.2017.1322253.

Abed, J. *et al.* (2016) 'Fap2 Mediates Fusobacterium nucleatum Colorectal Adenocarcinoma Enrichment by Binding to Tumor-Expressed Gal-GalNAc', *Cell Host and Microbe*. Cell Press, 20(2), pp. 215–225. doi: 10.1016/j.chom.2016.07.006.

Adelstein, B. A. *et al.* (2011) 'Most bowel cancer symptoms do not indicate colorectal cancer and polyps: A systematic review', *BMC Gastroenterology*. BioMed Central, 11(1), p. 65. doi: 10.1186/1471-230X-11-65.

Adlercreutz, H. (1990) 'Western diet and western diseases: Some hormonal and biochemical mechanisms and associations', *Scandinavian Journal of Clinical and Laboratory Investigation*. Informa Healthcare, 50(s201), pp. 3–23. doi: 10.1080/00365519009085798.

Alhajjar, N. *et al.* (2020) 'Genome-wide mutagenesis identifies factors involved in Enterococcus faecalis vaginal adherence and persistence.', *bioRxiv*. bioRxiv, p. 2020.04.30.069468. doi: 10.1101/2020.04.30.069468.

Altomare, D. F. *et al.* (2013) 'Exhaled volatile organic compounds identify patients with colorectal cancer', *British Journal of Surgery*. Wiley-Blackwell, 100(1), pp. 144–150. doi: 10.1002/bjs.8942.

Amitay, E. L. *et al.* (2017) 'Fusobacterium and colorectal cancer: Causal factor or passenger? Results from a large colorectal cancer screening study', *Carcinogenesis*. Oxford University Press, 38(8), pp. 781–788. doi: 10.1093/carcin/bgx053.

Arnold, M. *et al.* (2017) 'Global patterns and trends in colorectal cancer incidence and mortality', *Gut*. BMJ Publishing Group, 66(4), pp. 683–691. doi: 10.1136/gutjnl-2015-310912.

Aron-Wisnewsky, J. *et al.* (2019) 'Major microbiota dysbiosis in severe obesity: Fate after bariatric surgery', *Gut*. BMJ Publishing Group, 68(1), pp. 70–82. doi: 10.1136/gutjnl-2018-316103.

Arrazuria, R. *et al.* (2018) 'Diet induced changes in the microbiota and cell composition of rabbit gut associated lymphoid tissue (GALT)', *Scientific Reports*. Nature Publishing Group, 8(1), pp. 1–11. doi: 10.1038/s41598-018-32484-1.

Arthur, J. C. *et al.* (2014) 'Microbial genomic analysis reveals the essential role of inflammation in bacteria-induced colorectal cancer', *Nature Communications*. NIH Public Access, 5, p. 4724. doi: 10.1038/ncomms5724.

Audie, J. P. *et al.* (1993) 'Expression of human mucin genes in respiratory, digestive, and reproductive tracts ascertained by in situ hybridization.', *Journal of*

Histochemistry & Cytochemistry. SAGE Publications, 41(10), pp. 1479–1485. doi: 10.1177/41.10.8245407.

Bäckhed, F. *et al.* (2004) 'The gut microbiota as an environmental factor that regulates fat storage', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 101(44), pp. 15718–15723. doi: 10.1073/pnas.0407076101.

Bäckhed, F. *et al.* (2015) 'Dynamics and stabilization of the human gut microbiome during the first year of life', *Cell Host and Microbe*. Cell Press, 17(5), pp. 690–703. doi: 10.1016/j.chom.2015.04.004.

Baird, D. L. H. *et al.* (2020) 'Factors associated with metachronous metastases and survival in locally advanced and recurrent rectal cancer', *BJS Open*. Oxford University Press (OUP), 4(6), pp. 1172–1179. doi: 10.1002/bjs5.50341.

Balamurugan, R. *et al.* (2008) 'Real-time polymerase chain reaction quantification of specific butyrate-producing bacteria, Desulfovibrio and Enterococcus faecalis in the feces of patients with colorectal cancer', *Journal of Gastroenterology and Hepatology (Australia)*. Blackwell Publishing, 23(8 PART1), pp. 1298–1303. doi: 10.1111/j.1440-1746.2008.05490.x.

Barker, N. *et al.* (2007) 'Identification of stem cells in small intestine and colon by marker gene Lgr5', *Nature*, 449(7165), pp. 1003–1007. doi: 10.1038/nature06196.

Barker, N. (2014) 'Adult intestinal stem cells: Critical drivers of epithelial homeostasis and regeneration', *Nature Reviews Molecular Cell Biology*, pp. 19–33. doi: 10.1038/nrm3721.

Bates, R. C., Deleo, M. J. and Mercurio, A. M. (2004) 'The epithelial-mesenchymal transition of colon carcinoma involves expression of IL-8 and CXCR-1-mediated chemotaxis', *Experimental Cell Research*. Academic Press Inc., 299(2), pp. 315–324. doi: 10.1016/j.yexcr.2004.05.033.

Batty, C. A. *et al.* (2015) 'Use of the analysis of the volatile faecal metabolome in screening for colorectal cancer', *PLoS ONE*. Edited by M. Chamaillard. Public Library of Science, 10(6), p. e0130301. doi: 10.1371/journal.pone.0130301.

Baughn, A. D. and Malamy, M. H. (2004) 'The strict anaerobe Bacteroides fragilis grows in and benefits from nanomolar concentrations of oxygen', *Nature*. Nature, 427(6973), pp. 441–444. doi: 10.1038/nature02285.

Behar, S. M. and Briken, V. (2019) 'Apoptosis inhibition by intracellular bacteria and its consequence on host immunity', *Current Opinion in Immunology*. Elsevier Ltd, pp. 103–110. doi: 10.1016/j.coi.2019.05.007.

Bendardaf, R. *et al.* (2008) 'VEGF-1 expression in colorectal cancer is associated with disease localization, stage, and long-term disease-specific survival', *Anticancer Research*, 28(6 B), pp. 3865–3870.

Benedicto, A. *et al.* (2017) 'Decreased expression of the β2 integrin on tumor cells is associated with a reduction in liver metastasis of colorectal cancer in mice', *BMC Cancer*. BioMed Central Ltd., 17(1), p. 827. doi: 10.1186/s12885-017-3823-2.

Benno, Y., Sawada, K. and Mitsuoka, T. (1984) 'The Intestinal Microflora of Infants: Composition of Fecal Flora in Breast-Fed and Bottle-Fed Infants',

232

Microbiology and Immunology. Microbiol Immunol, 28(9), pp. 975–986. doi: 10.1111/j.1348-0421.1984.tb00754.x.

Bernstein, H. *et al.* (2009) 'Bile acids as endogenous etiologic agents in gastrointestinal cancer', *World Journal of Gastroenterology*. Baishideng Publishing Group Inc, pp. 3329–3340. doi: 10.3748/wjg.15.3329.

Bertuccini, L. *et al.* (2002) 'Invasion of HeLa cells by Enterococcus faecalis clinical isolates', *Medical Microbiology and Immunology*. Med Microbiol Immunol, 191(1), pp. 25–31. doi: 10.1007/s00430-002-0115-4.

Billig, S. *et al.* (2017) 'Lactate oxidation facilitates growth of Mycobacterium tuberculosis in human macrophages', *Scientific Reports*. Nature Publishing Group, 7(1), pp. 1–12. doi: 10.1038/s41598-017-05916-7.

Bird, A. R. *et al.* (2010) 'Resistant starch, large bowel fermentation and a broader perspective of prebiotics and probiotics', *Beneficial Microbes*, 1(4), pp. 423–431. doi: 10.3920/BM2010.0041.

Bird, R. P. (1987) 'Observation and quantification of aberrant crypts in the murine colon treated with a colon carcinogen: Preliminary findings', *Cancer Letters*. Cancer Lett, 37(2), pp. 147–151. doi: 10.1016/0304-3835(87)90157-1.

Bisson, J. A. *et al.* (2015) 'Wnt5a and Wnt11 inhibit the canonical Wnt pathway and promote cardiac progenitor development via the Caspase-dependent degradation of AKT', *Developmental Biology*. Academic Press Inc., 398(1), pp. 80–96. doi: 10.1016/j.ydbio.2014.11.015.

Biton, M. *et al.* (2018) 'T helper cell cytokines modulate intestinal stem cell renewal and differentiation', *Cell*. NIH Public Access, 175(5), p. 1307. doi: 10.1016/J.CELL.2018.10.008.

Blander, J. M. (2016) 'Death in the intestinal epithelium—basic biology and implications for inflammatory bowel disease', *FEBS Journal*. Blackwell Publishing Ltd, pp. 2720–2730. doi: 10.1111/febs.13771.

Bokulich, N. A. *et al.* (2016) 'Antibiotics, birth mode, and diet shape microbiome maturation during early life', *Science Translational Medicine*. American Association for the Advancement of Science, 8(343), pp. 343ra82-343ra82. doi: 10.1126/scitranslmed.aad7121.

Boldrick, J. C. *et al.* (2002) 'Stereotyped and specific gene expression programs in human innate immune responses to bacteria', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 99(2), pp. 972–977. doi: 10.1073/pnas.231625398.

Boleij, A. *et al.* (2011) 'Novel clues on the specific association of Streptococcus gallolyticus subsp gallolyticus with colorectal cancer', *Journal of Infectious Diseases*. J Infect Dis, 203(8), pp. 1101–1109. doi: 10.1093/infdis/jiq169.

Bond, A. *et al.* (2019) 'Volatile organic compounds emitted from faeces as a biomarker for colorectal cancer', *Alimentary Pharmacology and Therapeutics*. Blackwell Publishing Ltd, 49(8), pp. 1005–1012. doi: 10.1111/apt.15140.

Boonanantanasarn, K. *et al.* (2012) 'Enterococcus faecalis enhances cell proliferation through hydrogen peroxide-mediated epidermal growth factor receptor activation',

Infection and Immunity. American Society for Microbiology Journals, 80(10), pp. 3545–3558. doi: 10.1128/IAI.00479-12.

Bowden, S. D. *et al.* (2009) 'Glucose and glycolysis are required for the successful infection of macrophages and mice by Salmonella enterica serovar Typhimurium', *Infection and Immunity*. American Society for Microbiology Journals, 77(7), pp. 3117–3126. doi: 10.1128/IAI.00093-09.

Brabletz, T. *et al.* (2018) 'EMT in cancer', *Nature Reviews Cancer*. Nature Publishing Group, pp. 128–134. doi: 10.1038/nrc.2017.118.

Brader, P. *et al.* (2008) 'Escherichia coli Nissle 1917 facilitates tumor detection by positron emission tomography and optical imaging', *Clinical Cancer Research*. American Association for Cancer Research, 14(8), pp. 2295–2302. doi: 10.1158/1078-0432.CCR-07-4254.

Bradford, M. M. (1976) 'A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding', *Analytical Biochemistry*. Academic Press, 72(1–2), pp. 248–254. doi: 10.1016/0003-2697(76)90527-3.

Brahimi-Horn, M. C., Chiche, J. and Pouysségur, J. (2007) 'Hypoxia and cancer', *Journal of Molecular Medicine*. J Mol Med (Berl), pp. 1301–1307. doi: 10.1007/s00109-007-0281-3.

Brand, A. *et al.* (2016) 'LDHA-Associated Lactic Acid Production Blunts Tumor Immunosurveillance by T and NK Cells', *Cell Metabolism*. Cell Press, 24(5), pp. 657–671. doi: 10.1016/j.cmet.2016.08.011.

Brattain, M. G. *et al.* (1981) 'Heterogeneity of Malignant Cells from a Human Colonic Carcinoma', *Cancer Research*, 41(5), pp. 1751–1756. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7214343 (Accessed: 2 May 2019).

Bretland, A. J., Lawry, J. and Sharrard, R. M. (2001) 'A study of death by anoikis in cultured epithelial cells', *Cell Proliferation*. Cell Prolif, 34(4), pp. 199–210. doi: 10.1046/j.1365-2184.2001.00198.x.

Buc, E. *et al.* (2013) 'High Prevalence of Mucosa-Associated E. coli Producing Cyclomodulin and Genotoxin in Colon Cancer', *PLoS ONE*. Edited by J. R. Battista. Public Library of Science, 8(2), p. e56964. doi: 10.1371/journal.pone.0056964.

Budinska, E. *et al.* (2013) 'Gene expression patterns unveil a new level of molecular heterogeneity in colorectal cancer', *Journal of Pathology*. J Pathol, 231(1), pp. 63–76. doi: 10.1002/path.4212.

Bullman, S. *et al.* (2017) 'Analysis of Fusobacterium persistence and antibiotic response in colorectal cancer', *Science*, 358(6369), pp. 1443–1448. doi: 10.1126/science.aal5240.

Bundgaard-Nielsen, C. *et al.* (2019) 'The presence of bacteria varies between colorectal adenocarcinomas, precursor lesions and non-malignant tissue', *BMC Cancer*. BioMed Central Ltd., 19(1). doi: 10.1186/s12885-019-5571-y.

Cairns, R. A., Harris, I. S. and Mak, T. W. (2011) *Regulation of cancer cell metabolism*, *Nature Reviews Cancer*. Nature Publishing Group. doi: 10.1038/nrc2981.

Calderwood, A. H., Lasser, K. E. and Roy, H. K. (2016) 'Colon Adenoma Features and Their Impact on Risk of Future Advanced Adenomas and Colorectal Cancer', *World Journal of Gastrointestinal Oncology*. Baishideng Publishing Group Co, 8(12), pp. 826–834. doi: 10.4251/wjgo.v8.i12.826.

Campo, E. *et al.* (1994) 'Cathepsin B expression in colorectal carcinomas correlates with tumor progression and shortened patient survival', *American Journal of Pathology*. American Society for Investigative Pathology, 145(2), pp. 301–309. Available at: /pmc/articles/PMC1887383/?report=abstract (Accessed: 8 April 2021).

Cancer Genome Atlas Network (2012) 'Comprehensive molecular characterization of human colon and rectal cancer', *Nature*. Nature Publishing Group, 487(7407), pp. 330–337. doi: 10.1038/nature11252.

Candela, M. *et al.* (2014) 'Inflammation and colorectal cancer, when microbiota-host mutualism breaks', *World Journal of Gastroenterology*, 20(4), pp. 908–922. doi: 10.3748/wjg.v20.i4.908.

Cao, F. *et al.* (2018) 'Fibroblast activation protein-α in tumor cells promotes colorectal cancer angiogenesis via the Akt and ERK signaling pathways', *Molecular Medicine Reports*. Spandidos Publications, 17(2), pp. 2593–2599. doi: 10.3892/mmr.2017.8155.

Cao, Y. *et al.* (2018) 'Long-term use of antibiotics and risk of colorectal adenoma', *Gut*, 67(4), pp. 672–678. doi: 10.1136/gutjnl-2016-313413.

Castellarin, M. *et al.* (2012) 'Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma', *Genome Research*. Cold Spring Harbor Laboratory Press, 22(2), pp. 299–306. doi: 10.1101/gr.126516.111.

Cebra, J. J. (1999) 'Influences of microbiota on intestinal immune system development', in *American Journal of Clinical Nutrition*. American Society for Nutrition. doi: 10.1093/ajcn/69.5.1046s.

Cha, P. H. *et al.* (2021) 'APC loss induces Warburg effect via increased PKM2 transcription in colorectal cancer', *British Journal of Cancer*. Springer Nature, 124(3), pp. 634–644. doi: 10.1038/s41416-020-01118-7.

Chandran, D. *et al.* (2019) 'The use of selected ion flow tube-mass spectrometry technology to identify breath volatile organic compounds for the detection of head and neck squamous cell carcinoma: A pilot study', *Medicina (Lithuania)*. MDPI AG, 55(6). doi: 10.3390/medicina55060306.

Chen, J., Domingue, J. C. and Sears, C. L. (2017) 'Microbiota dysbiosis in select human cancers: Evidence of association and causality', *Seminars in Immunology*. Academic Press, pp. 25–34. doi: 10.1016/j.smim.2017.08.001.

Chen, X. *et al.* (2018) 'Supervillin promotes epithelial-mesenchymal transition and metastasis of hepatocellular carcinoma in hypoxia via activation of the RhoA/ROCK-ERK/p38 pathway', *Journal of Experimental and Clinical Cancer Research*. BioMed Central Ltd., 37(1), p. 128. doi: 10.1186/s13046-018-0787-2.

Chen, Y. *et al.* (2017) 'Invasive *Fusobacterium* nucleatum activates beta-catenin signaling in colorectal cancer via a TLR4/P-PAK1 cascade', *Oncotarget*, 8(19). doi: 10.18632/oncotarget.15992.

Chénard, T. *et al.* (2020) 'The influence of blood on the human gut microbiome', *BMC Microbiology*. BioMed Central Ltd., 20(1), p. 44. doi: 10.1186/s12866-020-01724-8.

Cheng, R. *et al.* (2014) 'Wnt5a suppresses colon cancer by inhibiting cell proliferation and epithelial-mesenchymal transition', *Journal of Cellular Physiology*. Wiley-Liss Inc., 229(12), pp. 1908–1917. doi: 10.1002/jcp.24566.

Chmelař, D. *et al.* (2016) 'The effect of oxygen on endotoxin production in bacteria of the Bacteroides fragilis group isolated from patients with colorectal carcinoma', *Epidemiologie, Mikrobiologie, Imunologie*, 65(2), pp. 129–135. Available at: https://pubmed.ncbi.nlm.nih.gov/27467330/ (Accessed: 7 April 2021).

Chrysostomou, A. *et al.* (2018) 'Cervical Cancer Screening Programs in Europe: The Transition Towards HPV Vaccination and Population-Based HPV Testing', *Viruses*. MDPI AG, 10(12), p. 729. doi: 10.3390/v10120729.

Chu, H. and Mazmanian, S. K. (2013) 'Innate immune recognition of the microbiota promotes host-microbial symbiosis', *Nature Immunology*. NIH Public Access, pp. 668–675. doi: 10.1038/ni.2635.

Clark, E. S. and Weaver, A. M. (2008) 'A new role for cortactin in invadopodia: Regulation of protease secretion', *European Journal of Cell Biology*. NIH Public Access, 87(8–9), pp. 581–590. doi: 10.1016/j.ejcb.2008.01.008.

Clevers, H. (2006) 'Wnt/β-Catenin Signaling in Development and Disease', *Cell*. Cell, pp. 469–480. doi: 10.1016/j.cell.2006.10.018.

Clevers, H. (2013) 'The intestinal crypt, a prototype stem cell compartment', *Cell*. Cell Press, p. 274. doi: 10.1016/j.cell.2013.07.004.

Clevers, H. and Nusse, R. (2012) 'Wnt/β-Catenin Signaling and Disease', *Cell*, 149(6), pp. 1192–1205. doi: 10.1016/j.cell.2012.05.012.

Cole, S. R. *et al.* (2003) 'A Randomised Trial of the Impact of New Faecal Haemoglobin Test Technologies on Population Participation in Screening for Colorectal Cancer', *Journal of Medical Screening*. SAGE Publications, 10(3), pp. 117–122. doi: 10.1177/096914130301000304.

Corley, D. A. *et al.* (2013) 'Variation of Adenoma Prevalence by Age, Sex, Race, and Colon Location in a Large Population: Implications for Screening and Quality Programs', *Clinical Gastroenterology and Hepatology*. NIH Public Access, 11(2), pp. 172–180. doi: 10.1016/j.cgh.2012.09.010.

Cougnoux, A. *et al.* (2014) 'Bacterial genotoxin colibactin promotes colon tumour growth by inducing a senescence-associated secretory phenotype', *Gut*, 63(12), pp. 1932–1942. doi: 10.1136/gutjnl-2013-305257.

Cougnoux, A. *et al.* (2016) 'Small-molecule inhibitors prevent the genotoxic and protumoural effects induced by colibactin-producing bacteria', *Gut*. BMJ Publishing Group, 65(2), pp. 278–285. doi: 10.1136/gutjnl-2014-307241.

Crowley, J. L. *et al.* (2009) 'Supervillin reorganizes the actin cytoskeleton and increases invadopodial efficiency', *Molecular Biology of the Cell*. American Society for Cell Biology, 20(3), pp. 948–962. doi: 10.1091/mbc.E08-08-0867.

Czyz, D. M., Willett, J. W. and Crosson, S. (2017) 'Brucella abortus induces a Warburg shift in host metabolism that is linked to enhanced intracellular survival of the pathogen', *Journal of Bacteriology*. American Society for Microbiology (ASM), 199(15). doi: 10.1128/JB.00227-17.

Dallosso, A. R. *et al.* (2012) 'Long-range epigenetic silencing of chromosome 5q31 protocadherins is involved in early and late stages of colorectal tumorigenesis through modulation of oncogenic pathways', *Oncogene*. Europe PMC Funders, 31(40), pp. 4409–4419. doi: 10.1038/onc.2011.609.

Dalmasso, G. *et al.* (2015) 'The bacterial genotoxin colibactin promotes colon tumor growth by modifying the tumor microenvironment', *Gut Microbes*. Taylor & Francis, 5(5), pp. 675–680. doi: 10.4161/19490976.2014.969989.

David, L. A. *et al.* (2014) 'Diet rapidly and reproducibly alters the human gut microbiome', *Nature*. NIH Public Access, 505(7484), pp. 559–563. doi: 10.1038/nature12820.

Degirmenci, B. *et al.* (2018) 'GLI1-expressing mesenchymal cells form the essential Wnt-secreting niche for colon stem cells', *Nature*. Nature Publishing Group, 558(7710), pp. 449–453. doi: 10.1038/s41586-018-0190-3.

Dejea, C. M. *et al.* (2018) 'Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria', *Science*. American Association for the Advancement of Science, 359(6375), pp. 592–597. doi: 10.1126/science.aah3648.

Deng, C., Zhang, X. and Li, N. (2004) 'Investigation of volatile biomarkers in lung cancer blood using solid-phase microextraction and capillary gas chromatography-mass spectrometry', *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*. Elsevier, 808(2), pp. 269–277. doi: 10.1016/j.jchromb.2004.05.015.

Deplanche, M. *et al.* (2019) 'Staphylococcus aureus induces DNA damage in host cell', *Scientific Reports*. Nature Publishing Group, 9(1), pp. 1–13. doi: 10.1038/s41598-019-44213-3.

Deriu, E. *et al.* (2013) 'Probiotic Bacteria Reduce Salmonella Typhimurium Intestinal Colonization by Competing for Iron', *Cell Host & Microbe*, 14(1). doi: 10.1016/j.chom.2013.06.007.

Derrien, M. and van Hylckama Vlieg, J. E. T. (2015) 'Fate, activity, and impact of ingested bacteria within the human gut microbiota', *Trends in Microbiology*. Elsevier Ltd, pp. 354–366. doi: 10.1016/j.tim.2015.03.002.

Desai, M. S. *et al.* (2016) 'A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility', *Cell*. Cell Press, 167(5), pp. 1339-1353.e21. doi: 10.1016/j.cell.2016.10.043.

Devaux, C. A., Mezouar, S. and Mege, J. L. (2019) 'The E-Cadherin Cleavage Associated to Pathogenic Bacteria Infections Can Favor Bacterial Invasion and Transmigration, Dysregulation of the Immune Response and Cancer Induction in Humans', *Frontiers in Microbiology*. Frontiers Media S.A., p. 2598. doi: 10.3389/fmicb.2019.02598.

Devic, S. (2016) 'Warburg effect - a consequence or the cause of carcinogenesis?', *Journal of Cancer*, pp. 817–822. doi: 10.7150/jca.14274.

Dillekås, H., Rogers, M. S. and Straume, O. (2019) 'Are 90% of deaths from cancer caused by metastases?', *Cancer Medicine*. Blackwell Publishing Ltd, 8(12), pp. 5574–5576. doi: 10.1002/cam4.2474.

Ding, T. and Schloss, P. D. (2014) 'Dynamics and associations of microbial community types across the human body', *Nature*. Nature Publishing Group, 509(7500), pp. 357–360. doi: 10.1038/nature13178.

Donaldson, G. P., Lee, S. M. and Mazmanian, S. K. (2015) 'Gut biogeography of the bacterial microbiota', *Nature Reviews Microbiology*, pp. 20–32. doi: 10.1038/nrmicro3552.

Donohoe, D. R. *et al.* (2012) 'The Warburg Effect Dictates the Mechanism of Butyrate-Mediated Histone Acetylation and Cell Proliferation', *Molecular Cell*. NIH Public Access, 48(4), pp. 612–626. doi: 10.1016/j.molcel.2012.08.033.

Doran, D. M. *et al.* (2011) 'Hypoxia activates constitutive luciferase reporter constructs', *Biochimie*. NIH Public Access, 93(2), pp. 361–368. doi: 10.1016/j.biochi.2010.10.009.

Dravis, C. *et al.* (2015) 'Sox10 Regulates Stem/Progenitor and Mesenchymal Cell States in Mammary Epithelial Cells', *Cell Reports*. Elsevier B.V., 12(12), pp. 2035–2048. doi: 10.1016/j.celrep.2015.08.040.

Drewes, J. L. *et al.* (2017) 'High-resolution bacterial 16S rRNA gene profile meta-analysis and biofilm status reveal common colorectal cancer consortia', *npj Biofilms and Microbiomes*. Nature Publishing Group, 3(1), pp. 1–12. doi: 10.1038/s41522-017-0040-3.

Durai, R. *et al.* (2007) 'Role of insulin-like growth factor binding protein-4 in prevention of colon cancer', *World Journal of Surgical Oncology*, 5, pp. 128–128. doi: 10.1186/1477-7819-5-128.

Edwards, A. M. and Massey, R. C. (2011) 'Invasion of human cells by a bacterial pathogen', *Journal of Visualized Experiments*. Journal of Visualized Experiments, 49(49), p. 2693. doi: 10.3791/2693.

Ellis, H. and Mahadevan, V. (2014) 'Anatomy of the caecum, appendix and colon', *Surgery (United Kingdom)*. Elsevier Ltd, pp. 155–158. doi: 10.1016/j.mpsur.2014.02.001.

Escoll, P. and Buchrieser, C. (2018) 'Metabolic reprogramming of host cells upon bacterial infection: Why shift to a Warburg-like metabolism?', *FEBS Journal*, pp. 2146–2160. doi: 10.1111/febs.14446.

Faïs, T. *et al.* (2018) 'Colibactin: More Than a New Bacterial Toxin', *Toxins*, 10(4), p. 151. doi: 10.3390/toxins10040151.

Faith, J. J. *et al.* (2013) 'The long-term stability of the human gut microbiota', *Science*. American Association for the Advancement of Science, 341(6141). doi: 10.1126/science.1237439.

Fearnhead, N. S., Britton, M. P. and Bodmer, W. F. (2001) 'The ABC of APC',

Human Molecular Genetics. Oxford University Press, pp. 721–733. doi: 10.1093/hmg/10.7.721.

Fearon, E. R. and Vogelstein, B. (1990) 'A genetic model for colorectal tumorigenesis', *Cell*. Cell Press, pp. 759–767. doi: 10.1016/0092-8674(90)90186-I.

Feng, Q. *et al.* (2015) 'Gut microbiome development along the colorectal adenoma-carcinoma sequence', *Nature Communications*, 6. doi: 10.1038/ncomms7528.

Fernández, L. *et al.* (2013) 'The human milk microbiota: Origin and potential roles in health and disease', *Pharmacological Research*. Academic Press, pp. 1–10. doi: 10.1016/j.phrs.2012.09.001.

Filipe, M. I. and Branfoot, A. C. (1974) 'Abnormal patterns of mucus secretion in apparently normal mucosa of large intestine with carcinoma', *Cancer*. John Wiley & Sons, Ltd, 34(2), pp. 282–290. doi: 10.1002/1097-0142(197408)34:2<282::AID-CNCR2820340211>3.0.CO;2-W.

Flanagan, L. *et al.* (2014) 'Fusobacterium nucleatum associates with stages of colorectal neoplasia development, colorectal cancer and disease outcome', *European Journal of Clinical Microbiology and Infectious Diseases*. Springer Berlin Heidelberg, 33(8), pp. 1381–1390. doi: 10.1007/s10096-014-2081-3.

Flemer, B. *et al.* (2017) 'Tumour-associated and non-tumour-associated microbiota in colorectal cancer', *Gut*, 66(4), pp. 633–643. doi: 10.1136/gutjnl-2015-309595.

Flint, H. J. *et al.* (2007) 'Interactions and competition within the microbial community of the human colon: Links between diet and health: Minireview', *Environmental Microbiology*. Environ Microbiol, pp. 1101–1111. doi: 10.1111/j.1462-2920.2007.01281.x.

Flint, H. J. *et al.* (2012) 'Microbial degradation of complex carbohydrates in the gut', *Gut Microbes*. Taylor & Francis, pp. 289–306. doi: 10.4161/gmic.19897.

Flood, D. M. *et al.* (2000) 'Colorectal cancer incidence in Asian migrants to the United States and their descendants', *Cancer Causes and Control*. Cancer Causes Control, 11(5), pp. 403–411. doi: 10.1023/A:1008955722425.

Fordtran, J. S. and Dietschy, J. M. (1966) 'Water and Electrolyte Movement in the Intestine', *Gastroenterology*, 50(2), pp. 263–285. doi: 10.1016/S0016-5085(66)80060-4.

Frey, P. *et al.* (2020) 'Canonical BMP signaling executes epithelial-mesenchymal transition downstream of snail1', *Cancers*. MDPI AG, 12(4), p. 1019. doi: 10.3390/cancers12041019.

Friedl, P. and Wolf, K. (2003) 'Tumour-cell invasion and migration: Diversity and escape mechanisms', *Nature Reviews Cancer*. Nature Publishing Group, pp. 362–374. doi: 10.1038/nrc1075.

Fukumoto, M. *et al.* (2015) 'α-Actinin-4 enhances colorectal cancer cell invasion by suppressing focal adhesion maturation', *PLoS ONE*. Public Library of Science, 10(4). doi: 10.1371/journal.pone.0120616.

Ganesh, K. and Massagué, J. (2021) 'Targeting metastatic cancer', *Nature Medicine*. Nature Research, pp. 34–44. doi: 10.1038/s41591-020-01195-4.

Gantois, I. *et al.* (2006) 'Butyrate specifically down-regulates Salmonella pathogenicity island 1 gene expression', *Applied and Environmental Microbiology.* Appl Environ Microbiol, 72(1), pp. 946–949. doi: 10.1128/AEM.72.1.946-949.2006.

Gargini, R. *et al.* (2016) 'WIP Drives Tumor Progression through YAP/TAZ-Dependent Autonomous Cell Growth', *Cell Reports.* Elsevier B.V., 17(8), pp. 1962–1977. doi: 10.1016/j.celrep.2016.10.064.

van der Geest, L. G. M. *et al.* (2015) 'Nationwide trends in incidence, treatment and survival of colorectal cancer patients with synchronous metastases', *Clinical and Experimental Metastasis.* Kluwer Academic Publishers, 32(5), pp. 457–465. doi: 10.1007/s10585-015-9719-0.

Geng, J. *et al.* (2014) 'Co-occurrence of driver and passenger bacteria in human colorectal cancer', 6, pp. 1–5. doi: 10.1186/1757-4749-6-26.

Del Giudice, M. E. *et al.* (2014) 'Systematic review of clinical features of suspected colorectal cancer in primary care', *Canadian Family Physician.* College of Family Physicians of Canada, 60(8), p. e405. Available at: /pmc/articles/PMC4131977/ (Accessed: 12 February 2021).

Goetze, K. *et al.* (2011) 'Lactate enhances motility of tumor cells and inhibits monocyte migration and cytokine release', *International Journal of Oncology.* Int J Oncol, 39(2), pp. 453–463. doi: 10.3892/ijo.2011.1055.

Gonzalez, H., Hagerling, C. and Werb, Z. (2018) 'Roles of the immune system in cancer: From tumor initiation to metastatic progression', *Genes and Development.* Cold Spring Harbor Laboratory Press, pp. 1267–1284. doi: 10.1101/GAD.314617.118.

Goodrich, J. K. *et al.* (2014) 'Human genetics shape the gut microbiome', *Cell.* Cell Press, 159(4), pp. 789–799. doi: 10.1016/j.cell.2014.09.053.

Gordon, H. A. and Pesti, L. (1971) 'The gnotobiotic animal as a tool in the study of host microbial relationships.', *Bacteriological reviews.* American Society for Microbiology (ASM), pp. 390–429. doi: 10.1128/mmbr.35.4.390-429.1971.

Gorroño-Etxebarria, I. *et al.* (2019) 'Wnt-11 as a potential prognostic biomarker and therapeutic target in colorectal cancer', *Cancers.* MDPI AG, 11(7). doi: 10.3390/cancers11070908.

Gout, S. and Huot, J. (2008) 'Role of cancer microenvironment in metastasis: Focus on colon cancer', *Cancer Microenvironment.* Springer, pp. 69–83. doi: 10.1007/s12307-008-0007-2.

Greaves, M. and Maley, C. C. (2012) 'Clonal evolution in cancer', *Nature.* NIH Public Access, pp. 306–313. doi: 10.1038/nature10762.

Greenhough, A. *et al.* (2010) 'The proapoptotic BH3-only protein Bim is downregulated in a subset of colorectal cancers and is repressed by antiapoptotic COX-2/PGE2 signalling in colorectal adenoma cells', *Oncogene.* Europe PMC Funders, 29(23), pp. 3398–3410. doi: 10.1038/onc.2010.94.

Grivennikov, S. I. *et al.* (2012) 'Adenoma-linked barrier defects and microbial products drive IL-23/IL-17-mediated tumour growth', *Nature.* Nature Publishing Group, 491(7423), pp. 254–258. doi: 10.1038/nature11465.

Gronbach, K. *et al.* (2010) 'Safety of probiotic Escherichia coli strain Nissle 1917 depends on intestinal microbiota and adaptive immunity of the host', *Infection and Immunity*. American Society for Microbiology (ASM), 78(7), pp. 3036–3046. doi: 10.1128/IAI.00218-10.

Guenther, K. *et al.* (2010) 'Sever Sepsis After Probiotic Treatment With Escherichia coli NISSLE 1917', *Pediatric Infectious Disease Journal*. Lippincott Williams and Wilkins, 29(2), pp. 188–189. doi: 10.1097/INF.0b013e3181c36eb9.

Guinney, J. *et al.* (2015) 'The consensus molecular subtypes of colorectal cancer', *Nature Medicine*. Nat Med, 21(11), pp. 1350–1356. doi: 10.1038/nm.3967.

Günther, J. and Seyfert, H. M. (2018) 'The first line of defence: insights into mechanisms and relevance of phagocytosis in epithelial cells', *Seminars in Immunopathology*. Springer Verlag, 40(6), pp. 555–565. doi: 10.1007/s00281-018-0701-1.

Günther, S. D. *et al.* (2020) 'Cytosolic Gram-negative bacteria prevent apoptosis by inhibition of effector caspases through lipopolysaccharide', *Nature Microbiology*. Nature Research, 5(2), pp. 354–367. doi: 10.1038/s41564-019-0620-5.

Guo, P. *et al.* (2020) 'FadA promotes DNA damage and progression of Fusobacterium nucleatum-induced colorectal cancer through up-regulation of chk2', *Journal of Experimental & Clinical Cancer Research*. Springer Science and Business Media LLC, 39(1), p. 202. doi: 10.1186/s13046-020-01677-w.

Gursoy, U. K., Könönen, E. and Uitto, V. J. (2008) 'Intracellular replication of fusobacteria requires new actin filament formation of epithelial cells', *APMIS*. John Wiley & Sons, Ltd, 116(12), pp. 1063–1070. doi: 10.1111/j.1600-0463.2008.00868.x.

Hagland, H. R. and Søreide, K. (2015) 'Cellular metabolism in colorectal carcinogenesis: Influence of lifestyle, gut microbiome and metabolic pathways', *Cancer Letters*. Elsevier, pp. 273–280. doi: 10.1016/j.canlet.2014.02.026.

Hague, A. *et al.* (1993) 'Sodium butyrate induces apoptosis in human colonic tumour cell lines in a p53-independent pathway: Implications for the possible role of dietary fibre in the prevention of large-bowel cancer', *International Journal of Cancer*, 55(3), pp. 498–505. doi: 10.1002/ijc.2910550329.

Hajishengallis, G., Darveau, R. P. and Curtis, M. A. (2012) 'The keystone-pathogen hypothesis', *Nature Reviews Microbiology*, pp. 717–725. doi: 10.1038/nrmicro2873.

Hajishengallis, G. and Lamont, R. J. (2016) 'Dancing with the Stars: How Choreographed Bacterial Interactions Dictate Nososymbiocity and Give Rise to Keystone Pathogens, Accessory Pathogens, and Pathobionts', *Trends in Microbiology*. Elsevier Ltd, pp. 477–489. doi: 10.1016/j.tim.2016.02.010.

Hamer, H. M. *et al.* (2008) 'Review article: The role of butyrate on colonic function', *Alimentary Pharmacology and Therapeutics*. John Wiley & Sons, Ltd, pp. 104–119. doi: 10.1111/j.1365-2036.2007.03562.x.

Han, Y. W. *et al.* (2000) 'Interactions between periodontal bacteria and human oral epithelial cells: Fusobacterium nucleatum adheres to and invades epithelial cells', *Infection and Immunity*. American Society for Microbiology Journals, 68(6), pp.

3140–3146. doi: 10.1128/IAI.68.6.3140-3146.2000.

Hanahan, D. and Weinberg, R. A. (2011) 'Hallmarks of cancer: The next generation', *Cell*, pp. 646–674. doi: 10.1016/j.cell.2011.02.013.

Hara-Kudo, Y. *et al.* (2013) 'Prevalence of the main food-borne pathogens in retail food under the national food surveillance system in Japan', *Food Additives & Contaminants: Part A*. Taylor & Francis , 30(8), pp. 1450–1458. doi: 10.1080/19440049.2012.745097.

Hartsock, A. and Nelson, W. J. (2008) 'Adherens and tight junctions: Structure, function and connections to the actin cytoskeleton', *Biochimica et Biophysica Acta - Biomembranes*. NIH Public Access, pp. 660–669. doi: 10.1016/j.bbamem.2007.07.012.

He, P. and Jin, X. (2018) 'SOX10 induces epithelial-mesenchymal transition and contributes to nasopharyngeal carcinoma progression', *Biochemistry and Cell Biology*. Canadian Science Publishing, 96(3), pp. 326–331. doi: 10.1139/bcb-2017-0160.

Hecht, I. *et al.* (2015) 'Tumor invasion optimization by mesenchymal-amoeboid heterogeneity', *Scientific Reports*. Nature Publishing Group, 5. doi: 10.1038/srep10622.

Vander Heiden, M., Cantley, L. and Thompson, C. (2009) 'Understanding the Warburg effect: The metabolic Requirements of cell proliferation', *Science*. NIH Public Access, 324(5930), pp. 1029–1033. doi: 10.1126/science.1160809.Understanding.

Henry, L. R. *et al.* (2007) 'Clinical implications of fibroblast activation protein in patients with colon cancer', *Clinical Cancer Research*. Clin Cancer Res, 13(6), pp. 1736–1741. doi: 10.1158/1078-0432.CCR-06-1746.

Herzer, P. J. *et al.* (1990) 'Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of Escherichia coli', *Journal of Bacteriology*, 172(11), pp. 6175–6181. doi: 10.1128/jb.172.11.6175-6181.1990.

Hirayama, A. *et al.* (2009) 'Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry', *Cancer Research*. American Association for Cancer Research, 69(11), pp. 4918–4925. doi: 10.1158/0008-5472.CAN-08-4806.

Ho, N. T. *et al.* (2018) 'Meta-analysis of effects of exclusive breastfeeding on infant gut microbiota across populations', *Nature Communications*. Nature Publishing Group, 9(1). doi: 10.1038/s41467-018-06473-x.

Holmgren, L., O'reilly, M. S. and Folkman, J. (1995) 'Dormancy of micrometastases: Balanced proliferation and apoptosis in the presence of angiogenesis suppression', *Nature Medicine*. Nat Med, 1(2), pp. 149–153. doi: 10.1038/nm0295-149.

Hooper, L. V and Gordon, J. I. (2001) 'Commensal host-bacterial relationships in the gut', *Science*. American Association for the Advancement of Science, pp. 1115–1118. doi: 10.1126/science.1058709.

Hou, J. M. *et al.* (2012) 'Clinical significance and molecular characteristics of

circulating tumor cells and circulating tumor microemboli in patients with small-cell lung cancer', *Journal of Clinical Oncology*. J Clin Oncol, 30(5), pp. 525–532. doi: 10.1200/JCO.2010.33.3716.

Hsu, H. Y. and Wen, M. H. (2002) 'Lipopolysaccharide-mediated reactive oxygen species and signal transduction in the regulation of interleukin-1 gene expression', *Journal of Biological Chemistry*. Elsevier, 277(25), pp. 22131–22139. doi: 10.1074/jbc.M111883200.

Huang, J. Y., Lee, S. M. and Mazmanian, S. K. (2011) 'The human commensal Bacteroides fragilis binds intestinal mucin', *Anaerobe*, 17(4), pp. 137–141. doi: 10.1016/j.anaerobe.2011.05.017.

Huang, T.-C. *et al.* (2017) 'Distinct roles and differential expression levels of Wnt5a mRNA isoforms in colorectal cancer cells', *PLOS ONE*. Edited by M. Katoh. Public Library of Science, 12(8), p. e0181034. doi: 10.1371/journal.pone.0181034.

Hughes, C. S., Postovit, L. M. and Lajoie, G. A. (2010) 'Matrigel: a complex protein mixture required for optimal growth of cell culture.', *Proteomics*. John Wiley & Sons, Ltd, 10(9), pp. 1886–1890. doi: 10.1002/pmic.200900758.

Hughes, R., Magee, E. A. and Bingham, S. (2000) 'Protein degradation in the large intestine: relevance to colorectal cancer.', *Current issues in intestinal microbiology*, pp. 51–58. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11709869 (Accessed: 28 November 2017).

Huh, J. W. and Roh, T. Y. (2020) 'Opportunistic detection of Fusobacterium nucleatum as a marker for the early gut microbial dysbiosis', *BMC Microbiology*. BioMed Central, 20(1), p. 208. doi: 10.1186/s12866-020-01887-4.

Huycke, M. M., Abrams, V. and Moore, D. R. (2002) 'Enterococcus faecalis produces extracellular superoxide and hydrogen peroxide that damages colonic epithelial cell DNA', *Carcinogenesis*. Oxford University Press, 23(3), pp. 529–536. doi: 10.1093/carcin/23.3.529.

Ibáñez, C. *et al.* (2017) 'GC-MS based metabolomics of colon cancer cells using different extraction solvents', *Analytica Chimica Acta*. Elsevier, 986, pp. 48–56. doi: 10.1016/j.aca.2017.07.019.

Ivanov, I. I. *et al.* (2009) 'Induction of Intestinal Th17 Cells by Segmented Filamentous Bacteria', *Cell*. NIH Public Access, 139(3), pp. 485–498. doi: 10.1016/j.cell.2009.09.033.

Jacob, A. and Prekeris, R. (2015) 'The regulation of MMP targeting to invadopodia during cancer metastasis', *Frontiers in Cell and Developmental Biology*. Frontiers Media S.A. doi: 10.3389/fcell.2015.00004.

Ji, S. *et al.* (2010) 'Intracellular degradation of Fusobacterium nucleatum in human gingival epithelial cells', *Molecules and Cells*, 30(6), pp. 519–526. doi: 10.1007/s10059-010-0142-8.

Jiang, H. and Li, H. (2021) 'Prognostic values of tumoral MMP2 and MMP9 overexpression in breast cancer: a systematic review and meta-analysis', *BMC Cancer*. Springer Science and Business Media LLC, 21(1), p. 149. doi: 10.1186/s12885-021-07860-2.

Johansson, M. E. V. *et al.* (2008) 'The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 105(39), pp. 15064–15069. doi: 10.1073/pnas.0803124105.

Johansson, M. E. V. *et al.* (2014) 'Bacteria penetrate the normally impenetrable inner colon mucus layer in both murine colitis models and patients with ulcerative colitis', *Gut*, 63(2), pp. 281–291. doi: 10.1136/gutjnl-2012-303207.

Johansson, M. E. V. *et al.* (2015) 'Normalization of host intestinal mucus layers requires long-term microbial colonization', *Cell Host and Microbe*. Cell Press, 18(5), pp. 582–592. doi: 10.1016/j.chom.2015.10.007.

Jonsson, H. *et al.* (2020) 'Genome sequence of segmented filamentous bacteria present in the human intestine', *Communications Biology*. Nature Research, 3(1), pp. 1–9. doi: 10.1038/s42003-020-01214-7.

Jung, G. *et al.* (2020) 'Epigenetics of colorectal cancer: biomarker and therapeutic potential', *Nature Reviews Gastroenterology and Hepatology*. Nature Research, pp. 111–130. doi: 10.1038/s41575-019-0230-y.

Kandoth, C. *et al.* (2013) 'Mutational landscape and significance across 12 major cancer types', *Nature*. Nature Publishing Group, 502(7471), pp. 333–339. doi: 10.1038/nature12634.

Kaoutari, A. El *et al.* (2013) 'The abundance and variety of carbohydrate-active enzymes in the human gut microbiota', *NATURE REVIEWS | MICROBIOLOGY*, 11. doi: 10.1038/nrmicro3050.

Kasper, S. H. *et al.* (2020) 'Colorectal cancer-associated anaerobic bacteria proliferate in tumor spheroids and alter the microenvironment', *Scientific Reports*, 10(1). doi: 10.1038/s41598-020-62139-z.

Khanna, C. and Hunter, K. (2005) 'Modeling metastasis in vivo', *Carcinogenesis*. Oxford Academic, pp. 513–523. doi: 10.1093/carcin/bgh261.

Khoursheed, M. A. *et al.* (2003) 'Expression of E-cadherin in human colorectal cancer', *Surgeon*. Edinburgh University Press, 1(2), pp. 86–91. doi: 10.1016/S1479-666X(03)80121-X.

Kiela, P. R. and Ghishan, F. K. (2016) 'Physiology of intestinal absorption and secretion', *Best Practice and Research: Clinical Gastroenterology*. Bailliere Tindall Ltd, pp. 145–159. doi: 10.1016/j.bpg.2016.02.007.

Kim, B. R. *et al.* (2015) 'BMP-2 induces motility and invasiveness by promoting colon cancer stemness through STAT3 activation', *Tumor Biology*. Springer Netherlands, 36(12), pp. 9475–9486. doi: 10.1007/s13277-015-3681-y.

Kim, E. R. and Chang, D. K. (2014) 'Colorectal cancer in inflammatory bowel disease: The risk, pathogenesis, prevention and diagnosis', *World Journal of Gastroenterology*, pp. 9872–9881. doi: 10.3748/wjg.v20.i29.9872.

Kim, J. M. *et al.* (1998) 'Apoptosis of human intestinal epithelial cells after bacterial invasion', *Journal of Clinical Investigation*. The American Society for Clinical Investigation, 102(10), pp. 1815–1823. doi: 10.1172/JCI2466.

Kinzler, K. W. *et al.* (2014) 'Microbiota organization is a distinct feature of proximal colorectal cancers', *Proceedings of the National Academy of Sciences*. National Academy of Sciences, 111(51), pp. 18321–18326. doi: 10.1073/pnas.1406199111.

Korinek, V. *et al.* (1998) 'Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4', *Nature Genetics*. Nature Publishing Group, 19(4), pp. 379–383. doi: 10.1038/1270.

Kornprat, P. *et al.* (2011) 'Value of tumor size as a prognostic variable in colorectal cancer: A critical reappraisal', *American Journal of Clinical Oncology: Cancer Clinical Trials*. Am J Clin Oncol, 34(1), pp. 43–49. doi: 10.1097/COC.0b013e3181cae8dd.

Kostic, A. D. *et al.* (2013) 'Fusobacterium nucleatum Potentiates Intestinal Tumorigenesis and Modulates the Tumor-Immune Microenvironment', *Cell Host and Microbe*. NIH Public Access, 14(2), pp. 207–215. doi: 10.1016/j.chom.2013.07.007.

Kowalczyk, M. *et al.* (2018) 'Occurrence of colorectal aberrant crypt foci depending on age and dietary patterns of patients', *BMC Cancer*. BioMed Central Ltd., 18(1), pp. 1–9. doi: 10.1186/s12885-018-4100-8.

Kruis, W. *et al.* (2004) 'Maintaining remission of ulcerative colitis with the probiotic Escherichia coli Nissle 1917 is as effective as with standard mesalazine', *Gut*. Gut, 53(11), pp. 1617–1623. doi: 10.1136/gut.2003.037747.

Kunzmann, A. T. *et al.* (2015) 'Dietary fiber intake and risk of colorectal cancer and incident and recurrent adenoma in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial', *American Journal of Clinical Nutrition*. American Society for Nutrition, 102(4), pp. 881–890. doi: 10.3945/ajcn.115.113282.

Kurdi, P. *et al.* (2006) 'Mechanism of growth inhibition by free bile acids in lactobacilli and bifidobacteria', *Journal of Bacteriology*. American Society for Microbiology (ASM), 188(5), pp. 1979–1986. doi: 10.1128/JB.188.5.1979-1986.2006.

Kwak, E. L. and Chung, D. C. (2007) 'Hereditary colorectal cancer syndromes: An overview', *Clinical Colorectal Cancer*. Elsevier Inc., pp. 340–344. doi: 10.3816/CCC.2007.n.002.

Kwong, T. N. Y. *et al.* (2018) 'Association Between Bacteremia From Specific Microbes and Subsequent Diagnosis of Colorectal Cancer', *Gastroenterology*. W.B. Saunders, 155(2), pp. 383-390.e8. doi: 10.1053/j.gastro.2018.04.028.

Labi, V. and Erlacher, M. (2015) 'How cell death shapes cancer', *Cell Death and Disease*. Nature Publishing Group, pp. e1675–e1675. doi: 10.1038/cddis.2015.20.

Lang, J. M., Eisen, J. A. and Zivkovic, A. M. (2014) 'The microbes we eat: Abundance and taxonomy of microbes consumed in a day's worth of meals for three diet types', *PeerJ*. PeerJ Inc., 2014(2), p. e659. doi: 10.7717/peerj.659.

Lefebvre, P. *et al.* (2009) 'Role of Bile Acids and Bile Acid Receptors in Metabolic Regulation'. doi: 10.1152/physrev.00010.2008.-The.

Lei, Y. *et al.* (2011) 'Proteomics identification of ITGB3 as a key regulator in reactive oxygen species-induced migration and invasion of colorectal cancer cells',

Molecular and Cellular Proteomics. Mol Cell Proteomics, 10(10). doi: 10.1074/mcp.M110.005397.

Lengauer, C., Kinzler, K. W. and Vogelstein, B. (1997) 'Genetic instability in colorectal cancers', *Nature*. Nature Publishing Group, 386(6625), pp. 623–627. doi: 10.1038/386623a0.

Leschelle, X. *et al.* (2000) 'Butyrate metabolism upstream and downstream acetyl-CoA synthesis and growth control of human colon carcinoma cells', *European Journal of Biochemistry*. Blackwell Science Ltd, 267(21), pp. 6435–6442. doi: 10.1046/j.1432-1327.2000.01731.x.

Leve, F. and Morgado-Díaz, J. A. (2012) 'Rho GTPase signaling in the development of colorectal cancer', *Journal of Cellular Biochemistry*. John Wiley & Sons, Ltd, 113(8), pp. 2549–2559. doi: 10.1002/jcb.24153.

Levine, D. S. and Haggitt, R. C. (1989) 'Normal histology of the colon', *American Journal of Surgical Pathology*. Am J Surg Pathol, pp. 966–984. doi: 10.1097/00000478-198911000-00008.

Li, B. H. *et al.* (2005) 'Matrix metalloproteinase-2 and tissue inhibitor of metalloproteinase-2 in colorectal carcinoma invasion and metastasis', *World Journal of Gastroenterology*. WJG Press, 11(20), pp. 3046–3050. doi: 10.3748/wjg.v11.i20.3046.

Li, Q. and Chen, H. (2012) 'Silencing of wnt5a during colon cancer metastasis involves histone modifications', *Epigenetics*. Taylor and Francis Inc., 7(6), pp. 551–558. doi: 10.4161/epi.20050.

Li, S. K. H. and Martin, A. (2016) 'Mismatch Repair and Colon Cancer: Mechanisms and Therapies Explored', *Trends in Molecular Medicine*, pp. 274–289. doi: 10.1016/j.molmed.2016.02.003.

Li, Z. *et al.* (2017) 'Tissue factor pathway inhibitor-2 induced hepatocellular carcinoma cell differentiation', *Saudi Journal of Biological Sciences*. Elsevier B.V., 24(1), pp. 95–102. doi: 10.1016/j.sjbs.2016.09.003.

Li, Z. L. *et al.* (2020) 'Changes in extracellular matrix in different stages of colorectal cancer and their effects on proliferation of cancer cells', *World Journal of Gastrointestinal Oncology*. Baishideng Publishing Group Co, 12(3), pp. 267–275. doi: 10.4251/WJGO.V12.I3.267.

Liberti, M. V and Locasale, J. W. (2016) 'The Warburg Effect: How Does it Benefit Cancer Cells?', *Trends in Biochemical Sciences*. NIH Public Access, pp. 211–218. doi: 10.1016/j.tibs.2015.12.001.

Lintz, M., Muñoz, A. and Reinhart-King, C. A. (2017) 'The Mechanics of Single Cell and Collective Migration of Tumor Cells', *Journal of Biomechanical Engineering*. American Society of Mechanical Engineers (ASME), p. 0210051. doi: 10.1115/1.4035121.

Liu, Y., Beyer, A. and Aebersold, R. (2016) 'On the Dependency of Cellular Protein Levels on mRNA Abundance', *Cell*, pp. 535–550. doi: 10.1016/j.cell.2016.03.014.

Lloyd-Price, J., Abu-Ali, G. and Huttenhower, C. (2016) 'The healthy human microbiome', *Genome Medicine*. BioMed Central, 8(1), p. 51. doi: 10.1186/s13073-

016-0307-y.

Loeb, L. A., Loeb, K. R. and Anderson, J. P. (2003) 'Multiple mutations and cancer', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, pp. 776–781. doi: 10.1073/pnas.0334858100.

Loening-Baucke, V. *et al.* (2012) 'Mucosal invasion by fusobacteria is a common feature of acute appendicitis in Germany, Russia, and China', *Saudi Journal of Gastroenterology*. Wolters Kluwer -- Medknow Publications, 18(1), pp. 55–58. doi: 10.4103/1319-3767.91734.

Louis, P., Hold, G. L. and Flint, H. J. (2014) 'The gut microbiota, bacterial metabolites and colorectal cancer', *Nature Reviews Microbiology*. Nature Publishing Group, pp. 661–672. doi: 10.1038/nrmicro3344.

Lu, K., Dong, J. L. and Fan, W. J. (2018) 'Twist1/2 activates MMP2 expression via binding to its promoter in colorectal cancer', *European Review for Medical and Pharmacological Sciences*, 22(23), pp. 8210–8219. doi: 10.26355/eurrev_201812_16514.

Lutgens, M. W. M. D. *et al.* (2008) 'High frequency of early colorectal cancer in inflammatory bowel disease', *Gut*. BMJ Publishing Group, 57(9), pp. 1246–1251. doi: 10.1136/gut.2007.143453.

Ma, C. T. *et al.* (2018) 'Fusobacterium nucleatum promotes the progression of colorectal cancer by interacting with e-cadherin', *Oncology Letters*. Spandidos Publications, 16(2), pp. 2606–2612. doi: 10.3892/ol.2018.8947.

Ma, J. *et al.* (1999) 'Prospective study of colorectal cancer risk in men and plasma levels of insulin-like growth factor (IGF)-I and IGF-binding protein-3', *Journal of the National Cancer Institute*. Oxford University Press, 91(7), pp. 620–625. doi: 10.1093/jnci/91.7.620.

Macfarlane, G. T., Cummings, J. H. and Allison, C. (1986) 'Protein degradation by human intestinal bacteria', *Journal of General Microbiology*, 132(6), pp. 1647–1656. doi: 10.1099/00221287-132-6-1647.

Mambu, J. *et al.* (2020) 'Rck of Salmonella Typhimurium Delays the Host Cell Cycle to Facilitate Bacterial Invasion', *Frontiers in Cellular and Infection Microbiology*. Frontiers Media S.A., 10, p. 656. doi: 10.3389/fcimb.2020.586934.

Mamlouk, S. *et al.* (2020) 'Malignant transformation and genetic alterations are uncoupled in early colorectal cancer progression', *BMC Biology*. BioMed Central Ltd, 18(1), p. 116. doi: 10.1186/s12915-020-00844-x.

Martin, H. M. *et al.* (2004) 'Enhanced Escherichia coli adherence and invasion in Crohn's disease and colon cancer', *Gastroenterology*. W.B. Saunders, 127(1), pp. 80–93. doi: 10.1053/j.gastro.2004.03.054.

Martin, R. *et al.* (2010) 'Early life: Gut microbiota and immune development in infancy', *Beneficial Microbes*. Wageningen Academic Publishers, 1(4), pp. 367–382. doi: 10.3920/BM2010.0027.

Martinez-Medina, M. *et al.* (2014) 'Western diet induces dysbiosis with increased e coli in CEABAC10 mice, alters host barrier function favouring AIEC colonisation', *Gut*. BMJ Publishing Group, 63(1), pp. 116–124. doi: 10.1136/gutjnl-2012-304119.

Mazmanian, S. K. *et al.* (2005) 'An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system', *Cell*. Cell, 122(1), pp. 107–118. doi: 10.1016/j.cell.2005.05.007.

McBurney, M. I. *et al.* (2019) 'Establishing What Constitutes a Healthy Human Gut Microbiome: State of the Science, Regulatory Considerations, and Future Directions', *Journal of Nutrition*. Oxford University Press, pp. 1882–1895. doi: 10.1093/jn/nxz154.

McNeil, N. I. (1984) 'The contribution of the large intestine to energy supplies in man', *American Journal of Clinical Nutrition*, 39(2), pp. 338–342. doi: 10.1093/ajcn/39.2.338.

Mehlen, P. and Puisieux, A. (2006) 'Metastasis: A question of life or death', *Nature Reviews Cancer*. Nature Publishing Group, pp. 449–458. doi: 10.1038/nrc1886.

Mehta, R. S. *et al.* (2017) 'Dietary Patterns and Risk of Colorectal Cancer: Analysis by Tumor Location and Molecular Subtypes', *Gastroenterology*. W.B. Saunders, 152(8), pp. 1944-1953.e1. doi: 10.1053/j.gastro.2017.02.015.

Mejri, N. *et al.* (2017) 'Prognostic value of tumor size in stage II and III colorectal cancer in Tunisian population', *Colorectal Cancer*. Future Medicine Ltd, 6(4), pp. 113–119. doi: 10.2217/crc-2017-0011.

Menezes, M. E. *et al.* (2016) 'Detecting Tumor Metastases: The Road to Therapy Starts Here', in *Advances in Cancer Research*. Academic Press Inc., pp. 1–44. doi: 10.1016/bs.acr.2016.07.001.

Merritt, J. H., Kadouri, D. E. and O'Toole, G. A. (2005) 'Growing and Analyzing Static Biofilms', in *Current Protocols in Microbiology*. John Wiley & Sons, Inc. doi: 10.1002/9780471729259.mc01b01s00.

Miles, A. A., Misra, S. S. and Irwin, J. O. (1938) 'The estimation of the bactericidal power of the blood', *Journal of Hygiene*. Cambridge University Press, 38(6), pp. 732–749. doi: 10.1017/S002217240001158X.

Molenaar, M. *et al.* (1996) 'XTcf-3 transcription factor mediates β-catenin-induced axis formation in xenopus embryos', *Cell*. Cell Press, 86(3), pp. 391–399. doi: 10.1016/S0092-8674(00)80112-9.

Molly, K. *et al.* (1994) 'Validation of the simulator of the human intestinal microbial ecosystem (SHIME) reactor using microorganism-associated activities', *Microbial Ecology in Health and Disease*, 7(4), pp. 191–200. doi: 10.3109/08910609409141354.

Monroe, K. R. *et al.* (2003) 'Correlation of Dietary Intake and Colorectal Cancer Incidence Among Mexican-American Migrants: The Multiethnic Cohort Study'. doi: 10.1207/S15327914NC4502_01.

Mooi, J. K. *et al.* (2018) 'The prognostic impact of consensus molecular subtypes (CMS) and its predictive effects for bevacizumab benefit in metastatic colorectal cancer: Molecular analysis of the AGITG MAX clinical trial', *Annals of Oncology*. Elsevier, 29(11), pp. 2240–2246. doi: 10.1093/annonc/mdy410.

Mori, G. *et al.* (2018) 'Shifts of Faecal Microbiota during Sporadic Colorectal Carcinogenesis', *Scientific Reports*. Nature Publishing Group, 8(1), pp. 1–11. doi:

10.1038/s41598-018-28671-9.

Moss, S. F. *et al.* (1996) 'Inward growth of colonic adenomatous polyps', *Gastroenterology*. W.B. Saunders, 111(6), pp. 1425–1432. doi: 10.1016/S0016-5085(96)70003-3.

Mozdiak, E. *et al.* (2019) 'Colorectal cancer and adenoma screening using urinary volatile organic compound (VOC) detection: early results from a single-centre bowel screening population (UK BCSP)', *Techniques in Coloproctology*. Springer-Verlag Italia s.r.l., 23(4), pp. 343–351. doi: 10.1007/s10151-019-01963-6.

Muhammad, N. *et al.* (2017) 'Involvement of c-Fos in the promotion of cancer stem-like cell properties in head and neck squamous cell carcinoma', *Clinical Cancer Research*. American Association for Cancer Research Inc., 23(12), pp. 3120–3128. doi: 10.1158/1078-0432.CCR-16-2811.

Mullish, B. H. and Williams, H. R. T. (2018) 'Clostridium diffi cile infection and antibiotic-associated diarrhoea', *Clinical Medicine, Journal of the Royal College of Physicians of London*. Royal College of Physicians, 18(3), pp. 237–241. doi: 10.7861/clinmedicine.18-3-237.

Munemitsu, S. *et al.* (1995) 'Regulation of intracellular β-catenin levels by the adenomatous polyposis coli (APC) tumor-suppressor protein', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 92(7), pp. 3046–3050. doi: 10.1073/pnas.92.7.3046.

Murphy, G. *et al.* (1985) 'Purification and characterization of a bone metalloproteinase that degrades gelatin and types IV and V collagen', *Biochimica et Biophysica Acta (BBA)/Protein Structure and Molecular*. Elsevier, 831(1), pp. 49–58. doi: 10.1016/0167-4838(85)90148-7.

Muthakkaruppan, V. R., Kubai, L. and Auerbach, R. (1982) 'Tumor-induced neovascularization in the mouse eye', *Journal of the National Cancer Institute*, 69(3), pp. 699–708. doi: 10.1093/jnci/69.3.699.

Muzny, D. M. *et al.* (2012) 'Comprehensive molecular characterization of human colon and rectal cancer', *Nature*. Nature, 487(7407), pp. 330–337. doi: 10.1038/nature11252.

Myers, L. L. *et al.* (1985) 'Association of enterotoxigenic bacteroides fragilis with diarrheal disease in calves', *Journal of Infectious Diseases*. J Infect Dis, 152(6), pp. 1344–1347. doi: 10.1093/infdis/152.6.1344.

Nakano, V. *et al.* (2008) 'Adherence and invasion of Bacteroidales isolated from the human intestinal tract', *Clinical Microbiology and Infection*. Blackwell Publishing Ltd, 14(10), pp. 955–963. doi: 10.1111/j.1469-0691.2008.02069.x.

Nakatsu, G. *et al.* (2015) 'Gut mucosal microbiome across stages of colorectal carcinogenesis', *Nature Communications*, 6. doi: 10.1038/ncomms9727.

Nakayama, M. and Oshima, M. (2019) 'Mutant p53 in colon cancer', *Journal of Molecular Cell Biology*. Oxford University Press, pp. 267–276. doi: 10.1093/jmcb/mjy075.

Näthke, I. S. *et al.* (1996) 'The adenomatous polyposis coli tumor suppressor protein localizes to plasma membrane sites involved in active cell migration', *Journal of*

Cell Biology. The Rockefeller University Press, 134(1), pp. 165–179. doi: 10.1083/jcb.134.1.165.

Navarro, M. *et al.* (2017) 'Colorectal cancer population screening programs worldwide in 2016: An update', *World Journal of Gastroenterology*. Baishideng Publishing Group Co., Limited, pp. 3632–3642. doi: 10.3748/wjg.v23.i20.3632.

Nejman, D. *et al.* (2020) 'The human tumor microbiome is composed of tumor type-specific intracellular bacteria', *Science*. American Association for the Advancement of Science, 368(6494), pp. 973–980. doi: 10.1126/science.aay9189.

Nelson, C. M. *et al.* (2008) 'Change in cell shape is required for matrix metalloproteinase-induced epithelial-mesenchymal transition of mammary epithelial cells', *Journal of Cellular Biochemistry*. NIH Public Access, 105(1), pp. 25–33. doi: 10.1002/jcb.21821.

Neu, J. and Rushing, J. (2011) 'Cesarean Versus Vaginal Delivery: Long-term Infant Outcomes and the Hygiene Hypothesis', *Clinics in Perinatology*. NIH Public Access, pp. 321–331. doi: 10.1016/j.clp.2011.03.008.

Nougayrède, J. P. *et al.* (2005) 'Cyclomodulins: Bacterial effectors that modulate the eukaryotic cell cycle', *Trends in Microbiology*. Elsevier Current Trends, pp. 103–110. doi: 10.1016/j.tim.2005.01.002.

Nougayrède, J. P. *et al.* (2006) 'Escherichia coli induces DNA double-strand breaks in eukaryotic cells', *Science*. American Association for the Advancement of Science, 313(5788), pp. 848–851. doi: 10.1126/science.1127059.

Nowak, D. *et al.* (2005) 'Beta-actin in human colon adenocarcinoma cell lines with different metastatic potential', *Acta Biochimica Polonica*, 52(2), pp. 461–468. doi: 10.18388/abp.2005_3460.

O'Toole, G. A. (2011) 'Microtiter Dish Biofilm Formation Assay', *Journal of Visualized Experiments*. MyJoVE Corporation, (47). doi: 10.3791/2437.

Ogino, S. *et al.* (2018) 'Integrative analysis of exogenous, endogenous, tumour and immune factors for precision medicine', *Gut*, 67(6), p. gutjnl-2017-315537. doi: 10.1136/gutjnl-2017-315537.

Okuda, T. *et al.* (2012) 'Synergy in biofilm formation between Fusobacterium nucleatum and Prevotella species', *Anaerobe*, 18(1), pp. 110–116. doi: 10.1016/j.anaerobe.2011.09.003.

Olier, M. *et al.* (2012) 'Genotoxicity of escherichia coli nissle 1917 strain cannot be dissociated from its probiotic activity', *Gut Microbes*. Taylor & Francis, 3(6), pp. 501–509. doi: 10.4161/gmic.21737.

Omar, M. I. *et al.* (2016) 'A Study on VOCs Released by Lung Cancer Cell Line Using GCMS-SPME', *Procedia Chemistry*. Elsevier, 20, pp. 1–7. doi: 10.1016/j.proche.2016.07.027.

Osman, M. A. *et al.* (2018) '16S rRNA gene sequencing for deciphering the colorectal cancer gut microbiome: Current protocols and workflows', *Frontiers in Microbiology*. Frontiers Media S.A. doi: 10.3389/fmicb.2018.00767.

Pan, J. *et al.* (2020) 'Prevalence and risk factors for colorectal polyps in a Chinese

population: a retrospective study', *Scientific Reports*. Nature Research, 10(1), pp. 1–8. doi: 10.1038/s41598-020-63827-6.

Paňková, K. *et al.* (2010) 'The molecular mechanisms of transition between mesenchymal and amoeboid invasiveness in tumor cells', *Cellular and Molecular Life Sciences*. Springer, pp. 63–71. doi: 10.1007/s00018-009-0132-1.

Pantel, K. and Brakenhoff, R. H. (2004) 'Dissecting the metastatic cascade', *Nature Reviews Cancer*. Nature Publishing Group, pp. 448–456. doi: 10.1038/nrc1370.

Papandreou, I. *et al.* (2006) 'HIF-1 mediates adaptation to hypoxia by actively downregulating mitochondrial oxygen consumption', *Cell Metabolism*. Cell Metab, 3(3), pp. 187–197. doi: 10.1016/j.cmet.2006.01.012.

Paraskeva, C. *et al.* (1984) 'The isolation and characterization of colorectal epithelial cell lines at different stages in malignant transformation from familial polyposis coli patients', *International Journal of Cancer*, 34(1), pp. 49–56. doi: 10.1002/ijc.2910340109.

Paraskeva, C. *et al.* (1989) 'Specific Cytogenetic Abnormalities in Two New Human Colorectal Adenoma-derived Epithelial Cell Lines', *Cancer Research*, 49(5), pp. 1282–1286. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2917357 (Accessed: 2 May 2019).

Parhi, L. *et al.* (2020) 'Breast cancer colonization by Fusobacterium nucleatum accelerates tumor growth and metastatic progression', *Nature Communications*. Nature Research, 11(1), pp. 1–12. doi: 10.1038/s41467-020-16967-2.

Pelaseyed, T. *et al.* (2014) 'The mucus and mucins of the goblet cells and enterocytes provide the first defense line of the gastrointestinal tract and interact with the immune system', *Immunological Reviews*. NIH Public Access, pp. 8–20. doi: 10.1111/imr.12182.

Perez-Muñoz, M. E. *et al.* (2017) 'A critical assessment of the "sterile womb" and "in utero colonization" hypotheses: Implications for research on the pioneer infant microbiome', *Microbiome*. BioMed Central Ltd., pp. 1–19. doi: 10.1186/s40168-017-0268-4.

Petersen, C. and Round, J. L. (2014) 'Defining dysbiosis and its influence on host immunity and disease', *Cellular Microbiology*. Blackwell Publishing Ltd, pp. 1024–1033. doi: 10.1111/cmi.12308.

Peterson, D. A. *et al.* (2007) 'IgA Response to Symbiotic Bacteria as a Mediator of Gut Homeostasis', *Cell Host and Microbe*. Cell Host Microbe, 2(5), pp. 328–339. doi: 10.1016/j.chom.2007.09.013.

Peterson, J. *et al.* (2009) 'The NIH Human Microbiome Project', *Genome Research*. Cold Spring Harbor Laboratory Press, 19(12), pp. 2317–2323. doi: 10.1101/gr.096651.109.

Pfeiffer, T., Schuster, S. and Bonhoeffer, S. (2001) 'Cooperation and competition in the evolution of ATP-producing pathways', *Science*. American Association for the Advancement of Science, 292(5516), pp. 504–507. doi: 10.1126/science.1058079.

Pino, M. S. and Chung, D. C. (2010) 'The Chromosomal Instability Pathway in Colon Cancer', *Gastroenterology*. NIH Public Access, 138(6), pp. 2059–2072. doi:

10.1053/j.gastro.2009.12.065.

Plummer, M. *et al.* (2015) 'Global burden of gastric cancer attributable to *Helicobacter pylori*', *International Journal of Cancer*, 136(2). doi: 10.1002/ijc.28999.

Pretlow, T. G. *et al.* (1991) 'Aberrant Crypts: Putative Preneoplastic Foci in Human Colonie Mucosa', *Cancer Research*, 51(5), pp. 1564–1567. Available at: https://pubmed.ncbi.nlm.nih.gov/1997197/ (Accessed: 24 April 2021).

Pretzsch, E. *et al.* (2019) 'Mechanisms of Metastasis in Colorectal Cancer and Metastatic Organotropism: Hematogenous versus Peritoneal Spread', *Journal of Oncology*. Hindawi Limited. doi: 10.1155/2019/7407190.

Pryde, S. E. *et al.* (2002) 'The microbiology of butyrate formation in the human colon', *FEMS Microbiology Letters*. Blackwell Publishing Ltd, 217(2), pp. 133–139. doi: 10.1111/j.1574-6968.2002.tb11467.x.

Purcell, R. V. *et al.* (2017) 'Distinct gut microbiome patterns associate with consensus molecular subtypes of colorectal cancer', *Scientific Reports*. Sci Rep, 7(1). doi: 10.1038/s41598-017-11237-6.

Putze, J. *et al.* (2009) 'Genetic structure and distribution of the colibactin genomic island among members of the family Enterobacteriaceae', *Infection and Immunity*. Infect Immun, 77(11), pp. 4696–4703. doi: 10.1128/IAI.00522-09.

Qin, J. *et al.* (2010) 'A human gut microbial gene catalogue established by metagenomic sequencing', *Nature*. Nature Publishing Group, 464(7285), pp. 59–65. doi: 10.1038/nature08821.

Qualtrough, D. *et al.* (2004) 'Hedgehog signalling in colorectal tumour cells: Induction of apoptosis with cyclopamine treatment', *International Journal of Cancer*. John Wiley & Sons, Ltd, 110(6), pp. 831–837. doi: 10.1002/ijc.20227.

Qualtrough, D. *et al.* (2007) 'Prostaglandin F2α stimulates motility and invasion in colorectal tumor cells', *International Journal of Cancer*. John Wiley & Sons, Ltd, 121(4), pp. 734–740. doi: 10.1002/ijc.22755.

Qualtrough, D. *et al.* (2009) 'The actin-bundling protein fascin is overexpressed in colorectal adenomas and promotes motility in adenoma cells in vitro', *British Journal of Cancer*. Nature Publishing Group, 101(7), pp. 1124–1129. doi: 10.1038/sj.bjc.6605286.

Qualtrough, D. *et al.* (2015) 'The hedgehog inhibitor cyclopamine reduces β-catenin-Tcf transcriptional activity, induces E-cadherin expression, and reduces invasion in colorectal cancer cells', *Cancers*. MDPI AG, 7(3), pp. 1885–1899. doi: 10.3390/cancers7030867.

Rahim, M. *et al.* (2019) '13C-Urea Breath Test Accuracy for Helicobacter pylori Infection in the Asian Population: A Meta-Analysis'. doi: 10.5334/aogh.2570.

Rawla, P., Sunkara, T. and Barsouk, A. (2019) 'Epidemiology of colorectal cancer: Incidence, mortality, survival, and risk factors', *Przeglad Gastroenterologiczny*. Termedia Publishing House Ltd., pp. 89–103. doi: 10.5114/pg.2018.81072.

Reya, T. and Clevers, H. (2005) 'Wnt signalling in stem cells and cancer', *Nature*,

pp. 843–850. doi: 10.1038/nature03319.

Rhee, K. J. *et al.* (2009) 'Induction of persistent colitis by a human commensal, enterotoxigenic Bacteroides fragilis, in wild-type C57BL/6 mice', *Infection and Immunity*. Infect Immun, 77(4), pp. 1708–1718. doi: 10.1128/IAI.00814-08.

Riihimaki, M. *et al.* (2016) 'Patterns of metastasis in colon and rectal cancer', *Scientific Reports*. Nature Publishing Group, 6. doi: 10.1038/srep29765.

Rodríguez-Piñeiro, A. M. *et al.* (2013) 'Studies of mucus in mouse stomach, small intestine, and colon. II. Gastrointestinal mucus proteome reveals Muc2 and Muc5ac accompanied by a set of core proteins', *American Journal of Physiology - Gastrointestinal and Liver Physiology*. American Physiological Society, 305(5), p. G348. doi: 10.1152/ajpgi.00047.2013.

Roediger, W. E. W. (1980) 'Role of anaerobic bacteria in the metabolic welfare of the colonic mucosa in man', *Gut*. BMJ Publishing Group, 21(9), pp. 793–798. doi: 10.1136/gut.21.9.793.

Roediger, W. E. W. (1982) 'Utilization of Nutrients by Isolated Epithelial Cells of the Rat Colon', *Gastroenterology*. Elsevier, 83(2), pp. 424–429. doi: 10.1016/S0016-5085(82)80339-9.

Rondanelli, M. *et al.* (2019) 'Volatile organic compounds as biomarkers of gastrointestinal diseases and nutritional status', *Journal of Analytical Methods in Chemistry*. Hindawi Limited. doi: 10.1155/2019/7247802.

Round, A. N. *et al.* (2012) 'Lamellar structures of MUC2-rich mucin: A potential role in governing the barrier and lubricating functions of intestinal mucus', *Biomacromolecules*. American Chemical Society, 13(10), pp. 3253–3261. doi: 10.1021/bm301024x.

Round, J. L. *et al.* (2011) 'The toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota', *Science*. Science, 332(6032), pp. 974–977. doi: 10.1126/science.1206095.

Round, J. L. and Mazmanian, S. K. (2010) 'Inducible Foxp3+ regulatory T-cell development by a commensal bacterium of the intestinal microbiota', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 107(27), pp. 12204–12209. doi: 10.1073/pnas.0909122107.

Rubinstein, M. R. *et al.* (2013) 'Fusobacterium nucleatum Promotes Colorectal Carcinogenesis by Modulating E-Cadherin/??-Catenin Signaling via its FadA Adhesin', *Cell Host and Microbe*. NIH Public Access, 14(2), pp. 195–206. doi: 10.1016/j.chom.2013.07.012.

Rubinstein, M. R. *et al.* (2019) 'Fusobacterium nucleatum promotes colorectal cancer by inducing Wnt/β-catenin modulator Annexin A1', *EMBO reports*. EMBO, 20(4). doi: 10.15252/embr.201847638.

Sadabad, M. S. *et al.* (2015) 'A simple coculture system shows mutualism between anaerobic faecalibacteria and epithelial Caco-2 cells', *Scientific Reports*. Nature Publishing Group, 5(1), p. 17906. doi: 10.1038/srep17906.

Saffarian, A. *et al.* (2019) 'Crypt- and mucosa-associated core microbiotas in humans and their alteration in colon cancer patients', *mBio*, 10(4). doi:

10.1128/mBio.01315-19.

Saidi, R. F. and Sears, C. L. (1996) 'Bacteroides fragilis toxin rapidly intoxicates human intestinal epithelial cells (HT29/C1) in vitro', *Infection and Immunity*. American Society for Microbiology (ASM), 64(12), pp. 5029–5034. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8945542 (Accessed: 11 December 2018).

Sasaki, N. *et al.* (2016) 'Reg4+ deep crypt secretory cells function as epithelial niche for Lgr5+ stem cells in colon', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 113(37), pp. E5399–E5407. doi: 10.1073/pnas.1607327113.

Satoh, K. *et al.* (2017) 'Global metabolic reprogramming of colorectal cancer occurs at adenoma stage and is induced by MYC', *Proceedings of the National Academy of Sciences*. National Academy of Sciences, 114(37), pp. E7697–E7706. doi: 10.1073/pnas.1710366114.

Schmalhofer, O., Brabletz, S. and Brabletz, T. (2009) 'E-cadherin, β-catenin, and ZEB1 in malignant progression of cancer', *Cancer and Metastasis Reviews*, pp. 151–166. doi: 10.1007/s10555-008-9179-y.

Schroeder, B. O. (2019) 'Fight them or feed them: How the intestinal mucus layer manages the gut microbiota', *Gastroenterology Report*. Oxford University Press, pp. 3–12. doi: 10.1093/gastro/goy052.

Schulz, M. D. *et al.* (2014) 'High-fat-diet-mediated dysbiosis promotes intestinal carcinogenesis independently of obesity', *Nature*. Nature Publishing Group, 514(7253), pp. 508–512. doi: 10.1038/nature13398.

Sears, C. L. (2009) 'Enterotoxigenic Bacteroides fragilis: A rogue among symbiotes', *Clinical Microbiology Reviews*, pp. 349–369. doi: 10.1128/CMR.00053-08.

Sears, C. L. and Pardoll, D. M. (2011) 'Perspective: Alpha-bugs, their microbial partners, and the link to colon cancer', *Journal of Infectious Diseases*. Oxford University Press, pp. 306–311. doi: 10.1093/jinfdis/jiq061.

Secher, T. *et al.* (2013) 'Escherichia coli Producing Colibactin Triggers Premature and Transmissible Senescence in Mammalian Cells', *PLoS ONE*. Public Library of Science, 8(10), p. 77157. doi: 10.1371/journal.pone.0077157.

Segain, J. P. *et al.* (2000) 'Butyrate inhibits inflammatory responses through NFκB inhibition: Implications for Crohn's disease', *Gut*. BMJ Publishing Group, 47(3), pp. 397–403. doi: 10.1136/gut.47.3.397.

Sekharam, M. *et al.* (2003) 'Insulin-like Growth Factor 1 Receptor Enhances Invasion and Induces Resistance to Apoptosis of Colon Cancer Cells through the Akt/Bcl-xL Pathway', *Cancer Research*, 63(22), pp. 7708–7716. Available at: https://pubmed.ncbi.nlm.nih.gov/14633695/ (Accessed: 25 March 2021).

Sender, R., Fuchs, S. and Milo, R. (2016) 'Revised Estimates for the Number of Human and Bacteria Cells in the Body.', *PLoS biology*. Public Library of Science, 14(8), p. e1002533. doi: 10.1371/journal.pbio.1002533.

Sewda, K. *et al.* (2016) 'Cell-surface markers for colon adenoma and adenocarcinoma', *Oncotarget*. Impact Journals LLC, 7(14), pp. 17773–17789. doi:

10.18632/oncotarget.7402.

Sghir, A. *et al.* (2000) 'Quantification of bacterial groups within human fecal flora by oligonucleotide probe hybridization', *Applied and Environmental Microbiology*. Appl Environ Microbiol, 66(5), pp. 2263–2266. doi: 10.1128/AEM.66.5.2263-2266.2000.

Shah, P. *et al.* (2016) 'A microfluidics-based in vitro model of the gastrointestinal human-microbe interface', *Nature Communications*. Nature Publishing Group, 7(1), pp. 1–15. doi: 10.1038/ncomms11535.

Shanahan, F. (2013) 'The colonic microbiota in health and disease', *Current Opinion in Gastroenterology*, pp. 49–54. doi: 10.1097/MOG.0b013e32835a3493.

Sheng, Q. *et al.* (2019) 'Characteristics of fecal gut microbiota in patients with colorectal cancer at different stages and different sites', *Oncology Letters*. Spandidos Publications, 18(5), pp. 4834–4844. doi: 10.3892/ol.2019.10841.

Shestov, A. A. *et al.* (2014) 'Quantitative determinants of aerobic glycolysis identify flux through the enzyme GAPDH as a limiting step', *eLife*. eLife Sciences Publications Ltd, 3(July2014), pp. 1–18. doi: 10.7554/eLife.03342.

Shiratori, R. *et al.* (2019) 'Glycolytic suppression dramatically changes the intracellular metabolic profile of multiple cancer cell lines in a mitochondrial metabolism-dependent manner', *Scientific Reports*. Nature Research, 9(1), pp. 1–15. doi: 10.1038/s41598-019-55296-3.

Shuldiner, J., Liu, Y. and Lofters, A. (2018) 'Incidence of breast and colorectal cancer among immigrants in Ontario, Canada: A retrospective cohort study from 2004-2014', *BMC Cancer*. BioMed Central Ltd., 18(1), p. 537. doi: 10.1186/s12885-018-4444-0.

Sieber, O. M., Heinimann, K. and Tomlinson, I. P. M. (2003) 'Genomic instability - The engine of tumorigenesis?', *Nature Reviews Cancer*. European Association for Cardio-Thoracic Surgery, pp. 701–708. doi: 10.1038/nrc1170.

Silacci, P. *et al.* (2004) 'Gelsolin superfamily proteins: Key regulators of cellular functions', *Cellular and Molecular Life Sciences*. Springer, pp. 2614–2623. doi: 10.1007/s00018-004-4225-6.

Simiantonaki, N. *et al.* (2008) 'Hypoxia-inducible factor 1 alpha expression increases during colorectal carcinogenesis and tumor progression', *BMC Cancer*. BioMed Central, 8, p. 320. doi: 10.1186/1471-2407-8-320.

Simiczyjew, A. *et al.* (2017) 'Involvement of β-and γ-actin isoforms in actin cytoskeleton organization and migration abilities of bleb-forming human colon cancer cells', *PLoS ONE*. Public Library of Science, 12(3). doi: 10.1371/journal.pone.0173709.

Slade, E. A. *et al.* (2017) 'In vitro discrimination of wound-associated bacteria by volatile compound profiling using selected ion flow tube-mass spectrometry', *Journal of Applied Microbiology*. Blackwell Publishing Ltd, 123(1), pp. 233–245. doi: 10.1111/jam.13473.

Soler, A. P. *et al.* (1999) 'Increased tight junctional permeability is associated with the development of colon cancer', *Carcinogenesis*. Carcinogenesis, 20(8), pp. 1425–

1431. doi: 10.1093/carcin/20.8.1425.

Song, Z. *et al.* (2019) 'Taxonomic profiling and populational patterns of bacterial bile salt hydrolase (BSH) genes based on worldwide human gut microbiome', *Microbiome*. BioMed Central Ltd., 7(1), p. 9. doi: 10.1186/s40168-019-0628-3.

Sonnenburg, E. D. *et al.* (2016) 'Diet-induced extinctions in the gut microbiota compound over generations', *Nature*. Nature Publishing Group, 529(7585), pp. 212–215. doi: 10.1038/nature16504.

Sorenson, C. M., Barry, M. A. and Eastman, A. (1990) 'Analysis of events associated with cell cycle arrest at g2 phase and cell death induced by cisplatin', *Journal of the National Cancer Institute*, 82(9), pp. 749–755. doi: 10.1093/jnci/82.9.749.

Spencer, N. J. *et al.* (2016) 'Insights into the mechanisms underlying colonic motor patterns', *The Journal of Physiology*. Blackwell Publishing Ltd, 594(15), pp. 4099–4116. doi: 10.1113/JP271919.

Staley, C. *et al.* (2017) 'Interaction of gut microbiota with bile acid metabolism and its influence on disease states', *Applied Microbiology and Biotechnology*, 101(1), pp. 47–64. doi: 10.1007/s00253-016-8006-6.

Stamp, C. *et al.* (2018) 'Predominant Asymmetrical Stem Cell Fate Outcome Limits the Rate of Niche Succession in Human Colonic Crypts', *EBioMedicine*. Elsevier B.V., 31, pp. 166–173. doi: 10.1016/j.ebiom.2018.04.017.

Stecher, B. *et al.* (2007) 'Salmonella enterica serovar typhimurium exploits inflammation to compete with the intestinal microbiota', *PLoS Biology*. Public Library of Science, 5(10), pp. 2177–2189. doi: 10.1371/journal.pbio.0050244.

Steck, N. *et al.* (2011) 'Enterococcus faecalis metalloprotease compromises epithelial barrier and contributes to intestinal inflammation', *Gastroenterology*. W.B. Saunders, 141(3), pp. 959–971. doi: 10.1053/j.gastro.2011.05.035.

Stephens, M. and von der Weid, P.-Y. (2020) 'Lipopolysaccharides modulate intestinal epithelial permeability and inflammation in a species-specific manner', *Gut Microbes*. Taylor and Francis Inc., 11(3), pp. 421–432. doi: 10.1080/19490976.2019.1629235.

Sternlicht, M. D. *et al.* (1999) 'The stromal proteinase MMP3/stromelysin-1 promotes mammary carcinogenesis', *Cell*. Cell Press, 98(2), pp. 137–146. doi: 10.1016/S0092-8674(00)81009-0.

Stewart, C. J. *et al.* (2018) 'Temporal development of the gut microbiome in early childhood from the TEDDY study', *Nature*. Nature Publishing Group, pp. 583–588. doi: 10.1038/s41586-018-0617-x.

Sträter, J. *et al.* (1995) 'In situ detection of enterocytic apoptosis in normal colonic mucosa and in familial adenomatous polyposis', *Gut*. BMJ Publishing Group, 37(6), pp. 819–825. doi: 10.1136/gut.37.6.819.

Strauss, J. *et al.* (2011) 'Invasive potential of gut mucosa-derived fusobacterium nucleatum positively correlates with IBD status of the host', *Inflammatory Bowel Diseases*. Inflamm Bowel Dis, 17(9), pp. 1971–1978. doi: 10.1002/ibd.21606.

Sung, H. *et al.* (2021) 'Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries', *CA: A Cancer Journal for Clinicians*. Wiley, p. caac.21660. doi: 10.3322/caac.21660.

Swidsinski, A. *et al.* (1998) 'Association between intraepithelial Escherichia coli and colorectal cancer', *Gastroenterology*. W.B. Saunders, 115(2), pp. 281–286. doi: 10.1016/S0016-5085(98)70194-5.

Swidsinski, A. *et al.* (2005) 'Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease', *Journal of Clinical Microbiology*. American Society for Microbiology (ASM), 43(7), pp. 3380–3389. doi: 10.1128/JCM.43.7.3380-3389.2005.

Tan, T. Y. *et al.* (2017) 'Clinical characteristics and antimicrobial susceptibilities of anaerobic bacteremia in an acute care hospital', *Anaerobe*. Academic Press, 43, pp. 69–74. doi: 10.1016/j.anaerobe.2016.11.009.

Tanaka, S. *et al.* (2016) 'Snail1 expression in human colon cancer DLD-1 cells confers invasive properties without N-cadherin expression', *Biochemistry and Biophysics Reports*. Elsevier B.V., 8, pp. 120–126. doi: 10.1016/j.bbrep.2016.08.017.

Tanoue, T., Atarashi, K. and Honda, K. (2016) 'Development and maintenance of intestinal regulatory T cells', *Nature Reviews Immunology*. Nat Rev Immunol, pp. 295–309. doi: 10.1038/nri.2016.36.

Tay, D. L. M., Bhathal, P. S. and Fox, R. M. (1991) 'Quantitation of G0 and G1 phase cells in primary carcinomas: Antibody to M1 subunit of ribonucleotide reductase shows G1 phase restriction point block', *Journal of Clinical Investigation*. The American Society for Clinical Investigation, 87(2), pp. 519–527. doi: 10.1172/JCI115026.

Terry, P. *et al.* (2001) 'Prospective study of major dietary patterns and colorectal cancer risk in women', *American Journal of Epidemiology*. Am J Epidemiol, 154(12), pp. 1143–1149. doi: 10.1093/aje/154.12.1143.

The Human Microbiome Consortium (2012) 'Structure, function and diversity of the healthy human microbiome', *Nature*, pp. 207–214. doi: 10.1038/nature11234.

Tian, S. *et al.* (2018) 'Secreted AGR2 promotes invasion of colorectal cancer cells via Wnt11-mediated non-canonical Wnt signaling', *Experimental Cell Research*. Elsevier Inc., 364(2), pp. 198–207. doi: 10.1016/j.yexcr.2018.02.004.

Tjalsma, H. *et al.* (2012) 'A bacterial driver-passenger model for colorectal cancer: Beyond the usual suspects', *Nature Reviews Microbiology*, pp. 575–582. doi: 10.1038/nrmicro2819.

Tomasello, G. *et al.* (2014) 'Dismicrobism in inflammatory bowel disease and colorectal cancer: Changes in response of colocytes', *World Journal of Gastroenterology*, 20(48), pp. 18121–18130. doi: 10.3748/wjg.v20.i48.18121.

Tomkovich, S. *et al.* (2019) 'Human colon mucosal biofilms from healthy or colon cancer hosts are carcinogenic', *Journal of Clinical Investigation*. American Society for Clinical Investigation, 129(4), pp. 1699–1712. doi: 10.1172/JCI124196.

Tong, X. *et al.* (2014) 'SOX10, a novel HMG-box-containing tumor suppressor,

inhibits growth and metastasis of digestive cancers by suppressing the wnt/β-catenin pathway', *Oncotarget*. Impact Journals LLC, 5(21), pp. 10571–10583. doi: 10.18632/oncotarget.2512.

Traxler, S. *et al.* (2019) 'Volatile scents of influenza A and S. pyogenes (co-)infected cells', *Scientific Reports*. Nature Research, 9(1), pp. 1–12. doi: 10.1038/s41598-019-55334-0.

Tudek, B., Bird, R. P. and Bruce, W. R. (1989) 'Foci of Aberrant Crypts in the Colons of Mice and Rats Exposed to Carcinogens Associated with Foods', *Cancer Research*, 49(5), pp. 1236–1240. Available at: https://pubmed.ncbi.nlm.nih.gov/2917353/ (Accessed: 24 April 2021).

Turnbaugh, P. J. *et al.* (2009) 'A core gut microbiome in obese and lean twins', *Nature*. Nature Publishing Group, 457(7228), pp. 480–484. doi: 10.1038/nature07540.

Turnbaugh, P. J. and Gordon, J. I. (2009) 'The core gut microbiome, energy balance and obesity', in *Journal of Physiology*. Wiley/Blackwell (10.1111), pp. 4153–4158. doi: 10.1113/jphysiol.2009.174136.

Turroni, F. *et al.* (2019) 'Bifidobacterium bifidum: A key member of the early human gut microbiota', *Microorganisms*. MDPI AG. doi: 10.3390/microorganisms7110544.

UK, C. R. (2015) *Bowel cancer statistics | Cancer Research UK*. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer (Accessed: 31 August 2018).

Ulger Toprak, N. *et al.* (2006) 'A possible role of Bacteroides fragilis enterotoxin in the aetiology of colorectal cancer', *Clinical Microbiology and Infection*. Elsevier, 12(8), pp. 782–786. doi: 10.1111/j.1469-0691.2006.01494.x.

Vaga, S. *et al.* (2020) 'Compositional and functional differences of the mucosal microbiota along the intestine of healthy individuals', *Scientific Reports*. Nature Research, 10(1), p. 14977. doi: 10.1038/s41598-020-71939-2.

Visconti, A. *et al.* (2019) 'Interplay between the human gut microbiome and host metabolism', *Nature Communications*. Nature Publishing Group, 10(1), pp. 1–10. doi: 10.1038/s41467-019-12476-z.

Van Der Vlugt, M. *et al.* (2017) 'Adherence to colorectal cancer screening: Four rounds of faecal immunochemical test-based screening', *British Journal of Cancer*. Nature Publishing Group, 116(1), pp. 44–49. doi: 10.1038/bjc.2016.399.

Vogelstein, B. and Kinzler, K. W. (1993) 'The multistep nature of cancer', *Trends in Genetics*. Elsevier Current Trends, pp. 138–141. doi: 10.1016/0168-9525(93)90209-Z.

Wang, H. B. *et al.* (2012) 'Butyrate enhances intestinal epithelial barrier function via up-regulation of tight junction protein claudin-1 transcription', *Digestive Diseases and Sciences*. Dig Dis Sci, 57(12), pp. 3126–3135. doi: 10.1007/s10620-012-2259-4.

Wang, S. *et al.* (2014) 'Enterococcus faecalis from healthy infants modulates inflammation through MAPK signaling pathways', *PLoS ONE*. Public Library of Science, 9(5), p. 97523. doi: 10.1371/journal.pone.0097523.

Wang, X. *et al.* (2008) 'Enterococcus faecalis induces aneuploidy and tetraploidy in colonic epithelial cells through a bystander effect.', *Cancer research*. NIH Public Access, 68(23), pp. 9909–17. doi: 10.1158/0008-5472.CAN-08-1551.

Wang, X. *et al.* (2013) 'Comparative Microbial Analysis of Paired Amniotic Fluid and Cord Blood from Pregnancies Complicated by Preterm Birth and Early-Onset Neonatal Sepsis', *PLoS ONE*. Edited by A. J. Ratner. Public Library of Science, 8(2), p. e56131. doi: 10.1371/journal.pone.0056131.

Wang, X. and Huycke, M. M. (2007) 'Extracellular Superoxide Production by Enterococcus faecalis Promotes Chromosomal Instability in Mammalian Cells', *Gastroenterology*. W.B. Saunders, 132(2), pp. 551–561. doi: 10.1053/j.gastro.2006.11.040.

Warburg, O., Wind, F. and Negelein, E. (1927) 'The metabolism of tumors in the body', *Journal of General Physiology*. The Rockefeller University Press, 8(6), pp. 519–530. doi: 10.1085/jgp.8.6.519.

Wassenaar, T. M. (2018) 'E. coli and colorectal cancer: a complex relationship that deserves a critical mindset', *Critical Reviews in Microbiology*. Taylor & Francis, pp. 619–632. doi: 10.1080/1040841X.2018.1481013.

Wei, W. *et al.* (2016) 'MiR-30a-5p Suppresses Tumor Metastasis of Human Colorectal Cancer by Targeting ITGB3', *Cellular Physiology and Biochemistry*. S. Karger AG, 39(3), pp. 1165–1176. doi: 10.1159/000447823.

Weidle, U. H. *et al.* (2011) 'Intracellular proteins displayed on the surface of tumor cells as targets for therapeutic intervention with antibody-related agents', *Cancer Genomics and Proteomics*, pp. 49–64. Available at: https://cgp.iiarjournals.org/content/8/2/49 (Accessed: 1 May 2021).

Weinberg, B. A. *et al.* (2020) 'A comparison study of the intratumoral microbiome in younger verses older-onset colorectal cancer (COSMO CRC).', *Journal of Clinical Oncology*. American Society of Clinical Oncology (ASCO), 38(4_suppl), pp. 241–241. doi: 10.1200/jco.2020.38.4_suppl.241.

Wexler, H. M. (2007) 'Bacteroides: The good, the bad, and the nitty-gritty', *Clinical Microbiology Reviews*. American Society for Microbiology (ASM), pp. 593–621. doi: 10.1128/CMR.00008-07.

WHO (2015) *Q & A on the carcinogenicity of the consumption of red meat and processed meat*, *Int. Agency Res. Cancer*. doi: http://www.who.int/features/qa/cancer-red-meat/en/.

Williams, A. C. *et al.* (2000) 'Increased p53-dependent apoptosis by the insulin-like growth factor binding protein IGFBP-3 in human colonic adenoma-derived cells', *Cancer Research*, 60(1), pp. 22–27. Available at: https://pubmed.ncbi.nlm.nih.gov/10646845/ (Accessed: 6 March 2021).

Williams, A. M. *et al.* (2006) 'Effects of microflora on the neonatal development of gut mucosal T cells and myeloid cells in the mouse', *Immunology*, 119(4), pp. 470–478. doi: 10.1111/j.1365-2567.2006.02458.x.

Wolf, K. *et al.* (2007) 'Multi-step pericellular proteolysis controls the transition from individual to collective cancer cell invasion', *Nature Cell Biology*. Nature Publishing

Group, 9(8), pp. 893–904. doi: 10.1038/ncb1616.

Wolff, R. K. *et al.* (2018) 'Mutation analysis of adenomas and carcinomas of the colon: Early and late drivers', *Genes, Chromosomes and Cancer*. Blackwell Publishing Inc., 57(7), pp. 366–376. doi: 10.1002/gcc.22539.

Wong, J. M. W. *et al.* (2006) 'Colonic health: Fermentation and short chain fatty acids', in *Journal of Clinical Gastroenterology*, pp. 235–243. doi: 10.1097/00004836-200603000-00015.

Wong, S. H., Zhao, L., *et al.* (2017) 'Gavage of Fecal Samples From Patients With Colorectal Cancer Promotes Intestinal Carcinogenesis in Germ-Free and Conventional Mice', *Gastroenterology*. W.B. Saunders, 153(6), pp. 1621-1633.e6. doi: 10.1053/j.gastro.2017.08.022.

Wong, S. H., Kwong, T. N. Y., *et al.* (2017) 'Quantitation of faecal Fusobacterium improves faecal immunochemical test in detecting advanced colorectal neoplasia', *Gut*. BMJ Publishing Group, 66(8), pp. 1441–1448. doi: 10.1136/gutjnl-2016-312766.

Woodfield, G. *et al.* (2018) 'PTU-072 Breath testing for colorectal polyps and cancer- a load of hot air?', in *Gut*. BMJ, p. A206.3-A208. doi: 10.1136/gutjnl-2018-bsgabstracts.413.

Wu, J., Li, Q. and Fu, X. (2019) 'Fusobacterium nucleatum Contributes to the Carcinogenesis of Colorectal Cancer by Inducing Inflammation and Suppressing Host Immunity', *Translational Oncology*, 12(6). doi: 10.1016/j.tranon.2019.03.003.

Wu, K. *et al.* (2004) 'Dietary patterns and risk of colon cancer and adenoma in a cohort of men (United States)', *Cancer Causes and Control*. Cancer Causes Control, 15(9), pp. 853–862. doi: 10.1007/s10552-004-1809-2.

Wu, S. *et al.* (2003) 'Bacteroides fragilis enterotoxin induces c-Myc expression and cellular proliferation', *Gastroenterology*. W.B. Saunders, 124(2), pp. 392–400. doi: 10.1053/gast.2003.50047.

Wu, S. *et al.* (2009) 'A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses', *Nature Medicine*, 15(9). doi: 10.1038/nm.2015.

Xue, T.-C. *et al.* (2014) 'Goosecoid Promotes the Metastasis of Hepatocellular Carcinoma by Modulating the Epithelial-Mesenchymal Transition', *PLoS ONE*. Edited by J. W. Lee. Public Library of Science, 9(10), p. e109695. doi: 10.1371/journal.pone.0109695.

Yadavalli, S. *et al.* (2017) 'Data-Driven Discovery of Extravasation Pathway in Circulating Tumor Cells', *Scientific Reports*. Nature Publishing Group, 7(1), pp. 1–13. doi: 10.1038/srep43710.

Yao, C. K., Muir, J. G. and Gibson, P. R. (2016) 'Review article: Insights into colonic protein fermentation, its modulation and potential health implications', *Alimentary Pharmacology and Therapeutics*. Blackwell Publishing Ltd, 43(2), pp. 181–196. doi: 10.1111/apt.13456.

Yu, Y.-N. and Fang, J.-Y. (2015) 'Gut Microbiota and Colorectal Cancer', *Gastrointestinal Tumors*, 2(1), pp. 26–32. doi: 10.1159/000380892.

Yu, Y. *et al.* (2019) 'The long-term survival characteristics of a cohort of colorectal cancer patients and baseline variables associated with survival outcomes with or without time-varying effects', *BMC Medicine*. BioMed Central Ltd., 17(1), p. 150. doi: 10.1186/s12916-019-1379-5.

Zagari, F. *et al.* (2013) 'Lactate metabolism shift in CHO cell culture: The role of mitochondrial oxidative activity', *New Biotechnology*. N Biotechnol, 30(2), pp. 238–245. doi: 10.1016/j.nbt.2012.05.021.

Zeevi, D. *et al.* (2019) 'Structural variation in the gut microbiome associates with host health', *Nature*. Nature Publishing Group, 568(7750), pp. 43–48. doi: 10.1038/s41586-019-1065-y.

Zhang, G. and Weintraub, A. (1999) 'Prevalence of Enterotoxigenic Bacteroides fragilis in Adult Patients with Diarrhea and Healthy Controls', *Clinical Infectious Diseases*. University of Chicago Press, 29(3), pp. 590–594. doi: 10.1086/598639.

Zhang, J. *et al.* (2019) 'Oral antibiotic use and risk of colorectal cancer in the United Kingdom, 1989-2012: A matched case-control study', *Gut*. BMJ Publishing Group, 68(11), pp. 1971–1978. doi: 10.1136/gutjnl-2019-318593.

Zhang, S. *et al.* (2018) 'Induction of Escherichia coli Into a VBNC State by Continuous-Flow UVC and Subsequent Changes in Metabolic Activity at the Single-Cell Level', *Frontiers in Microbiology*. Frontiers Media S.A., 9(SEP), p. 2243. doi: 10.3389/fmicb.2018.02243.

Zhang, T. *et al.* (2017) 'Epigenetically upregulated WIPF1 plays a major role in BRAF V600E-promoted papillary thyroid cancer aggressiveness', *Oncotarget*. Impact Journals LLC, 8(1), pp. 900–914. doi: 10.18632/oncotarget.13400.

Zhao, C. *et al.* (2020) 'Supervillin promotes tumor angiogenesis in liver cancer', *Oncology Reports*. Spandidos Publications, 44(2), pp. 674–684. doi: 10.3892/or.2020.7621.

Zhao, D. *et al.* (2020) 'TFPI2 suppresses breast cancer progression through inhibiting TWIST-integrin α5 pathway', *Molecular Medicine*. BioMed Central Ltd., 26(1). doi: 10.1186/s10020-020-00158-2.

Zhao, J. *et al.* (2020) 'The role of interleukin-17 in tumor development and progression', *Journal of Experimental Medicine*. The Rockefeller University Press, 217(1). doi: 10.1084/jem.20190297.

Zhu, X. *et al.* (2019) 'Bone Morphogenetic Protein 1 Targeting COL1A1 and COL1A2 to Regulate the Epithelial-Mesenchymal Transition Process of Colon Cancer SW620 Cells', *Journal of Nanoscience and Nanotechnology*. American Scientific Publishers, 20(3), pp. 1366–1374. doi: 10.1166/jnn.2020.17362.

Zimmermann, D. *et al.* (2007) 'Determination of volatile products of human colon cell line metabolism by GC/MS analysis', *Metabolomics*. Kluwer Academic Publishers-Plenum Publishers, 3(1), pp. 13–17. doi: 10.1007/s11306-006-0038-y.

Zoetendal, E. G. *et al.* (2002) 'Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces', *Applied and Environmental Microbiology*. American Society for Microbiology, 68(7), pp. 3401–3407. doi:

10.1128/AEM.68.7.3401-3407.2002.