Ollscoil na hÉireann, Corcaigh

# National University of Ireland, Cork


## Benign, potentially malignant & malignant oral lesions; an analysis of oral and gut microbiota

Thesis presented by

**Dr Junaid Nayyar** BA, BDentSc, MFDS, PG Cert

for the degree of

**Doctorate of Clinical Dentistry**

**Cork University Dental School and Hospital**

**APC Microbiome Institute**

**Teagasc, Moorepark, Fermoy, Cork**

Head of Cork University Dental School and Hospital: Professor Paul Brady

Head of APC Microbiome Institute: Professor Paul Ross

Head of Teagasc: Professor Catherine Stanton

Supervisors: Professors Paul Brady, Paul Ross & Catherine Stanton

**April 2024**

# Contents

## Declaration

This is to certify that the work I am submitting is my own and has not been submitted for another degree, either at University College Cork or elsewhere.

All external references and sources are clearly acknowledged and identified within the contents.

This work has been carried out in the APC Microbiome Institute, University College Cork, Cork University Dental School and Hospital and at Teagasc, Moorepark, Fermoy, Cork between January 2022 and April 2024. The sample analyses were carried out by Dr Cassandre Bedu and Dr Dhrati Patangia, postdoctoral fellows in Prof. Catherine Stanton's Lab at Teagasc, both of whom contributed to the results and interpretation for discussion.

I have read and understood the regulations of University College Cork concerning plagiarism and intellectual property.

Signed: 

Junaid Nayyar, 2024

# Acknowledgements

I could not have completed this thesis and my clinical postgraduate training on my own.

Juwayriyah, my amazing little sister has always been there and looked out for me throughout my life. Helen my fiancé has supported me throughout the last few years and especially through the postgraduate training and exam pressures. My parents have dedicated so much of their life efforts to making sure that I and my siblings would have the best education.

I also want to thank Professor Paul Brady for his confidence, support and supervision. My colleagues Harriet and Eddie, for their flexibility, comradery and making life in the dental hospital easy. Finally I wish to thank both Dr Cassandre Bedu and Dr Dhrati Patangia as well as the teams in Teagasc and APC for their expertise and kind support in explaining and guiding me through this complex and novel area of research.

# Abstract

Introduction:

The human microbiome is widely known to be associated with health and disease. The oral microbiome has been linked with oral diseases and infections, though not many studies have explored the relation between oral and gut microbiome with oral cancer based on their histology. This study explores the oral and gut microbiota in 30 participants (n=30) divided in to three groups based on histology; benign (B) (n=15), potentially malignant (PM) (n=8), and malignant (M) (n=7) oral lesions.

Methods:

Using shotgun metagenomic sequencing, we analysed the microbiota profiles to determine their potential as biomarkers for oral malignancy. We looked at alpha diversity, beta diversity, taxonomy, differential abundance, and functional profiling.

Results:

Our results showed distinct gut microbial profiles between benign and malignant groups and the association of specific microbes in oral saliva, such as *Haemophilus parainfluenzae* and *Veillonella parvula* with malignancy. The influence of factors such as smoking, alcohol and oral hygiene was also studied, with oral hygiene being the leading factor explaining the variance in oral and gut microbial composition.

Conclusions:

Our study suggests that oral and gut microbiomes could act as possible biomarkers and aid in early detection and assessment of oral cancer risk. Further research is required to develop definitive biomarkers in both potentially malignant and malignant oral lesions.

## List of Tables



## List of Abbreviations

16S: 16S rRNA gene

ASV: Amplicon sequence variant

B: Benign

BA: Butyric acid

BMI: Body mass index

BPE: Basic periodontal exam

CRC: Colo-rectal cancer

DNA: Deoxyribonucleic acid

EGF: Epidermal growth factor

EMT: Epithelial-mesenchymal transition

FFQ: Food frequency questionnaire

HGP: Human genome project

HMP: Human microbiome project

HPA: Hypothalamic—Pituitary—Adrenal

IBD: Inflammatory bowel disease

LPS: Lipopolysaccharide

M: Malignant

MT: Malignant transformation

MTR: Malignant transformation rate

NGS: Next generation sequencing

OPMD: Oral potentially malignant disorder

OSCC: Oral squamous cell carcinoma

PCoA: Principal coordinates analysis

PCR: Polymerase chain reaction

PERMANOVA: Permutational multivariate analysis of variance

PM: Potentially malignant

qPCR: Quantitative real-time polymerase chain reaction

SCC: Squamous cell carcinoma

SCFA: Short chain fatty acids

Spp.: Latin for species pluralis, "multiple species"

TMA: Trimethylamine

TMAO: Trimethylamine-N-oxide

TNF: Tumour necrosis factor

WHO: World health organisation

# Chapter 1: Literature Review

## 1.1 Introduction

The sum of all microorganisms that reside in the human body is known as the microbiome, while the collection of micro-organisms that reside within a certain area are known as its' microbiota (Lederberg and McCray, 2001). This term, first coined by Lederberg, introduced the idea that an ecological community of commensal, symbiotic, and pathogenic microorganisms coexist and play a significant role in the human body in both health and disease.

To be able to detect and count the micro-organisms in a particular location several methods can be used. Traditionally this has been culture and growth on mediums, and more recently genomic sequencing to amplify the small quantities of DNA sampled. With modern techniques we can identify more microbiota and their abundance across different sites in the body.

The Human Microbiome Project (HMP) was influenced and inspired by Human Genome Project (HGP)(NIH HMP Working Group et al., 2009). While the HGP was purported to have been completed in 2003, it was not considered "complete" until 2022 (Nurk et al., 2022). The HGP attempted to catalogue the microbiome in key sites (oral cavity, gut, vagina and skin) from "normal" volunteers. It would then be possible to monitor the link between changes in the microbiome and states of health and disease, establish standardised reference data sources and further scientific study in this novel area. Just as the HGP project benefitted from the change from Next Generation Sequencing (NGS), studies of the microbiome have progressed significantly with the use of modern sequencing techniques and digital computational technologies (MetaHIT Consortium et al., 2010). The HMP has shown that, in healthy individuals, the microbiome varies across different body sites but is very similar across individuals, with some variations linked to race and ethnicity (Human Microbiome Project Consortium, 2012).

### 1.1.1 The microbiome as an organ

It is interesting to consider the microbiome as a “forgotten” or “hidden” organ, interacting and modulating with the rest of the body and its processes (Baquero and Nombela, 2012; Hou et al., 2022; O’Hara and Shanahan, 2006). The microbiome is critically involved in many physiological processes and biological signalling to distant organs in the body. Through the production of bioactive compounds from dietary macronutrients. Microbiota-derived metabolites communicate with the host immune system and influence essential host responses, much like any other organ in the body (Schroeder and Bäckhed, 2016). It has been said that if the gut microbiome alone could be removed from the body as an organ it would weigh about 2kg, similar to the weight of an adult liver (Patterson et al., 2016; Stephens et al., 2018).

### 1.1.2 Microbiome and host interactions in health and disease

The human microbiome plays a crucial role in influencing the normal structural and functional development of the mucosal immune system and is present during foetal development and continues rapidly in the first few years of life. This process continues to change and adapt to the host environment and vice versa (Flint et al., 2012; Kamada et al., 2013). Mucosal immune responses to resident microbes require precise control to distinguish between commensal and harmful pathogenic bacteria. While increased microbiome diversity and positive health outcomes are somewhat linked the opposite is seen to be the case in many diseases. Reduced diversity or an imbalance or disruption in the composition of the microbial community from “normal” leads to a state of microbial community dysfunction. This is known as dysbiosis (Eckburg et al., 2005). Dysbiosis in the microbiome can contribute to the pathogenesis of multiple disorders, including dental caries, periodontal disease, inflammatory bowel diseases, atopy, infection, diarrhoea, cancer, and arthritis, diabetes and chronic obesity emphasizing the significance of maintaining a balanced and diverse gut microbiota for overall health and well-being (O’Hara and Shanahan, 2006; O’Mahony et al., 2005)

## 1.2 Gut Microbiome in Disease

### 1.2.1 Gut Microbiome in Obesity and Diabetes

There are many studies supporting a complex link between gut microbiota and obesity, diabetes, and other metabolic disorders (Fan and Pedersen, 2021; Patterson et al., 2016). Studies have shown that gut microbiota possess glycoside hydrolase enzymes that cannot be formed from the human genome (Turnbaugh et al., 2006). These enzymes are involved in fermenting dietary polysaccharides, enhancing host energy status, and producing short-chain fatty acids (SCFA), which are crucial to host health and play a key role in the prevention and treatment of metabolic and bowel disorders and certain types of cancer (Scharlau et al., 2009).

While disorders such as obesity are multifactorial there is research to support microbiome changes as one such factor. Germfree mice develop less body fat than "conventionally" raised mice owing to the lack of microbiome diversity needed to support proper and complete absorption of nutrients (Bäckhed et al., 2004). When gut microbiota from "lean" and "obese" mice was introduced to these germfree mice there was a greater increase in body fat percentage of the mice with "obese" microbiota despite no other changes to their environment or dietary intake. This indicates an association between the composition of gut microbiota and body fat accumulation.

Gut microbiota have been shown to produce a lipopolysaccharide (LPS) which enters the bloodstream and is detected in plasma. This concept of "metabolic endotoxemia" (increased plasma LPS levels) is known to cause inflammatory processes that trigger host pathways resulting in diabetes and obesity (Cani et al., 2007a). The increased LPS levels are seen following consumption of a high-fat diet or direct subcutaneous infusions of LPS. In both cases increased insulin resistance and obesity is seen as the end result (Cani et al., 2007b).

Some gut bacterial species however offer protective effects. Prebiotic supplementation with oligofructose (a dietary fibre) dramatically increased the abundance of *Akkermansia muciniphila* (*A. muciniphila*) and improved metabolic function and endotoxemia (Everard et al., 2011). In an ex-vivo study it was shown that *A. muciniphila* and its SCFA regulate transcription factors and genes involved in

cell cycle control, lipolysis and satiety (Lukovac et al., 2014). In this same study *Faecalibacterium prausnitzii* was also identified as potentially protective against obesity though not as strongly as *A. muciniphila*. This type of association is often seen of microbes not exerting effects in isolation, but together in a complex and organised fashion.

### 1.2.2 The Gut Brain Axis

The concept of a microbiome–brain–gut axis has been described to explain the bidirectional interactions of the brain and gut. The brain can influence enteric microbiota via innervation related changes in gastrointestinal motility, secretion and intestinal permeability, or via signalling molecules released into the gut lumen. Equally the gut microbiota can result in the production of signalling molecules that affect the nervous system via endocrine, immune and neural signalling mechanisms (Rhee et al., 2009).

The hypothalamic—pituitary—adrenal (HPA) axis is a neuroendocrine system that can be shaped by exposure during its development to a range of stimuli. It has been shown that depending on the composition of the microbiome the neural regulation of the HPA axis during times of stress is affected. In germfree mice an exaggerated response to stress was seen by the detection of elevated plasma ACTH and corticosterone levels. A similar effect was seen when there was a lack of diversity and predominantly *Escherichia coli* were present. In contrast reconstitution with *Bifidobacterium infantis* helped diversify the microbiome and moderated the HPA response to stress (Sudo et al., 2004).

In humans changes to the gut microbiome has been linked with depression. While depression is a chronic condition with various genetic and environmental factors, the microbiota composition of an individual may be one such factor. Certain strains of *Oscillibacter* spp. have valeric acid as its main metabolic end product, which is a homolog of the neurotransmitter GABA (Naseribafrouei et al., 2014). GABA is known to play a significant role in anxiety and mood disorders.

There is a vast and growing body of knowledge around the role of the gut microbiome in conditions such as autism, anxiety, schizophrenia, Parkinson's disease, and

Alzheimer's disease (Cryan et al., 2019). There is a complex system of bidirectional communication of microbiota with the brain through innervation of the target organ, signalling via the sympathetic and parasympathetic systems, influencing gut motility and permeability, and triggering sensory information from the gut, as well as direct or indirect interactions with the enteric nervous system (Moloney et al., 2014). It is not yet fully known if the microbiome causes these conditions or is simply a factor in exacerbating the presentation of these conditions.

### 1.2.3 Gut Microbiome in Cardiac Diseases

Chronic metabolic diseases such as diabetes and obesity are often seen in conjunction with cardiac diseases such as hypertension, arteriosclerosis, high cholesterol and ischaemic heart disease. This makes It challenging to separate the effects of one distinct disease from another and makes It more difficult to determine causality from effect. However the following studies demonstrate significant microbiota association in various cardiac diseases (Fan and Pedersen, 2021).

There is evidence of a potential beneficial association between *Bacteroides stercoris* (*B. stercoris*) and cardiac health. Individuals with a high level of *B. stercoris* showed lower diastolic blood pressure, which is known to be a protective factor against hypertension (Gaundal et al., 2022). In contrast ischaemic heart failure has been shown to be associated with a dysbiotic gut microbiota with an increased abundance of *Ruminococcus*, *Acinetobacter* and *Veillonella* spp. and a decreased abundance of *Alistipes*, *Faecalibacterium* and *Oscillibacter* spp. (Cui et al., 2018).

Certain microbial metabolites, such as trimethylamine (TMA) and trimethylamine-N-oxide (TMAO) have been associated with an increased risk of atherosclerosis as they can enhance platelet aggregation and induce thrombosis as these metabolites induce a pro-inflammatory state. Gut microbiota synthesize TMA and TMAO from dietary phosphatidylcholine, lecithin and l- carnitine, which are found in meat from animals, poultry, seafood and eggs (Fan and Pedersen, 2021).

Epidemiological studies show positive relationships between the plasma concentrations of TMAO and an increased prevalence of stroke and myocardial infarction as well as an increase in premature mortality (Fan and Pedersen, 2021).

However, in some lab studies contradictory results have been shown, with increased plasma TMAO concentrations but unexpectedly reduced aortic plaque lesion sizes. This effect was observed in lab mice given dietary l-carnitine (needed for TMAO production) (Collins et al., 2016). These conflicting observations can be possibly explained by the variations in host microbiome compositions and diversity (Koay et al., 2021). As seen from previous examples, diet alone is not a standalone component, but the host-microbiota interaction is crucial to the impact on the individual.

Overall, there is evidence to suggest that the gut microbiome has a significant impact on the development of cardiac diseases such as atherosclerosis and ischemic heart disease through the production of microbial metabolites, dietary interactions, and its influence on host physiology.

### 1.2.4 Gut Microbiome and Oral Diseases

Dental caries and periodontal disease are the primary oral “diseases”, and the oral microbiome plays a crucial role in both (Kozak and Pawlik, 2023; Zhang et al., 2022). Both caries and periodontitis are influenced by the balance of microbial communities within the oral cavity. A healthy microbiome is protective against disease, while in dysbiosis the disease associated microbiome is influenced by ‘specialist’ microorganisms that possess metabolic functions and an elevated virulence potential resulting in disease initiation and progression.

In dental caries, there is increased abundance of cariogenic bacteria, such as *Streptococcus mutans* and *Lactobacillus* species, which are capable of fermenting dietary carbohydrates to produce acids. These acids demineralize the tooth enamel and dentin, leading to caries (Zhang et al., 2022). The oral microbiome's composition alone may not fully explain the relationship between caries and microbes, prompting research that combines metagenomics with transcriptomics and metabolomics to provide a comprehensive understanding. In addition to the classical caries prevention strategies such as fluoride treatments and dietary modifications to limit substrates for acidogenic bacteria there is an increased focus on maintaining a

healthy oral microbiome to target the ecological shift that is associated with caries (Baker and Edlund, 2019).

Periodontal diseases are primarily caused by bacterial biofilms forming along the teeth, the gingival crevice and on the root surfaces of teeth. There is a complex interplay of host immune and inflammatory responses to the bacterial presence that result in progressive degradation of bone and connective tissue that support the teeth (Chen et al., 2018; Di Stefano et al., 2022). Some key periodontal pathogens such as *Porphyromonas gingivalis*, *Tannerella forsythia*, and *Treponema denticola* have been termed "red complex" bacteria for their ability to drive periodontal disease pathogenesis (Sedghi et al., 2021).

## 1.3 The Oral Microbiome in Oral Malignancy & OPMD

### 1.3.1 Malignant Oral Lesions

Malignancy is another term used to mean cancer. Cancer is defined as: “a large group of diseases that can start in almost any organ or tissue of the body when abnormal cells grow uncontrollably, go beyond their usual boundaries to invade adjoining parts of the body and/or spread to other organs” (World Health Organization, n.d.).

Head and neck cancers, including malignancies of the lips, tongue, cheeks, floor of the mouth, hard and soft palate, sinuses, and pharynx, emerge as a significant global health concern. In 2022, head and neck cancer ranked as the sixteenth most common cancer worldwide, with oral squamous cell carcinomas (OSCC) accounting for 2% of all cancer incidences, resulting in 389,485 new cases and 188,230 deaths (Bray et al., 2024). Oral cancer is a malignant neoplasia which arises on the lip or oral cavity and is traditionally defined as a squamous cell carcinoma (OSCC) because 90% of oral cancers are histologically originated in the squamous cells (Lingen et al., 2008).

The development of OSCC is influenced by a confluence of environmental and genetic factors. Key risk factors include tobacco use, alcohol consumption, betel quid chewing, excessive sun exposure, and infection with Human Papillomavirus (HPV), with approximately 25% of OSCC patients presenting HPV positivity. Additionally, genetic predispositions contribute to OSCC susceptibility, suggesting a multifactorial

aetiology encompassing lifestyle choices, viral infections, and genetic susceptibility (Rivera, 2015).

OSCC manifests through a range of symptoms, often starting as non-specific or benign appearing alterations in the oral cavity. There may also be areas of epithelial dysplasia from which cancer can develop, or it may develop *de novo*. Some symptoms and signs include (Bagan et al., 2010):

- Persistent swellings, lumps, rough spots, crusts, or eroded areas on the lips, gums, or other areas inside the mouth.
- The appearance of velvety white, red, or speckled patches within the mouth.
- Unexplained bleeding, numbness, loss of feeling, or pain/tenderness in the face, mouth, or neck.
- Persistent sores on the face, neck, or mouth that bleed easily and do not heal within two weeks.
- Difficulties in chewing, swallowing, speaking, or moving the jaw or tongue.
- Changes in voice, such as hoarseness or chronic sore throat.
- Ear pain without an apparent cause.
- Unexplained alterations in dental occlusion or the fit of dentures.

The multifaceted nature of OSCC risk factors and symptoms highlight the importance of comprehensive approaches in prevention, early detection, and management strategies to address the morbidity and mortality associated with oral cancer.

### 1.3.2 Oral Potentially Malignant Disorders

An Oral Potentially Malignant Disorder (OPMD) is defined as “any oral mucosal abnormality that is associated with a statistically increased risk of developing oral cancer” (Warnakulasuriya et al., 2021). Terms such as “premalignant”, “precancerous”, “epithelial precursor lesions”, and “intra-epithelial lesion” are often used but there has been a shift away from these terms towards the use of “OPMD” to refer to these conditions. Additionally, the classification and staging of OPMDs has been refined by the WHO and experts to better frame these oral conditions. OPMDs present with diverse clinical attributes, such as colour variations (white, red, and mixed white-red), morphological changes (plaque/plateau, smooth, grooved,

wrinkled, granular, atrophic), different sizes, and involving different anatomical sites in the oral cavity (Farah et al., 2014).

Oral cancer doesn't always follow a linear progression. Not all OPMDs develop into oral cancer, and some oral cancers develop from non-dysplastic lesions. In other cases some morphological alterations are highly susceptible to malignant transformation; for example in leukoplakia patients, malignancy may arise elsewhere in clinically normal mucosa (Speight, 2007).

The overall prevalence of OPMDs worldwide is about 4.5% (Mello et al., 2018). Malignant transformation (MT) is the developing of a potentially malignant disorder into an invasive cancer.

The cumulative malignant transformation rate (MTR) across all OPMDs is 7.9%. This rate varies from as low as 1.4% for lichen planus up to 49.5% for proliferative verrucous leukoplakia. The MTR per year of lichen planus is 0.28% compared to 9.3% for proliferative verrucous leukoplakia (Iocca et al., 2020).

The most common OPMDs include leukoplakia (LE), lichen planus (LP), oral lichenoid lesions (OLL), oral erythroplakia (OE), oral submucous fibrosis (OSF), and proliferative verrucous leukoplakia (PVL).

### 1.3.3 Site-Specificity and Variability

The oral cavity is a complex ecosystem and hosts a vast array of microbial species with a rich biodiversity unique to its location. Prior to the Oral Microbiome Project, traditional culture-based methodologies had identified approximately 280 bacterial species within this niche (Dewhirst et al., 2010). However, the advent of culture-independent techniques has dramatically expanded our ability to detect the presence of over a 1000 species in the oral cavity (Bai et al., 2022; Olsen and Yamazaki, 2019). Further investigations, pushing the boundaries of microbial diversity, suggest the existence of several thousand species, positioning the oral cavity second only to the colon in terms of microbial complexity (Human Microbiome Project Consortium, 2012).

This oral microbiome's dynamic nature is characterized by both transient and permanent microbial residents, reflective of the oral cavity's openness to the

external environment. The composition of the oral microbiota encompasses a wide spectrum of obligate aerobes, along with facultative and obligate anaerobes, illustrating the diverse metabolic capabilities required to thrive in varying oxygen levels (Paster et al., 2006). The recognition of such microbial diversity not only enriches our understanding of oral biology but also emphasizes the intricate balance required to maintain oral health and the potential implications of microbial dysbiosis in disease pathogenesis.

### 1.3.4 Microbiome and Carcinogenesis

Research overlapping the microbiome and oncology has suggested associations between chronic infections and microbial dysbiosis that could be linked to cancer pathogenesis. It's now understood that approximately 13% of all cancers worldwide are directly attributable to infectious agents, signifying a significant portion of the global cancer burden (Martel et al., 2020). In this context, bacteria have been implicated in the development of a diverse array of malignancies. For instance, *Fusobacterium nucleatum* is associated with colon cancer (Flemer et al., 2018), while various species such as *Bacillus*, *Staphylococcus*, and *Bacteroidetes* have been linked to breast cancer (Vergara et al., 2019). Similarly, *Fusobacterium nucleatum* and *Granulicatella adiacens* have connections to pancreatic cancer (Gaiser et al., 2019), and *Helicobacter pylori* to gastric cancer (Polk and Peek, 2010).

With regards to oral cancer particular bacterial species such as *Streptococcus spp.*, *Prevotella spp.*, *Fusobacterium spp.*, *Porphyromonas gingivalis*, and *Capnocytophaga gingivalis* are strongly linked to oral squamous cell carcinoma (OSCC) (Karpiński, 2019). In other studies periodontal pathogens have been isolated in cases of colorectal and pancreatic cancers, suggesting their potential effects beyond the oral cavity (Flemer et al., 2018; Gaiser et al., 2019). Pathogens like *P. gingivalis* can induce significant dysbiosis contributing to conditions such as periodontal disease, which has been correlated with an increased risk of OSCC. Studies have shown that OSCC tissues exhibit higher levels of periodontal pathogens compared to healthy tissues, highlighting the need to investigate the role of of these microbes in oral carcinogenesis (Nagy et al., 1998; Tezal et al., 2005). An oral tumour, especially one

that has breached the epithelium and is leaking serum could create a similar environment for periodontal pathogens as gingival crevicular fluid.

The link between periodontal disease and an elevated risk of OSCC further strengthens the argument for the microbiome's role in oral cancer. Systematic reviews have shown that periodontal disease and oral dysbiosis can increase the likelihood of developing OSCC, even after adjusting for known risk factors (Amato, 2023; Gopinath et al., 2020).

It has been suggested that 15–20% of human cancers arise from inflammatory processes that promotes cell proliferation and malignant transformation (Colonia-García et al., 2020). Among the complex interactions within the human body, the role of the oral microbiota in cancer development, particularly through mechanisms such as chronic inflammation, the production of carcinogenic metabolites, immune modulation, alterations in epithelial barrier integrity and cell migration (Bai et al., 2022; Sun et al., 2020). Metabolites such as Butyric acid (BA), a short-chain fatty acid, produced by *P. gingivalis* and *F. nucleatum* have been shown to indirectly affect the migration of ameloblastomas through the expression of Epidermal Growth Factor (EGF) and Transforming Growth Factor β1 (Ishikawa et al., 2020). Microbial dysbiosis results in the degradation of epithelial tight junction proteins inducing part of the epithelial-mesenchymal transition (EMT) process. For instance, *P. gingivalis* and *F. nucleatum* have also been shown to regulate epithelial barrier function through the degradation of E-cadherin, EGF and tumour necrosis factor-α (TNF-α) (Abdulkareem et al., 2018).

### 1.3.5 Microbiome Biomarkers

A biomarker or “biological marker” is “an objective and quantifiable measure of a physiological process, pathological process or response to a treatment (excluding measurements of how an individual feels or functions)” (European Medicines Agency, 2024). Biomarkers can be subclassified into diagnostic, monitoring, response, predictive, prognostic, safety and susceptibility biomarkers depending on their application (Califf, 2018). The clinical application of a biomarker involves rigorous validation of analytical methods and demonstration of clinical relevance.

The development of oral microbiome biomarkers in the detection of potentially malignant and malignant conditions presents several challenges. The complex diversity of the oral microbiome makes the identification of specific microbes difficult due to the vast number and the interplay between multiple microbes and in their functional outcomes (Bourgeois et al., 2022). Additionally, there is significant inter-individual variation in oral microbiomes, based on factors such as genetics, diet, lifestyle, and oral hygiene. This variation can make it harder to identify consistent biomarkers across populations (Nearing et al., 2020). There are further challenges as microbiome-based biomarkers differ from "classical" biomarkers due to the complex interactions with the host's immune system and other microbes. These interactions can have varying degrees of impact on biological pathways (R.-J. Li et al., 2021).

Finally, while many studies have identified associations between specific oral microbes and cancer or potentially malignant disorders, establishing a causal relationship between can remain a challenge as it can difficult to determine if the microbiome causes or responds to the clinical outcomes that are measured (Chen et al., 2023; Zhang et al., 2023). For the above reasons as well as some technical and practical barriers there are still no qualified biomarkers in current use despite the potential clinical value.

## 1.4 Microbiome Studies

### 1.4.1 Inclusion and exclusion criteria

Several factors can have an impact on the microbiome that is sampled in an individual. These include antibiotic usage, diet, body mass index, age, pregnancy, and ethnicity. These factors must be eliminated of accounted for to minimise variability in microbiomes due to host differences.

Antibiotic usage and pregnancy are sometimes set as exclusion criteria. The time period since the most recent antibiotic usage is reported at varying length by different studies. The NIH human Microbiome Project excluded subjects who had used systemic antibiotics, antifungals, antivirals, or antiparasitic within 6 months of sampling (NIH HMP Working Group et al., 2009). However other studies have used a

shorter time period for exclusion criteria, as short as 3 months (Nel Van Zyl et al., 2022).

Pregnancy causes the body to undergo many changes and this is no different with regards to the microbiome, where the composition of the gut microbiota changes dramatically during pregnancy (Koren et al., 2012). As such to allow better analysis of results with less confounding factors. Several microbiome studies and their inclusion and exclusion criteria are listed in Table 1.

*Table 1 Inclusion and exclusion criteria mentioned in various studies*

| Study | Inclusion criteria | Exclusion criteria | Notes |
|---|---|---|---|
| (Pushalkar et al., 2012) | - Have a history of smoking and drinking | - antibiotic usage within the last month | |
| (Schmidt et al., 2014) | Smokers, alcohol, immunocompromised | Not mentioned | No clear inclusion/exclusion criteria mentioned |
| (Guerrero-Preston et al., 2016) | Histopathologically confirmed SCC patients | Not mentioned | Controls were healthy patients, non-smoker, non drinkers |
| (Wang et al., 2017) | Not mentioned | Not mentioned | 2003 – 2014 H&N SCC patients |
| (Flemer et al., 2018) | Not listed | History of CRC, IBD and irritable bowel syndrome | |

### 1.4.2 Storage and transport of samples

The gold standard conditions for sample storage is thought to be immediate freezing on dry ice or liquid nitrogen and then storage at -80 Celsius (Human Microbiome Project Consortium, 2012). This can be used in various sequencing and experimental methods such as amplicon, metagenomic, meta-transcriptomic sequencing, and metabolomic measurement but can often be very expensive and sometimes impractical due to the nature of at home stool sampling, which is often associated with varying durations of home freezing before being collected by researchers. Another suggested method is that the samples should be preserved at −20°C within 15 minutes and then transferred to a laboratory on dry ice within 24 hours and stored at −80°C until required for analysis (Gorzelak et al., 2015). Some studies have shown that the actual impact of storage at various temperatures (-20 to -80 Celsius) is quite small (Lauber et al., 2010). However what appears to be more influential to the results of the microbiome sample is the number of freeze thaw cycles (Sergeant et al., 2012). Alternatively samples can be preserved in the kits and stored at ambient temperatures for more than a week (Han et al., 2018). Sample preservation and storage should be consistent to minimize potential confounding variations (Qian et al., 2020). The sampling and storage methods of several studies are shown in Table 2.

*Table 2 Several microbiome studies and their methods for obtaining and storing samples*

| Study | Sample method | Sample storage | Notes |
|---|---|---|---|
| (Pushalkar et al., 2012) | Oral tissue samples obtained from a "tissue bank" | samples procured and stored at −80°C | Oral mucosa processed to include all bacteria |
| (Schmidt et al., 2014) | - Oral swab from lesion and contralateral "normal" side as control.<br>- L+R side tongue in healthy controls | Swab stored on ice for <30mins. | Swab held at 20* to dry lesion and stroked 10 times, swab rotated 180* and repeat. |
| (Guerrero-Preston et al., 2016) | Oral rinse (saline) | Not mentioned | 20mL of saline, gargled for 20+10seconds |
| (Wang et al., 2017) | Tissue samples from biorepository | Flash frozen and stored at –80 °C | Samples taken from tissue biorepository |
| (Flemer et al., 2018) | Oral sample – rub inside of both cheeks with swab | Not mentioned | |

### 1.4.3 Microbiome sequencing methods

The most commonly used sequencing methods are 16S rRNA (amplicon) sequencing, metagenomic sequencing (shotgun) (Goodrich et al., 2014).

While amplicon sequencing is quicker and cheaper to run it will only yield taxonomy profiles, and the PCR process can generate biases. Shotgun metagenomic sequencing provides more detailed genomic information and higher taxonomic resolution than the amplicon sequencing. It can also generate multi-dimensional functional profiles for what the microbes are doing and how they carry out their functions. It also provides more accurate taxonomy, including draft genomes of previously unreferenced microbes (Qian et al., 2020).

# Chapter 2: Study aims

The overall purpose of this study is:

To determine if gut and oral microbiota profiles can be used as biomarkers for the development of oral malignancy.

The aims of this study are to:

1) Observe and describe the gut and oral microbiota of patients with benign, potentially malignant and malignant oral lesions.

2) Investigate the relationship of smoking, alcohol and oral health (through periodontal status) on the gut and oral microbiota.

# Chapter 3: Material and Methods

## 3.1 Synopsis of Study

This is a single-centre prospective observational case-control study approved by the Clinical Research Ethics Committee of University College Cork (UCC) (Reference Number: ECM 4 (k) 28/06/2022) (Appendix A). The study was registered with APC Microbiome Ireland SFI Research Centre of UCC (study number APC167) and sample analysis and bioinformatics were carried at Teagasc Food Research Centre, Moorepark, Fermoy. Teagasc is the national body in Ireland providing integrated research, advisory and training services to the agriculture and food industry and rural communities.

Patients attending for biopsy of oral lesions had their gut and oral microbiota sampled. The participants were grouped into three categories of oral lesions based on biopsy and histological examination: benign (B) oral lesions, potentially malignant (PM) oral lesions, and malignant (M) oral lesions. Microbiota analyses were carried out to investigate if there is a relationship between gut and oral microbiota and its relationship on potentially malignant and malignant oral lesions. The benign oral lesions were chosen as a reference control group or "normal" control, though it is not possible to classify these strictly controls due to the possibility of inflammation and serum leaking in some benign conditions, which can affect microbiome composition.

## 3.2 Study population and recruitment

To determine a suitable sample size we discussed the use of a power calculation. The number of samples required for a microbiome study depends on the effect size. A tool called ''Evident'' has been developed to aid in estimating the sample size required based on the anticipated effect size and similar previous studies (Biocore, 2023). However, because there is no standardized way of reporting effect sizes and other studies were not readily comparable, a pragmatic sample size was chosen based on expert opinion and experience that the research team has in this area.

We recruited 30 participants in this study, from patients that attended for an oral biopsy to Cork University Dental School and Hospital, Cork, Ireland. Study

participation involved two to three visits over a period of two to six weeks. This is the usual time frame for patients attending for routine biopsy appointments. Recruitment began in October 2022 and concluded in October 2023.

### 3.2.1 Inclusion Criteria

Participants were considered eligible for the study if they:

1. Were able to give written informed consent.
2. Were between 18 and 90 years of age.
3. Were in generally good health as determined by the investigator.
4. Had an oral lesion that required biopsy.

### 3.2.2 Exclusion criteria

Participants were excluded from the study if they:

1. Were below 18 years of age or greater than 90 years of age.
2. Had a significant acute or chronic coexisting illness (cardiovascular, gastrointestinal, endocrinological, immunological, metabolic) or a previous oral cancer diagnosis, or any condition which contraindicated, in the investigator's judgement, entry to the study.
3. Were taking medication that the investigator believed would interfere with the objectives of the study, e.g. laxatives, enemas (within last 2 weeks).
4. Were not receiving treatment involving experimental drugs.

### 3.2.3 Recruitment

At the first visit participants were approached about the study (Visit 1). If they were eligible based on the inclusion and exclusion criteria, the details of the study were explained along with a Participant Information Leaflet (Appendix B). Participants were also provided with a Data Protection Notice (Appendix C) detailing what will happen to their personal data and samples; participant were required to read this document as well. Informed consent was then obtained and recorded on the Consent Form (Appendix D).

## 3.3 Clinical protocol

### 3.3.1 Visit One

Following a routine consultation appointment, and once deemed eligible, patients were offered the opportunity to participate in the study. The aim of the study and the procedures to be undertaken were explained to all potential participants. They were provided with a Participant Information Leaflet. Participants were also provided with a Data Protection Notice which detailed what will happen to their personal data and who has access to it. Once they agreed to take part in the study, they read and signed the Consent Form and received a signed copy to keep.

Each participant was assigned an anonymised participant number that was used for all further data collection and sample analysis. The only person who was aware of the identity of the participants and their participant number was the clinical researcher. This was necessary to correlate the participant number to the diagnosis and histological group.

Participants were provided with a Food Frequency Questionnaire (FFQ) (Appendix E) to complete before their next visit. If they are unable to complete the food frequency questionnaire by themselves the researcher will assist them in completing it at the next visit.

Participants were also provided with a faecal sample collection kit (500ml sealed container, Red Biohazard Autoclave bag, and thermally insulated sealed bag with an ice-pack) with written instructions on how to provide a sample, as per the APC167 - Aerobic stool collection instructions (Appendix F).

### 3.3.2 Visit Two

At visit 2 the food frequency questionnaire (FFQ) as well as the stool sample was collected from participants. If participants forgot to bring these, they could bring them at the following visit as well. The following participant data was collected using the Case Report Form (Appendix G): height, weight, BMI calculation, current medications and antibiotic use, periodontal status (Basic Periodontal Exam), smoking status and alcohol consumption. Periodontal status was used as an indicator of oral health. A BPE score was collected for each participant following British Society of

Periodontology guidelines (The British Society of Periodontology, 2019). The single worst BPE score was used to categorise the oral health on a scale ranging from 0 (excellent), 1 (Good) 2 (Moderate), 3 (Poor) to 4 (Very Poor) oral health. If participants had no teeth (edentulous) this was noted, but it was not possible to categorise their oral health.

As an example if someone's BPE score is:

Upper arch 3-1-1

Lower arch 1-2-2

Then their "oral hygiene" score is "3", as it is the worst score, categorising them as "Poor oral health".

Another example would be:

Upper arch 1-0-0

Lower arch 1-2-1

Then their "oral hygiene" score is "2", as it is the worst score, categorising them as "Moderate oral health".

An Oral Saliva sample was obtained before the biopsy. To try and minimise the impact of mouthwashes participants were asked not to brush their teeth or use mouthwash 3 hours before saliva sample collection. Three hours was chosen as a reasonable time for patients attending appointments. In either the morning or afternoon. Oral saliva was gathered from the floor of the mouth for 10 seconds on the right side and 10 seconds on the left side. A sterile swab in plastic polypropylene transport tube (MEUS S.R.L., Italy) was used.

An oral biopsy sample was obtained as scheduled by the dental surgeon as part of the patient's standard of care. A single biopsy was taken at the appropriate site. Biopsy performed were either incisional, where only a portion of the lesion is removed for analysis, or excisional, involving the complete removal of the lesion. The decision to carry out and incisional or excisional biopsy was made following a history and clinical examination. Factors such as size, location, and clinical suspicion regarding the nature of the lesion were taken into consideration.

Patient consent was obtained for the procedure and local anaesthetic was administered (Lidocaine Hydrochloride 2% w/v Adrenaline (epinephrine) 1:80,000)

using a 27 Gauge needle in a dental syringe. For incisional biopsies a 5mm biopsy punch was used to obtain the sample (Figure 1). In the case of an excisional biopsy complete removal of the lesion was carried out using a #15 scalpel blade (Figure 2).

*Figure 1 Biopsy punch 5mm diameter*

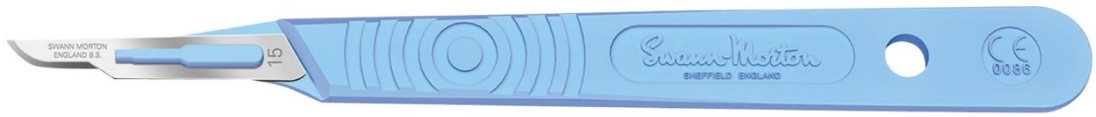



*Figure 2 Sterile #15 blade with disposable handle*

The biopsy sample was then sectioned equally into two pieces using a scalpel. One half was placed in a storage pot of 10% buffered Formalin and sent to histopathology. The other half of the sample was prepared for the microbiome analyses in this research (Oral Tissue sample). This sample was stored in a Sterile flat bottom tube with O-ring cap 2ml (Eppendorf AG, Germany) with 600 µL of RNAlater® (Sigma-Aldrich, USA) solution. This was done to preserve the genetic material contained from the biopsy site.

Both the Oral Saliva and Oral Tissue samples were then stored at -20°C within 30 minutes of sampling.

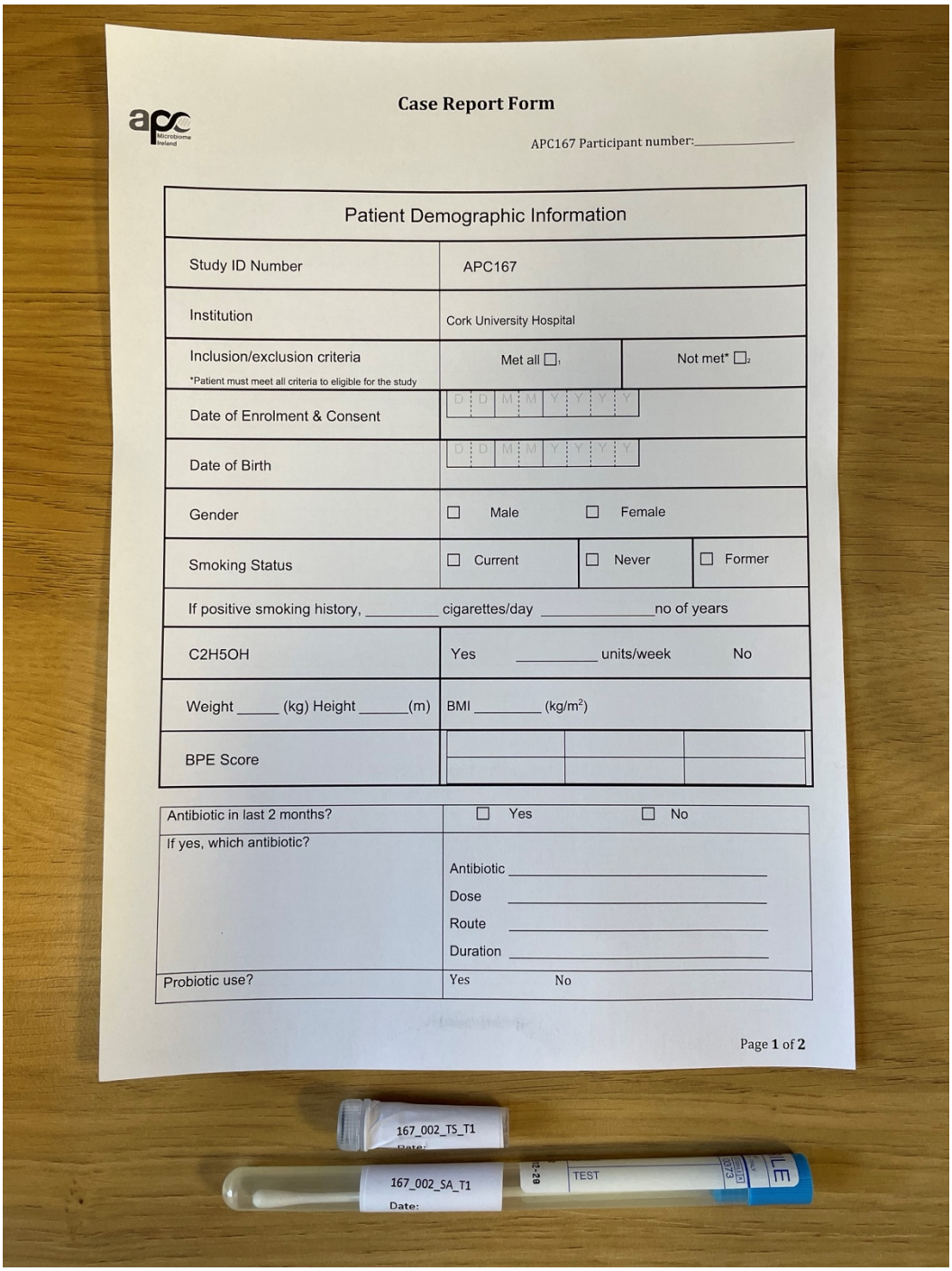
### 3.3.3 Visit Three

Biopsy results were delivered to the patient at this visit as part of their routine care. For patients who were diagnosed with a potentially malignant condition the follow up and recall was managed by the Oral Medicine team as appropriate. For patients had a suspicious lesion on visit 1 or who had a diagnosis of a malignant condition we requested that the patient attended with a close person of support during the clinical visit. A dedicated nurse was present to support the patient emotionally and they were offered water, tea and biscuits as appropriate. The follow up and treatment of these patients was closely communicated to the Head and Neck Cancer multi-disciplinary team urgently.

For the purposes of this study and based on the results of the biopsy, participants were allocated into one of three lesion groups: benign (B) oral lesions, potentially malignant (PM) oral lesions, and malignant (M) oral lesions. If participants forgot to bring the FFQ or Stool sample, it was collected during this visit.

## 3.4 Storage and transport

All of the above samples were stored at -20°C initially for up to four weeks and then transferred to Teagasc Moorepark for storage at -80°C for further downstream analysis. The samples were kept in thermally insulated bags and moved from -20°C to -80°C in under 60 minutes.



**Case Report Form**

APC167 Participant number:_____

| Patient Demographic Information | | | |
|---|---|---|---|
| Study ID Number | APC167 | | |
| Institution | Cork University Hospital | | |
| Inclusion/exclusion criteria<br>*Patient must meet all criteria to eligible for the study | Met all ☐1 | Not met* ☐2 | |
| Date of Enrolment & Consent | D D M M Y Y Y Y | | |
| Date of Birth | D D M M Y Y Y Y | | |
| Gender | ☐ Male | ☐ Female | |
| Smoking Status | ☐ Current | ☐ Never | ☐ Former |
| If positive smoking history, _____ cigarettes/day _____no of years | | | |
| C2H5OH | Yes _____ units/week | No | |
| Weight _____ (kg) Height _____(m) | BMI _____ ($kg/m^2$) | | |
| BPE Score | | | |

| | |
|---|---|
| Antibiotic in last 2 months? | ☐ Yes ☐ No |
| If yes, which antibiotic? | Antibiotic _____<br>Dose _____<br>Route _____<br>Duration _____ |
| Probiotic use? | Yes No |

Page **1** of **2**

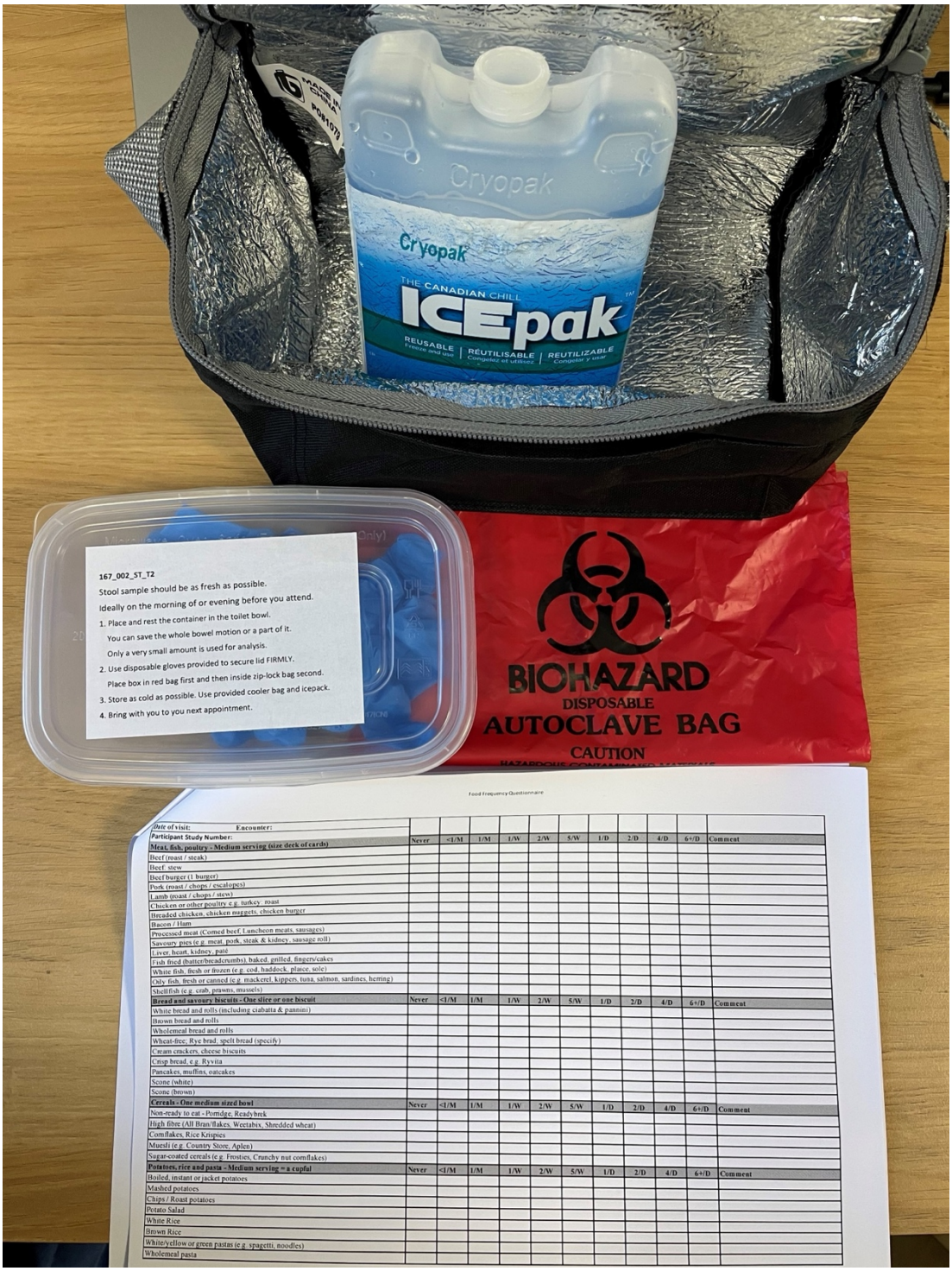
167_002_TS_T1

167_002_SA_T1

Date:

TEST

*Figure 3 Data and Samples collected at visit 2*




| Date of visit: Encounter: | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Participant Study Number: | Never | <1/M | 1/M | 1/W | 2/W | 5/W | 1/D | 2/D | 4/D | 6+/D | Comment |
| **Meat, fish, poultry - Medium serving (size deck of cards)** | | | | | | | | | | | |
| Beef (roast / steak) | | | | | | | | | | | |
| Beef. stew | | | | | | | | | | | |
| Beef burger (1 burger) | | | | | | | | | | | |
| Pork (roast / chops / escalopes) | | | | | | | | | | | |
| Lamb (roast / chops / stew) | | | | | | | | | | | |
| Chicken or other poultry e.g. turkey roast | | | | | | | | | | | |
| Breaded chicken, chicken nuggets, chicken burger | | | | | | | | | | | |
| Bacon / Ham | | | | | | | | | | | |
| Processed meat (Corned beef, Luncheon meats, sausages) | | | | | | | | | | | |
| Savoury pies (e.g. meat, pork, steak & kidney, sausage roll) | | | | | | | | | | | |
| Liver, heart, kidney, paté | | | | | | | | | | | |
| Fish fried (batter/breadcrumbs), baked, grilled, fingers/cakes | | | | | | | | | | | |
| White fish, fresh or frozen (e.g. cod, haddock, plaice, sole) | | | | | | | | | | | |
| Oily fish, fresh or canned (e.g. mackerel, kippers, tuna, salmon, sardines, herring) | | | | | | | | | | | |
| Shellfish (e.g. crab, prawns, mussels) | | | | | | | | | | | |
| **Bread and savoury biscuits - One slice or one biscuit** | Never | <1/M | 1/M | 1/W | 2/W | 5/W | 1/D | 2/D | 4/D | 6+/D | Comment |
| White bread and rolls (including ciabatta & panini) | | | | | | | | | | | |
| Brown bread and rolls | | | | | | | | | | | |
| Wholemeal bread and rolls | | | | | | | | | | | |
| Wheat-free, Rye brad, spelt bread (specify) | | | | | | | | | | | |
| Cream crackers, cheese biscuits | | | | | | | | | | | |
| Crisp bread, e.g. Ryvita | | | | | | | | | | | |
| Pancakes, muffins, oatcakes | | | | | | | | | | | |
| Scone (white) | | | | | | | | | | | |
| Scone (brown) | | | | | | | | | | | |
| **Cereals - One medium sized bowl** | Never | <1/M | 1/M | 1/W | 2/W | 5/W | 1/D | 2/D | 4/D | 6+/D | Comment |
| Non-ready to eat - Porridge, Readybrek | | | | | | | | | | | |
| High fibre (All Bran/flakes, Weetabix, Shredded wheat) | | | | | | | | | | | |
| Cornflakes, Rice Krispies | | | | | | | | | | | |
| Muesli (e.g. Country Store, Aplen) | | | | | | | | | | | |
| Sugar-coated cereals (e.g. Frosties, Crunchy nut cornflakes) | | | | | | | | | | | |
| **Potatoes, rice and pasta - Medium serving = a cupful** | Never | <1/M | 1/M | 1/W | 2/W | 5/W | 1/D | 2/D | 4/D | 6+/D | Comment |
| Boiled, instant or jacket potatoes | | | | | | | | | | | |
| Mashed potatoes | | | | | | | | | | | |
| Chips / Roast potatoes | | | | | | | | | | | |
| Potato Salad | | | | | | | | | | | |
| White Rice | | | | | | | | | | | |
| Brown Rice | | | | | | | | | | | |
| White/yellow or green pastas (e.g. spagetti, noodles) | | | | | | | | | | | |
| Wholemeal pasta | | | | | | | | | | | |

*Figure 4 Data and Samples collected at visit 3*

## 3.5 Lab protocol

### 3.5.1 DNA extraction

DNA extraction of saliva/oral cavity samples was conducted using the DNeasy Powersoil Pro Kit (QIAGEN, UK). Swabs were aseptically cut and placed into PowerBead tubes containing 60 µL of solution C1. The tubes were shaken horizontally at 3x1000 oscillations/min for 5 min using a Mini-Beadbeater (BioSpec Products, USA), and centrifuged for 1 min at 15000 g. Swabs were removed post-vortexing in the initial stage of DNA extraction, and the subsequent steps followed the manufacturer's instructions.

DNA extraction of fecal samples employed an optimised protocol combining a repeat bead beating (RBB) method and the QIAmp Fast DNA Stool Mini Kit (Qiagen, UK) (Yu and Morrison, 2004). In brief, 0.25 g of stool samples was incubated with 1 mL lysis buffer (500 mM NaCl, 50 mM tris-HCl pH 8.0, 50 mM EDTA and 4 % sodium dodecyl sulphate (SDS)) in sterile zirconia bead-beating tubes. The samples were horizontally homogenised at 3x1000 oscillations/min for 3 min using a Mini-Beadbeater (BioSpec Products, USA), followed by incubation at 70 °C for 15 min, and centrifugation at 16000 g for 4 min. The RBB steps were repeated with 0.3 mL lysis buffer for cell lysis. Supernatants were pooled and incubated with 10 M ammonium acetate. DNA precipitation was achieved using isopropanol, followed by washing with 70 % (v/v) ethanol, and re-suspension in Tris-EDTA buffer. Further purification utilized RNAse, Proteinase K, and buffer AL from the QIAmp Fast DNA Stool Mini Kit (Qiagen, UK). DNA was cleaned with Qiagen columns, AW1 and AW2 buffers, and eluted in AE buffer.

DNA quantity and quality were assessed using the QUBIT dsDNA BR Assay Kit (Invitrogen, Ireland) and Nanodrop 1000 (Thermo Scientific, Ireland), respectively. The eluted DNA was stored at -32°C until further experiment.

### 3.5.2 Illumina shotgun sequencing

Illumina DNA Prep Reference Guide (Illumina, 2021) was used for the preparation of shotgun libraries in accordance with manufacturer’s instructions. Notably, DNA inputs were normalised to 10.2 ng in 30 µL before genomic DNA tagmentation. The

tagmented DNA was amplified using appropriate index adapters and PCR mix with the following conditions: 68°C x 3 min, 98°C x 3 min, 8 cycles of 98°C x 45 s, 62°C x 30 s, 68°C x 2 min, then 68°C x 1 min and held at 10 °C.

The quality and quantity of individual DNA libraries were measured using the Agilent Technology 2100 Bioanalyzer with a High Sensitivity DNA kit and the QUBIT dsDNA HS Assay Kit (Invitrogen, Ireland), respectively. The DNA libraries were pooled to equimolarity. The quality of the pool was further verified using the Agilent 2100 Bioanalyzer with a High Sensitivity DNA kit before dilution and denaturation steps for the NextSeq 2000 system, following the recommended protocols. The pooled DNA library was outsourced to Teagasc sequencing facility where it was sequenced on a NovaSEq 6000 S1 300 cycle flow cell according to instructions from Illumina.

### 3.5.3 Laboratory steps

Detailed laboratory steps for the above methods can be found in (Appendix H)

## 3.6 Bioinformatics and statistical analysis

The metagenomic (shotgun) sequencing raw reads were subjected to a quality control process using FastQC (v0.11.8) and MultiQC (v1.9). This was followed by trimming and filtering processes using Trimmomatic (v0.39) and bowtie2 with KneadData (v0.10) program from the Huttenhower lab. Host genome (*Homo sapiens*) decontamination was done under the default settings provided by the KneadData program from the Huttenhower lab (“KneadData,” 2024).

Taxonomic classification at the species level was conducted using MetaPhlAn4 v4.0.6 (Blanco-Míguez et al., 2023). Reads that were not confidently assigned to any marker gene or when the alignment was ambiguous were labeled as "unclassified reads". These unclassified reads were treated as a group for subsequent analysis. The functional analysis was performed with HUMAnN (v3.8), both utilizing standard parameters.

Analysis of which taxa classes drive the difference between the three groups in the present study was done using Songbird (Morton et al., 2019). Songbird for differential abundance analysis was performed using default parameters to study the

bacterial strains most associated with each group and driving the differences in the groups. Only the top 15 associations (both positive and negative) were used for further analysis.

The outputs from MetaPhlAn4 were further processed for in-depth analysis within the R environment, making use of phyloseq, and ggplot2 packages to facilitate profiling and visual representation of the data.

Statistical analysis was performed in R using PERMANOVA to check the proportion of explained variance and significance of each variable using pairwiseAdonis (v0.4). Comparisons between groups and sample types were made using Wilcoxon test and PERMANOVA using Adonis and pariwiseAdonis functions in R. Statistical significance was determined by 999 permutations and p-values were corrected using Benjamin Hochberg method; p-values below 0.05 were considered significant.

# Chapter 4: Results

## 4.1 Description of the cohort

Table 3 summarizes the breakdown of data collected from the 30 participants in this study (n=30). Demographic information (Demo form), saliva sample (SA) and tissue sample (TS) was collected from all 30 participants. The stool sample (ST) was collected from 25 participants, while five participants did not provide ST samples. The food frequency questionnaire (FFQ) was returned by 27 participants, with three participants not returning the FFQs.

Following histological diagnoses 15 participants (50%) were categorised as Benign (B), 8 participants (26.6%) as Potentially Malignant (PM), and 7 participants (23.3%) as Malignant (M), as shown in Table 4.

Examples of participants in the benign (B) category are shown in Figure 5 and Figure 6, Potentially Malignant (PM) in Figure 7 and Figure 8, and Malignant (M) in Figure 9 and Figure 10.

*Table 3 Data collected for each participant*

| Participant ID | Demo form | SA Sample | TS Sample | ST sample | FFQ | Histology diagnosis |
|---|---|---|---|---|---|---|
| APC167_001 | yes | yes | yes | yes | yes | Potentially Malignant |
| APC167_002 | yes | yes | yes | **no** | yes | Benign |
| APC167_003 | yes | yes | yes | **no** | **no** | Potentially Malignant |
| APC167_004 | yes | yes | yes | yes | yes | Potentially Malignant |
| APC167_005 | yes | yes | yes | yes | yes | Malignant |
| APC167_006 | yes | yes | yes | yes | yes | Potentially Malignant |
| APC167_007 | yes | yes | yes | yes | yes | Benign |
| APC167_008 | yes | yes | yes | yes | yes | Potentially Malignant |
| APC167_009 | yes | yes | yes | yes | yes | Benign |
| APC167_010 | yes | yes | yes | yes | yes | Potentially Malignant |
| APC167_011 | yes | yes | yes | yes | yes | Potentially Malignant |
| APC167_012 | yes | yes | yes | yes | yes | Benign |
| APC167_013 | yes | yes | yes | **no** | yes | Benign |
| APC167_014 | yes | yes | yes | **no** | **no** | Benign |
| APC167_015 | yes | yes | yes | yes | yes | Benign |
| APC167_016 | yes | yes | yes | yes | yes | Benign |
| APC167_017 | yes | yes | yes | yes | yes | Benign |
| APC167_018 | yes | yes | yes | yes | yes | Benign |
| APC167_019 | yes | yes | yes | yes | yes | Benign |
| APC167_020 | yes | yes | yes | yes | yes | Malignant |
| APC167_021 | yes | yes | yes | yes | yes | Benign |
| APC167_022 | yes | yes | yes | yes | yes | Malignant |
| APC167_023 | yes | yes | yes | yes | yes | Malignant |
| APC167_024 | yes | yes | yes | yes | yes | Benign |
| APC167_025 | yes | yes | yes | yes | yes | Malignant |
| APC167_026 | yes | yes | yes | yes | yes | Benign |
| APC167_027 | yes | yes | yes | yes | yes | Benign |
| APC167_028 | yes | yes | yes | yes | yes | Malignant |
| APC167_029 | yes | yes | yes | yes | yes | Potentially Malignant |
| APC167_030 | yes | yes | yes | **no** | **no** | Malignant |

*Table 4 Participants in each histological group*

| Histological Diagnosis | Number of participants |
|---|---|
| Benign (B) | 15 |
| Potentially Malignant (PM) | 8 |
| Malignant (M) | 7 |

Table 5 describes the summary of demographics data collected in the Case Report Form (Appendix I). The ages of the participants ranged between 24 years and 78 years, with the mean age being 56.37 years (SD= 12.85). Of the participants 12 (40%) were female (F) and 18 (60%) were male (M). BMI was calculated from measured weight and height. The BMI of the participants ranged between 18 kg/$m^2$ and 38.28 kg/$m^2$, with the mean BMI being 27.13 kg/$m^2$ (SD= 5.03).

Oral health status, based on BPE scores, ranged from Good to Very Poor. 2 (6.67%) participants had no teeth and could not have a BPE score assigned. There were 3 (10%) with Good oral health, 8 (26.67%) with Moderate oral health, 13 (43.33%) with Poor oral health, and 4 (13.33%) with Very Poor oral health.

Of the 30 participants 4 (13.3%) had taken antibiotics in the preceding two months. There were 13 (43.33%) non-smokers, 10 (33.33%) were former smokers, and 7 (23.33%) were current smokers. With regards alcohol consumption, 11 participants reported that they did not consume any alcohol. The remaining 19 participants who did report to consume alcohol, reported a range between 1 and 40 units, with the mean units consumed being 12.79 units (SD= 11.69).

*Table 5 Participant demographics*

| Participant ID | Group | Age | Sex | BMI (kg/m$^2$) | Oral Health | Antibiotic use* | Smoker | Alcohol** |
|---|---|---|---|---|---|---|---|---|
| APC167_001 | PM | 31 | M | 28.7 | Moderate | No | Former | 0 |
| APC167_002 | B | 39 | M | 33.8 | Poor | No | Current | 12 |
| APC167_003 | PM | 60 | F | 28.9 | Moderate | No | Current | 2 |
| APC167_004 | PM | 71 | F | 26.9 | Good | No | Never | 10 |
| APC167_005 | M | 61 | M | 29.4 | Good | No | Never | 2 |
| APC167_006 | PM | 45 | F | 38.2 | Poor | No | Former | 2 |
| APC167_007 | B | 51 | M | 30.1 | Poor | No | Current | 30 |
| APC167_008 | PM | 78 | M | 24.8 | no teeth | No | Former | 16 |
| APC167_009 | B | 49 | F | 18.0 | Poor | No | Current | 15 |
| APC167_010 | PM | 68 | F | 34.0 | Poor | No | Current | 5 |
| APC167_011 | PM | 60 | M | 30.4 | Moderate | No | Never | 0 |
| APC167_012 | B | 62 | M | 28.1 | Moderate | No | Never | 20 |
| APC167_013 | B | 43 | F | 32.7 | Very Poor | Yes | Former | >40 |
| APC167_014 | B | 44 | M | 29.9 | Poor | No | Current | 10 |
| APC167_015 | B | 24 | M | 21.6 | Moderate | No | Never | >1 |
| APC167_016 | B | 43 | M | 26.4 | Poor | No | Never | 12 |
| APC167_017 | B | 57 | F | 33.5 | Moderate | No | Never | 0 |
| APC167_018 | B | 64 | F | 18.3 | Good | No | Never | 0 |
| APC167_019 | B | 62 | M | 26.0 | Moderate | No | Former | 0 |
| APC167_020 | M | 63 | F | 27.6 | Poor | No | Never | 0 |
| APC167_021 | B | 51 | M | 29.5 | Poor | No | Never | 10 |
| APC167_022 | M | 59 | F | 21.4 | Moderate | No | Former | 0 |
| APC167_023 | M | 45 | F | 21.9 | Poor | No | Former | 0 |
| APC167_024 | B | 76 | F | 18.5 | Very Poor | No | Never | 0 |
| APC167_025 | M | 64 | M | 27.4 | Poor | No | Former | 0 |
| APC167_026 | B | 59 | M | 29.6 | no teeth | Yes | Former | 4 |
| APC167_027 | B | 53 | M | 19.1 | Very Poor | Yes | Former | 0 |
| APC167_028 | M | 75 | M | 29.5 | Poor | Yes | Never | 8 |
| APC167_029 | PM | 65 | M | 26.3 | Very Poor | No | Never | 4 |
| APC167_030 | M | 69 | M | 22.9 | Poor | No | Current | >40 |

* in the previous 2 months

** units per week

## 4.2 Clinical photos

### 4.2.1 Benign lesions

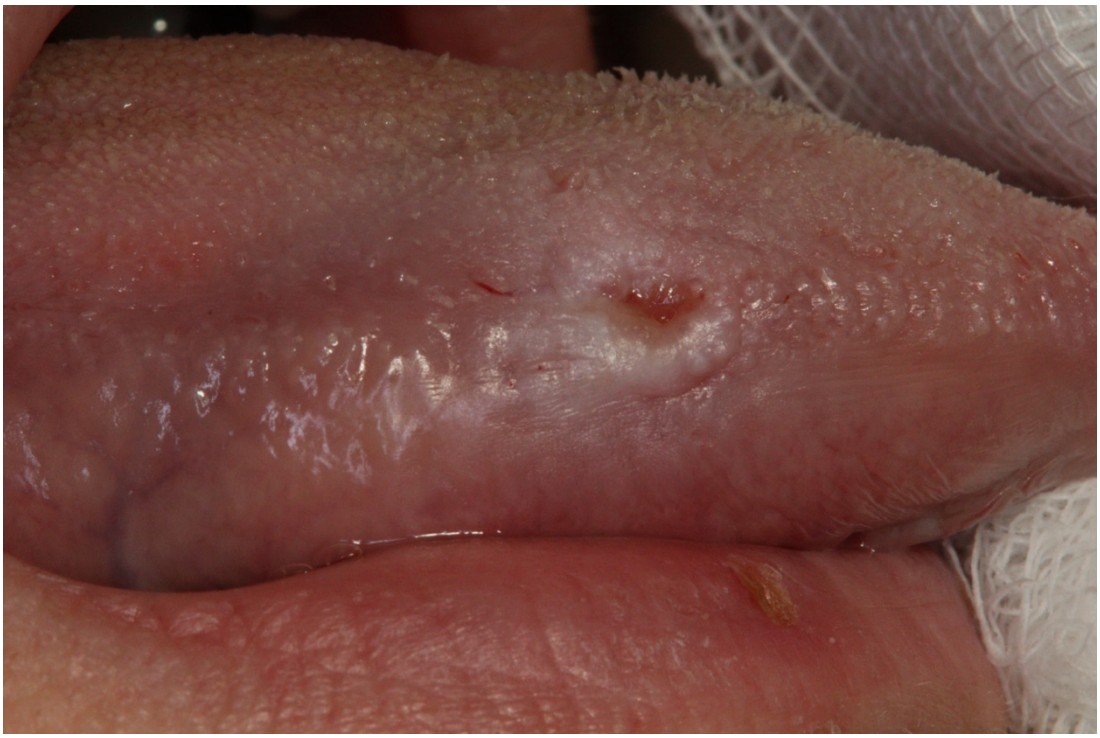


*Figure 5 APC167_007 – traumatic ulcer on side of tongue*

This traumatic ulcer was seen on the ventrolateral surface of the tongue next to an adjacent sharp tooth. The ulcer remained for more than 2 weeks after the sharp surface was smoothened, thus requiring incisional biopsy for definitive diagnosis.

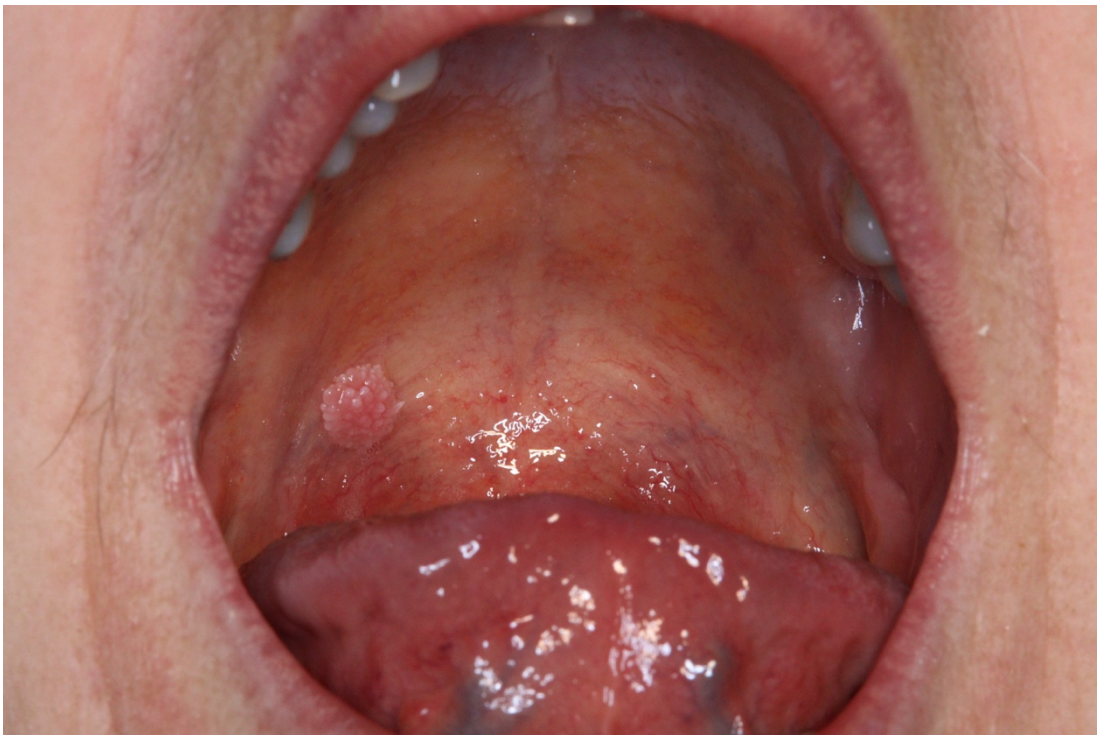


*Figure 6 APC167_017 - Squamous Papilloma on soft palate*

A Squamous Papilloma seen on the right soft palate. As this lesion did not display any worrying features an excisional biopsy was carried out to remove it whole.

### 4.2.2 Potentially malignant lesions

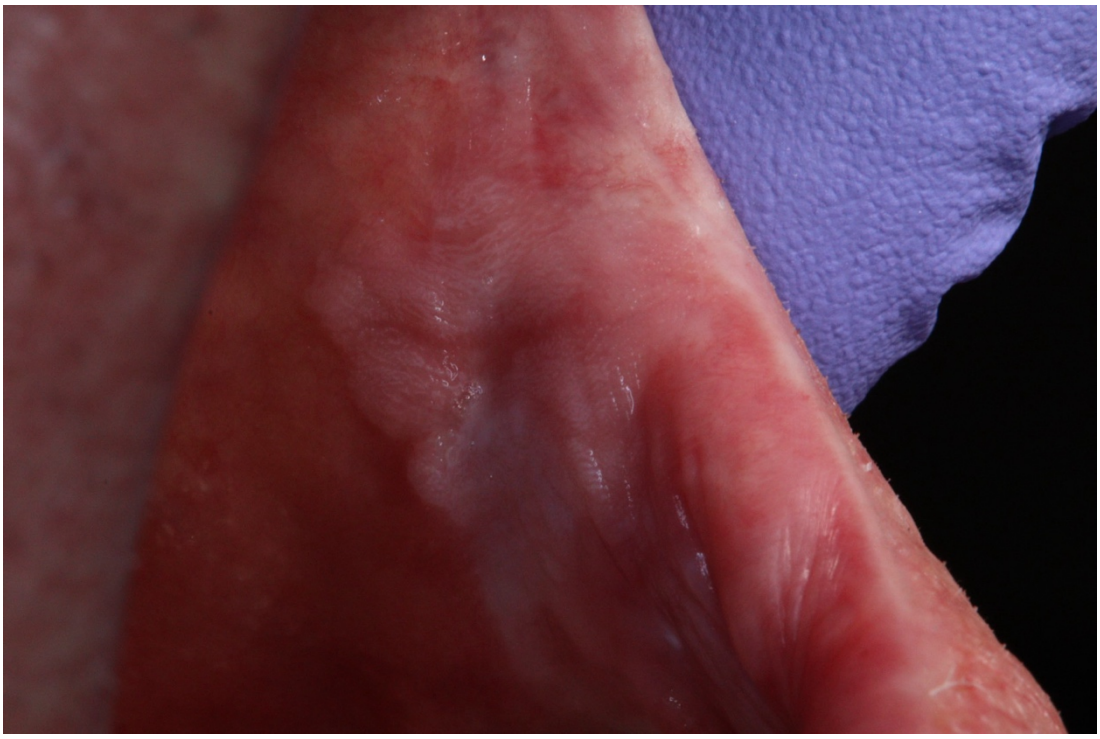


*Figure 7 APC167_008 - oral lichen planus*

This striated white patch was seen on the oral mucosa along the left commissure of the mouth. Histology confirmed a diagnosis of oral lichen planus.

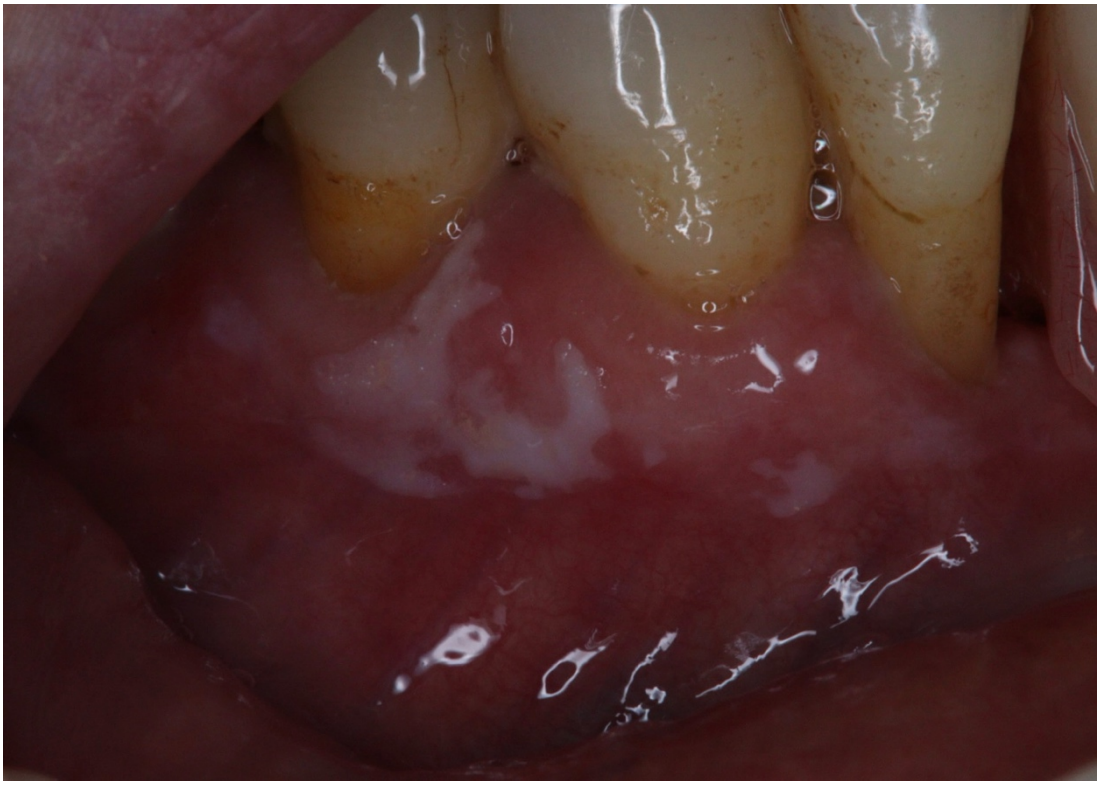


*Figure 8 APC167_010 - oral lichen planus*

This white patch was seen along the attached gingival mucosa in the mandible. Histology following an incisional biopsy confirmed a diagnosis of oral lichen planus.

### 4.2.3 Malignant lesions

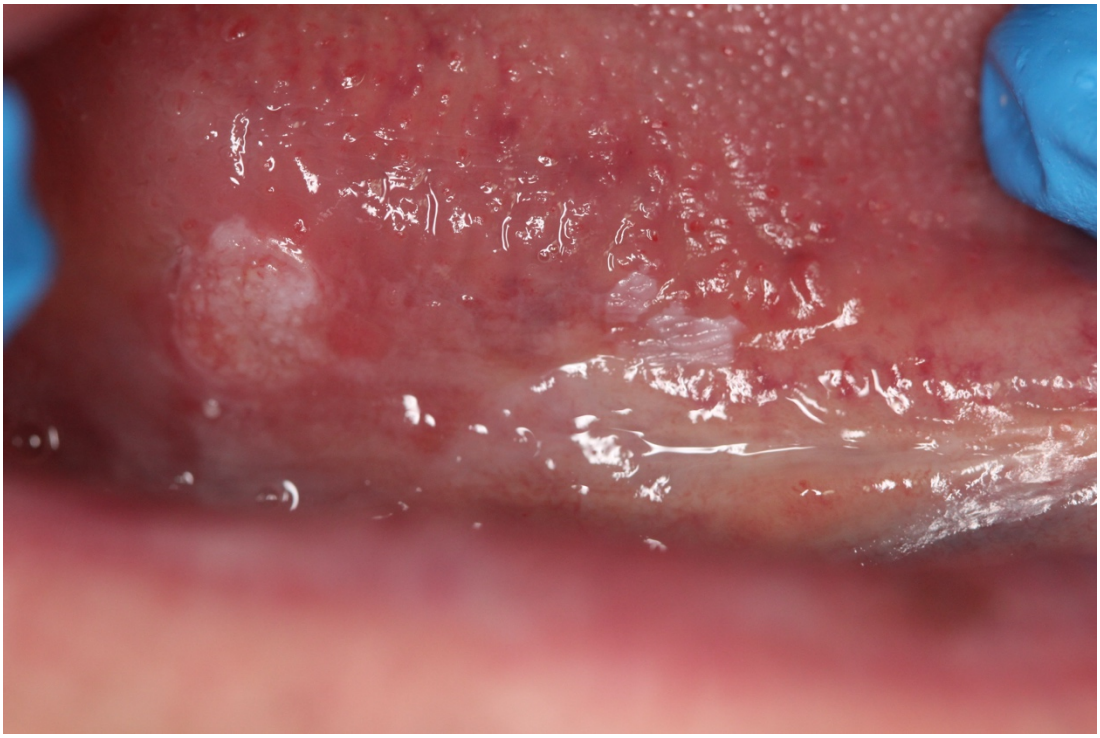


*Figure 9 APC167_005 - squamous cell carcinoma*

This ulcer was seen on the ventrolateral surface of the tongue next to a white plaque. It could not be attributed to any direct cause (e.g. sharp tooth). The ulcer had been present for several weeks. The patient was a non-smoker and alcohol intake was very low. An incisional biopsy confirmed a diagnosis of squamous cell carcinoma (SCC).

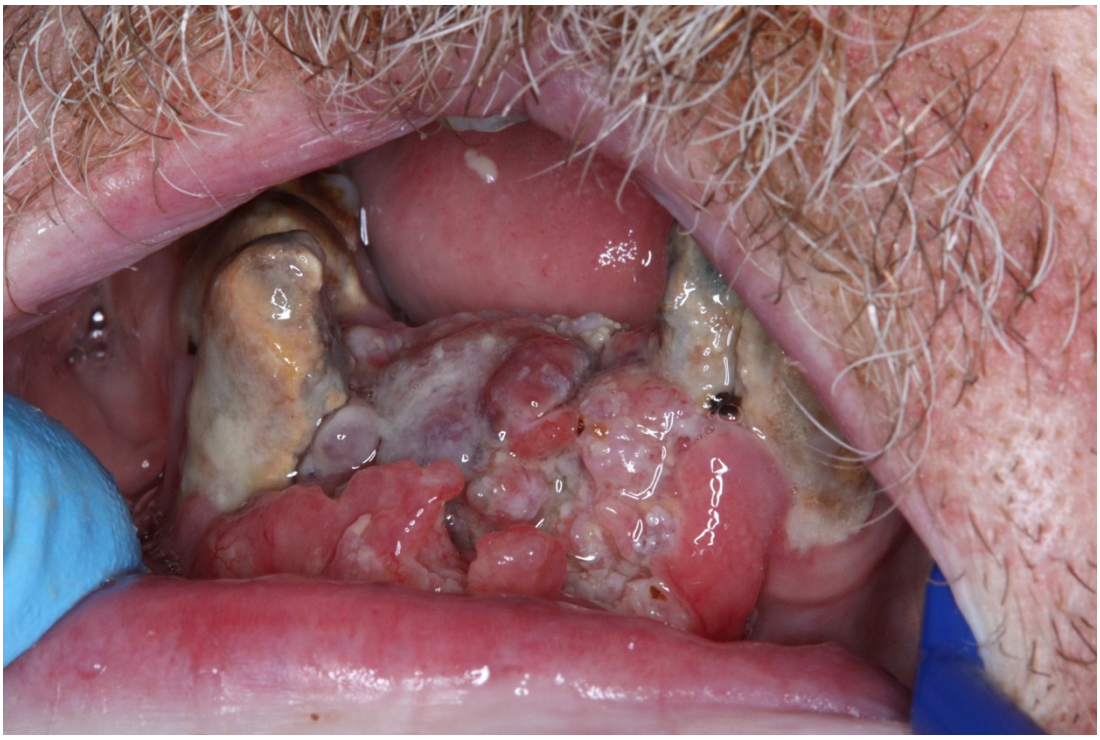


*Figure 10 APC167_030 - squamous cell carcinoma*

This soft tissue mass was seen in the anterior mandible of an elderly gentleman. It had been present for several months and was causing pain, numbness and difficulty eating, drinking and speech. The gentleman was a poor attender to his doctor and dentist. The mass was seen to be eroding into the lingual aspect on the mandible along the floor of mouth. Incisional biopsy confirmed a squamous cell carcinoma (SCC).

## 4.3 Alpha Diversity

The taxonomic compositions of the oral and gut microbiomes were first investigated based on alpha diversity using Shannon and Simpson diversity index. The Shannon index measures the diversity of species within a community considering both species richness (the number of different species present) and species evenness (how evenly the individuals are distributed among the different species). A higher Shannon index value indicates higher diversity within the microbial community, while a lower value indicates lower diversity. Similarly, the Simpson index quantifies the diversity of species within a community focusing more on the dominance or evenness of species rather than just the richness and evenness. Higher values of Simpson index indicate less diversity and more dominance of one or a few species, whereas lower values suggest higher diversity and more evenness among species (Lloréns-Rico et al., 2021).

In order to characterise the microbiomes and to understand the differences between the histological groups statistical comparisons between histological groups and sample types were made using Wilcoxon test and PERMANOVA using Adonis and pairwise Adonis functions in R. P-values were corrected using Benjamin Hochberg method and P-adjusted value < 0.05 were considered significant. Further to this we also attempted to establish if there was a relationship between diet, BMI, smoking, alcohol intake and oral health on the gut and oral microbiota.

### 4.3.1 Alpha Diversity - Overall

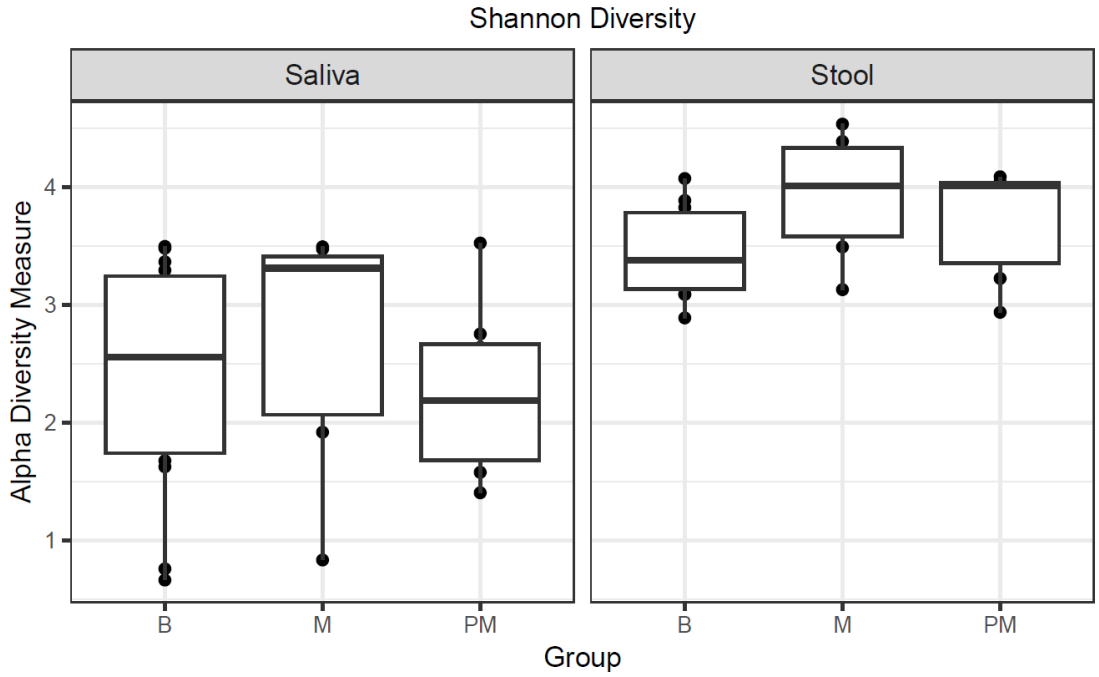



*Figure 11 – Alpha diversity as measured using Shannon diversity index for Saliva and Stool samples, comparing the histological groups benign (B), potentially malignant (PM) and malignant (M).*

The alpha diversity between histological groups for each sample type did not highlight any significant differences in diversity. There were no significant differences observed in the Shannon diversity index of the three groups, benign (B), potentially malignant (PM) and malignant (M), in either saliva or stool samples (Figure 11). Similar to Shannon diversity index, Simpson diversity index also did not show any significant difference between the groups in each sample type (figure not shown).

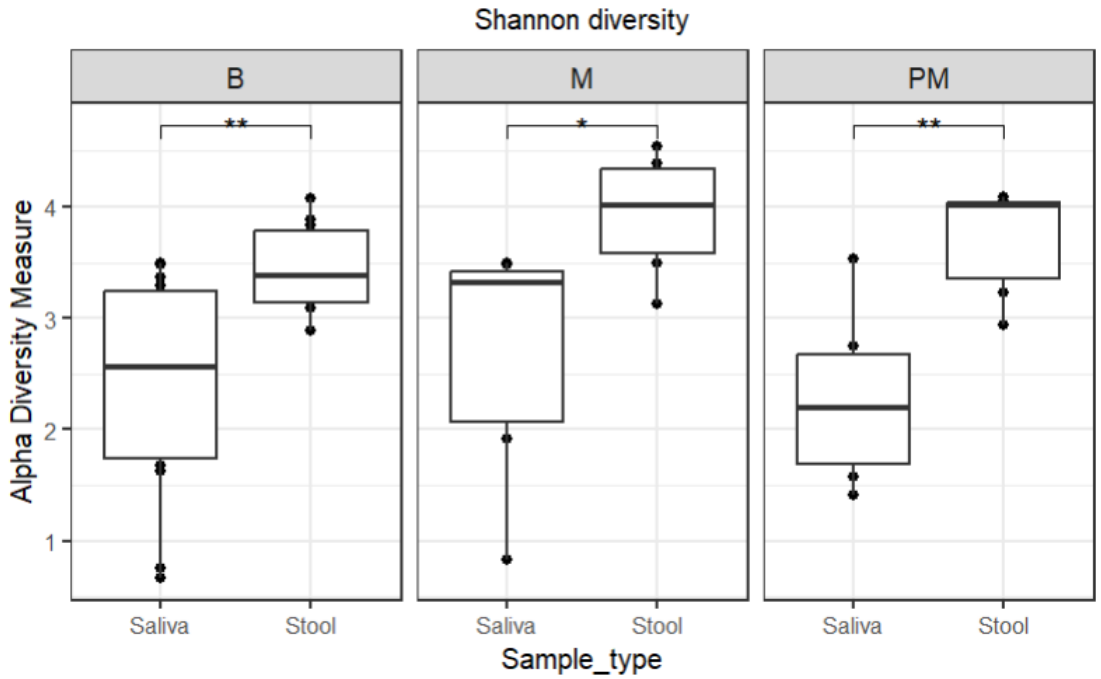



*Figure 12 - Alpha diversity as measured using Shannon diversity index for each histological group, comparing the Saliva and Stool samples. * Significance codes: '***' 0.001; '**' 0.01; '*' 0.05; '.' 0.1.*

Interestingly, the alpha diversity between saliva and stool samples for each histological groups reported significant differences in the Shannon diversity index observed between the three groups, (B), (PM) and (M) (Figure 12). In particular, the stool samples showed higher Shannon index values indicating higher diversity within the microbial community, when comparing with saliva samples.

### 4.3.2 Alpha Diversity – based on smoking status

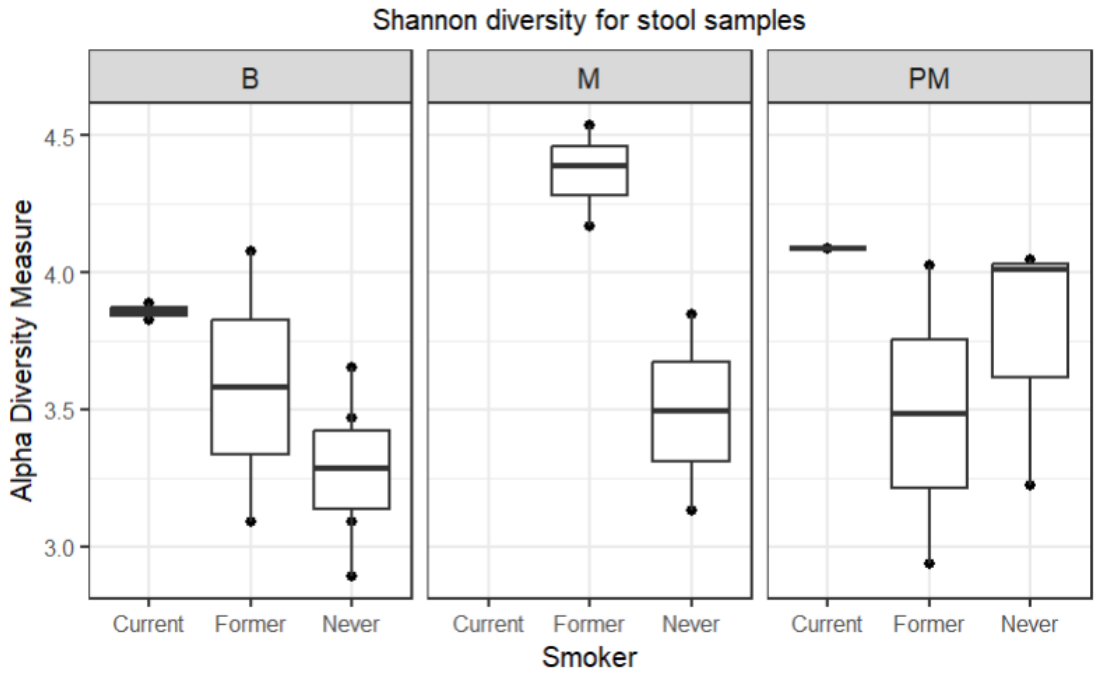



*Figure 13 - Alpha diversity as measured using Shannon diversity index for Stool samples comparing the smoking status of each histological group*

We looked at the diversity between groups for Stool samples, based on cigarette smoking status and did not observe any significant differences in diversity.

There were no significant differences observed between in the Shannon Diversity index of the three groups, (B), (PM) and (M), in smokers, former smoker or non-smokers (Figure 13).

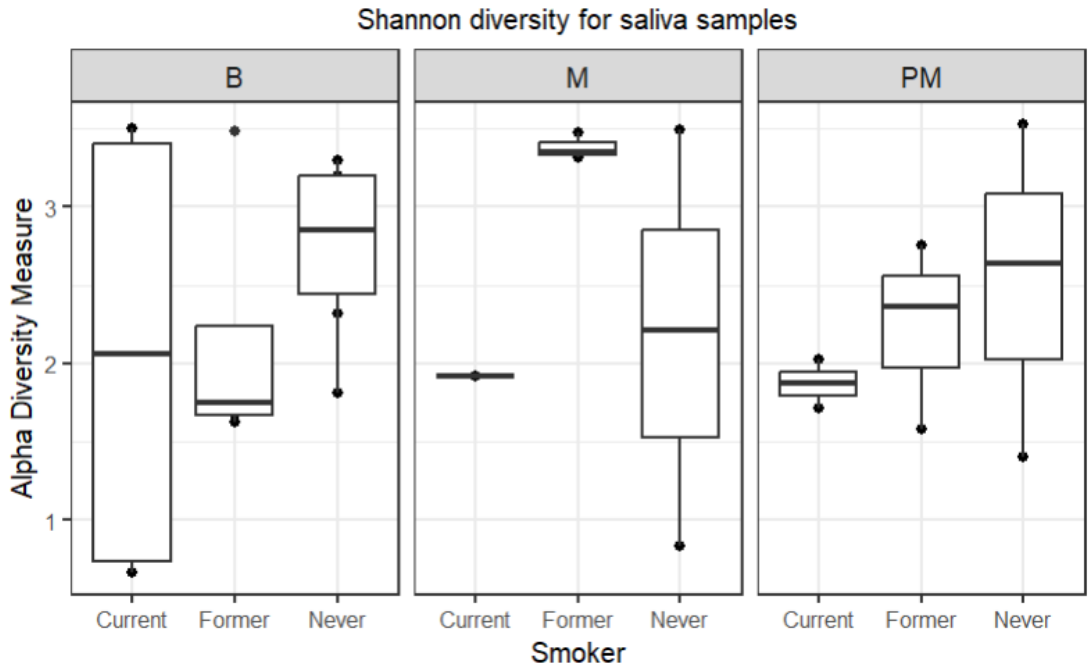



*Figure 14 - Alpha diversity as measured using Shannon diversity index for Saliva samples comparing the smoking status of each histological group*

We looked at the diversity between groups for Saliva samples, based on cigarette smoking status and did not observe any significant differences in diversity.

There were no significant differences observed between in the Shannon Diversity index of the three groups, (B), (PM) and (M), in smokers, former smoker or non-smokers (Figure 14).

## 4.4 Beta Diversity

The compositional characterization of the oral and gut microbiomes was then investigated based on beta diversity. Beta diversity was calculated using Jaccard and Bray-Curtis distance matrices and was plotted using PCoA in R.

Comparisons between histological groups and sample types were made using Wilcoxon test and PERMANOVA using Adonis and pairwise Adonis functions in R. P-values were corrected using Benjamin Hochberg method and P-adjusted value < 0.05 were considered significant.

### 4.4.1 Beta Diversity - Overall

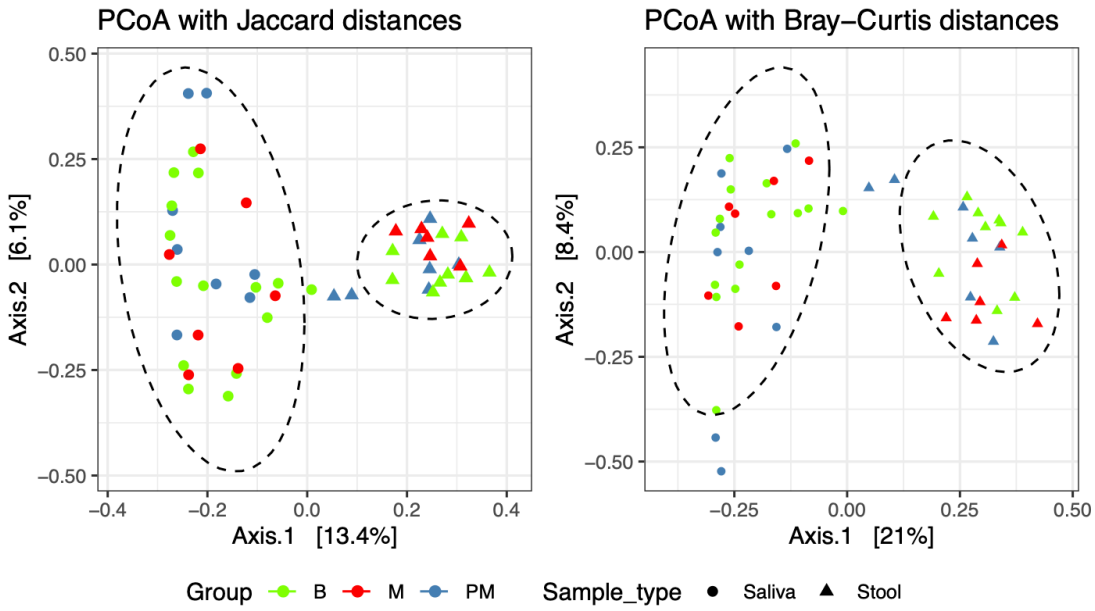



*Figure 15 – Beta diversity calculated with Jaccard and Bray-Curtis distances visualised using PCoA highlighting the differences of taxonomic variances between histological groups for Saliva and Stool samples.*

A clear difference in the microbial composition was observed between the stool and saliva samples, observed using beta diversity with PCoA using Jaccard and Bray-Curtis distance matrices (p=0.001, Adonis) (Figure 15).

### 4.4.2 Beta Diversity - Stool Samples

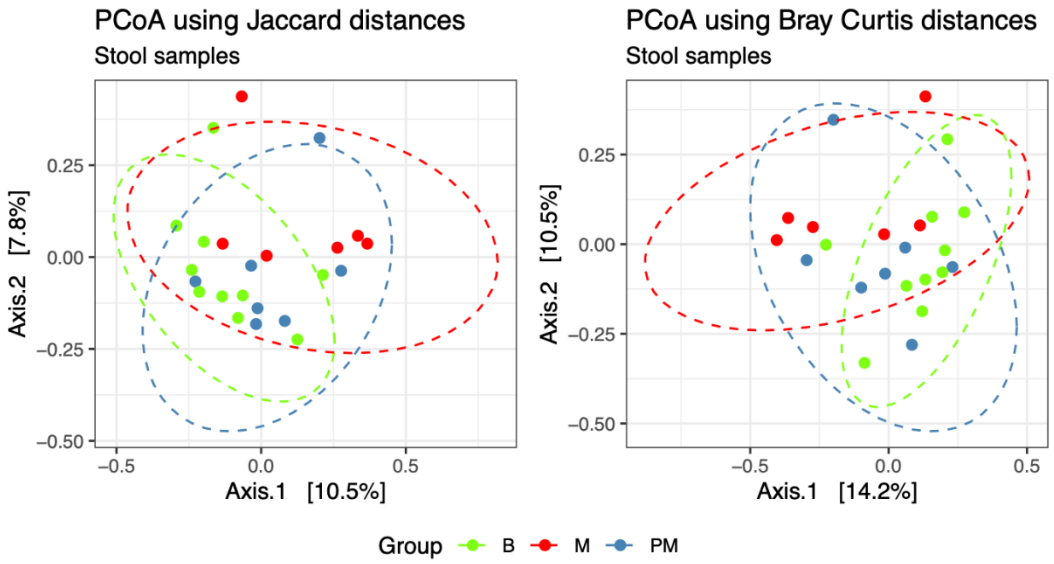



*Figure 16 - Beta diversity calculated with Jaccard and Bray-Curtis distances visualised using PCoA highlighting the differences of taxonomic variances between histological groups for Stool samples*

We further examined the beta diversity between our study groups in the stool and saliva samples. For stool samples, we observed distinct clustering between the benign (B) and malignant (M) groups (p.adjusted=0.045); while no other differences were observed (Figure 16).

*Table 6 Permutation Test: Multivariate Analysis – Stool samples*

| **Study Variable** | **$R^2$** | **p-value** |
|---|---|---|
| Sex | 0.039 | 0.899 |
| Alcohol Consumption | 0.049 | 0.21 |
| Antibiotic Use | 0.043 | 0.54 |
| Smoker | 0.097 | 0.176 |
| Oral Health | 0.133 | 0.537 |

As the variables listed in Table 6 can be confounding factors for oral and gut microbiomes we used the Adonis test for stool samples to evaluate the effects of the following parameters on beta diversity: Gender, Alcohol consumption, Antibiotic use, Oral Health. None of the variables had significant effect on beta diversity, with oral health explaining the highest variance amongst the variables recorded. A summary including P-values based on Jaccard distances is shown in Table 6.

### 4.4.3 Beta Diversity - Saliva Samples

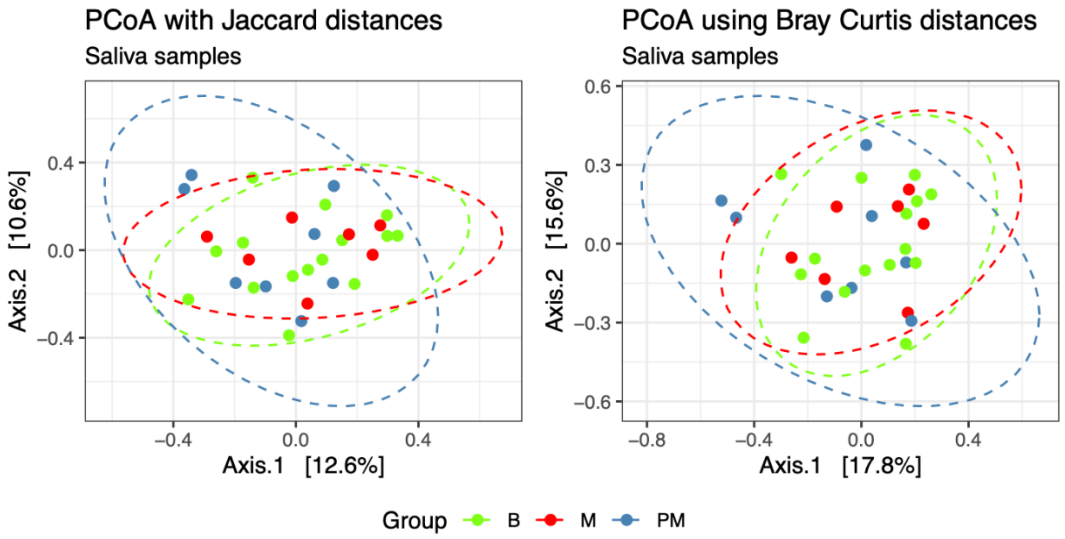



*Figure 17 - Beta diversity calculated with Jaccard and Bray-Curtis distances visualised using PCoA highlighting the differences of taxonomic variances between histological groups for Saliva samples*

Similar to stool samples we also looked at the difference in microbial composition for saliva samples. No significant differences in beta diversity were observed between any of the histological groups (Figure 17).

*Table 7 Permutation Test: Multivariate Analysis – Saliva samples*

| **Study Variable** | **$R^2$** | **p-value** |
|---|---|---|
| Sex | 0.039 | 0.426 |
| Alcohol Consumption | 0.036 | 0.334 |
| Antibiotic Use | 0.082 | 0.093 |
| Smoker | 0.066 | 0.551 |
| Oral Health | 0.111 | 0.253 |

As the variables listed in Table 7 can be confounding factors for oral and gut microbiomes we used the Adonis test for saliva samples to evaluate the effects of the following parameters on beta diversity: Gender, Alcohol consumption, Antibiotic use, Oral Health. None of the variables had significant effect on beta diversity. A summary including P-values based on Jaccard distances is shown in Table 7

## 4.5 Taxonomy

### 4.5.1 Relative abundance

Differences in the relative abundance of microbial composition between the three histological groups in both sample types were observed using plots at phylum and species levels.

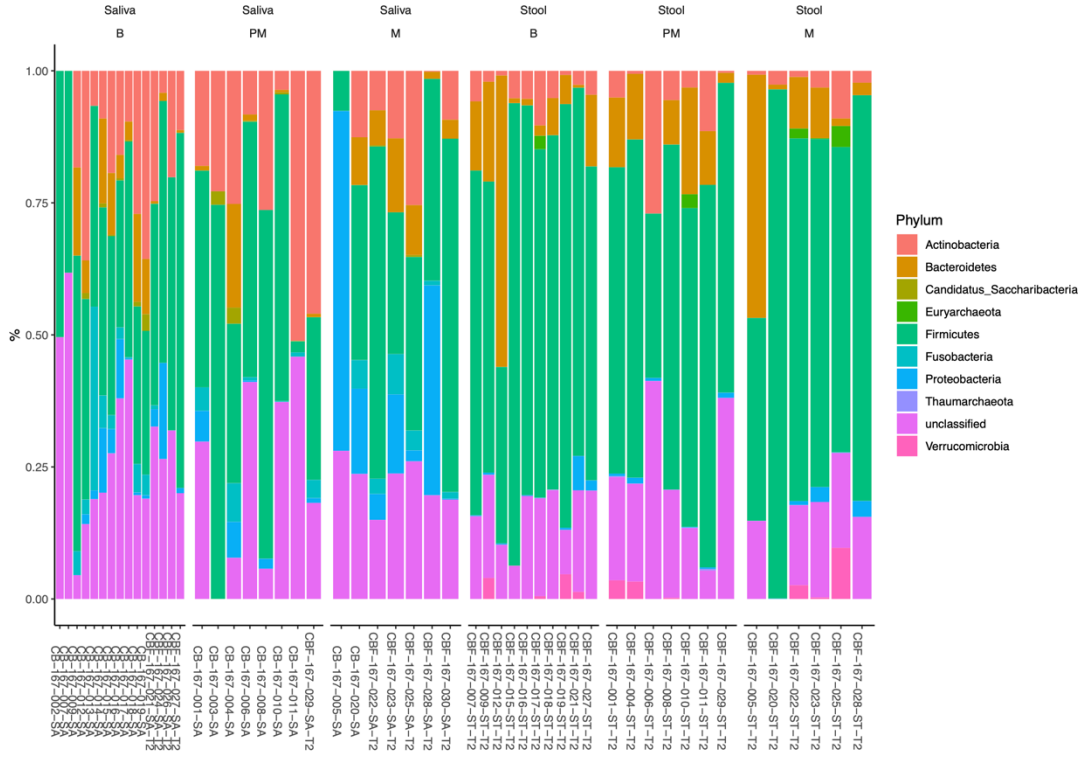



*Figure 18 Phylum relative abundance- Barplot*

Overall, the phylum Firmicutes was observed to have high relative abundance in all groups with higher relative abundance in stool samples. The next most abundant phylum was Actinobacteria. (Figure 18).

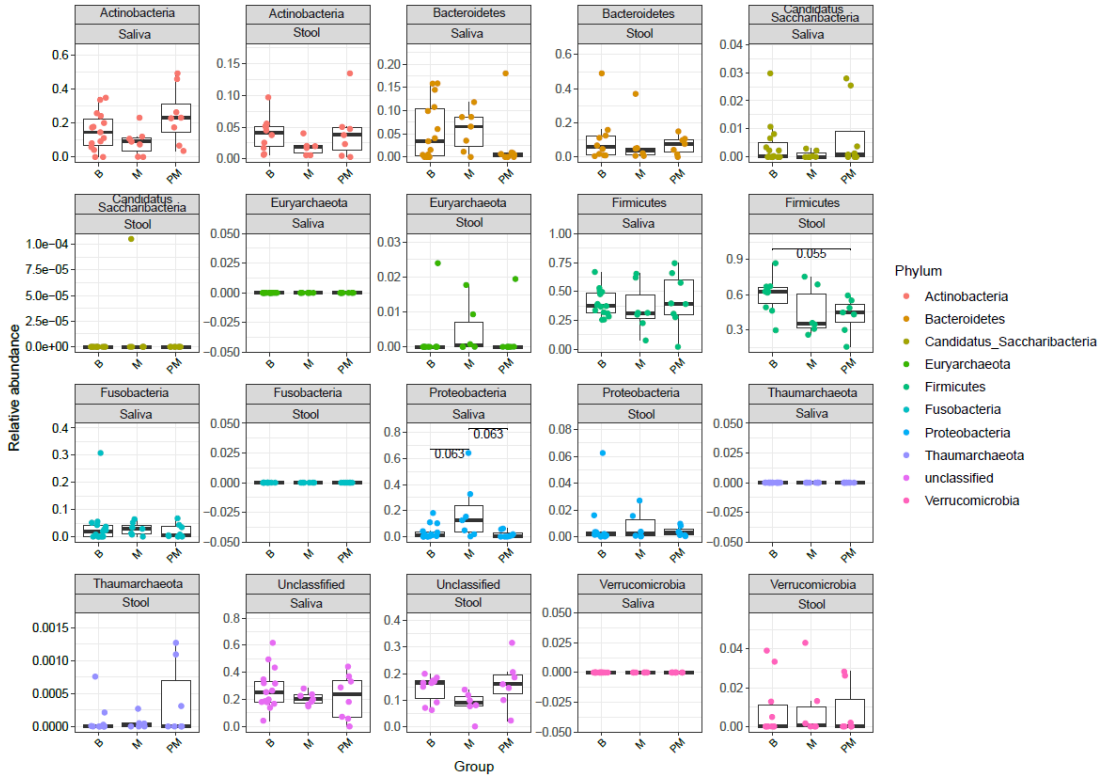



*Figure 19 Phylum relative abundance- Boxplot*

In stool samples, Firmicutes were observed to be highest in the Benign (B) group as compared to potentially malignant (p.adjusted=0.05) group (Figure 19). In saliva samples, Proteobacteria abundance was higher in the malignant group compared to benign (p.adjusted=0.06) and potentially malignant (p.adjusted=0.06) group (Figure 19).

### 4.5.2 Top 15 Strains

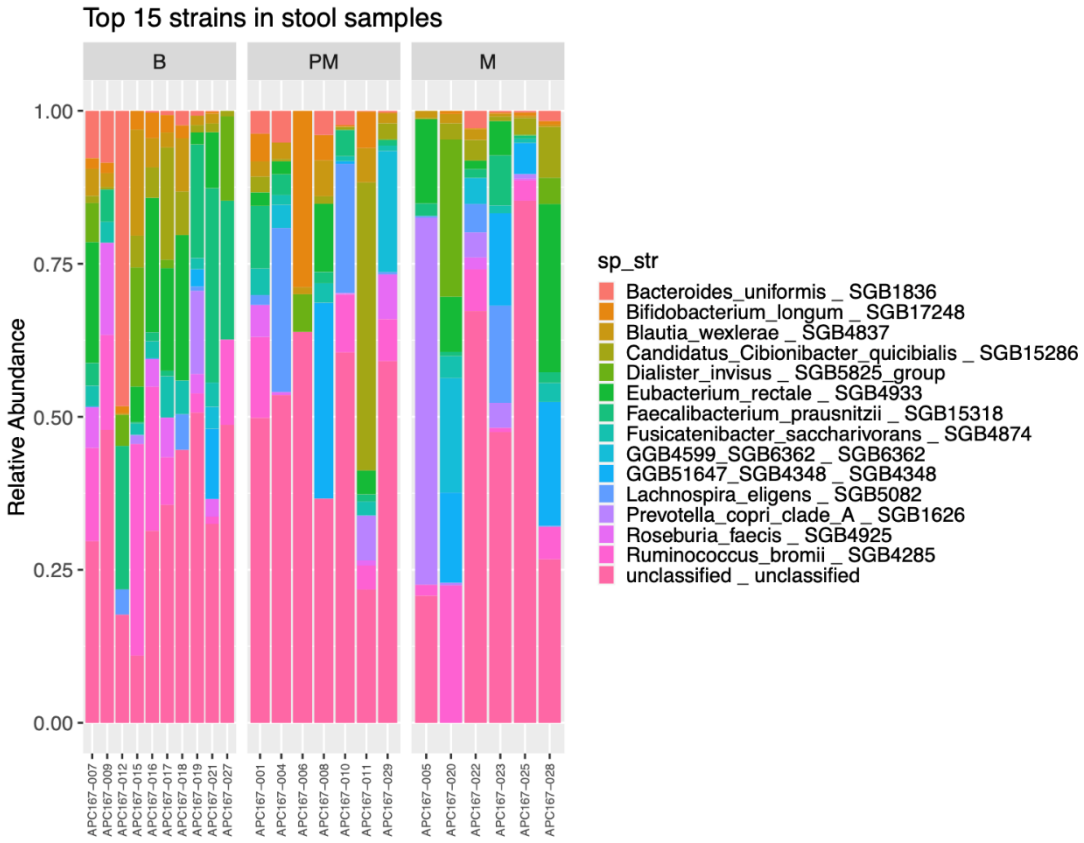



*Figure 20 Top 15 strains in Stool*

In benign (B) samples, a diverse population is observed with relatively even representation among strains. The transition to potentially malignant (PM) samples is marked by a notable shift in microbial composition. Certain strains, such as *Dialister invisus* and *Ruminococcus bromii*, appear to increase in relative abundance. In malignant (M) samples, the dominance of particular strains becomes more pronounced. *Prevotella copri* clade and *Ruminococcus bromii* exhibit an increase in their relative abundance. Conversely, the representation of *Bifidobacterium longum* is noticeably diminished in these samples compared to benign and potentially malignant conditions (Figure 20).

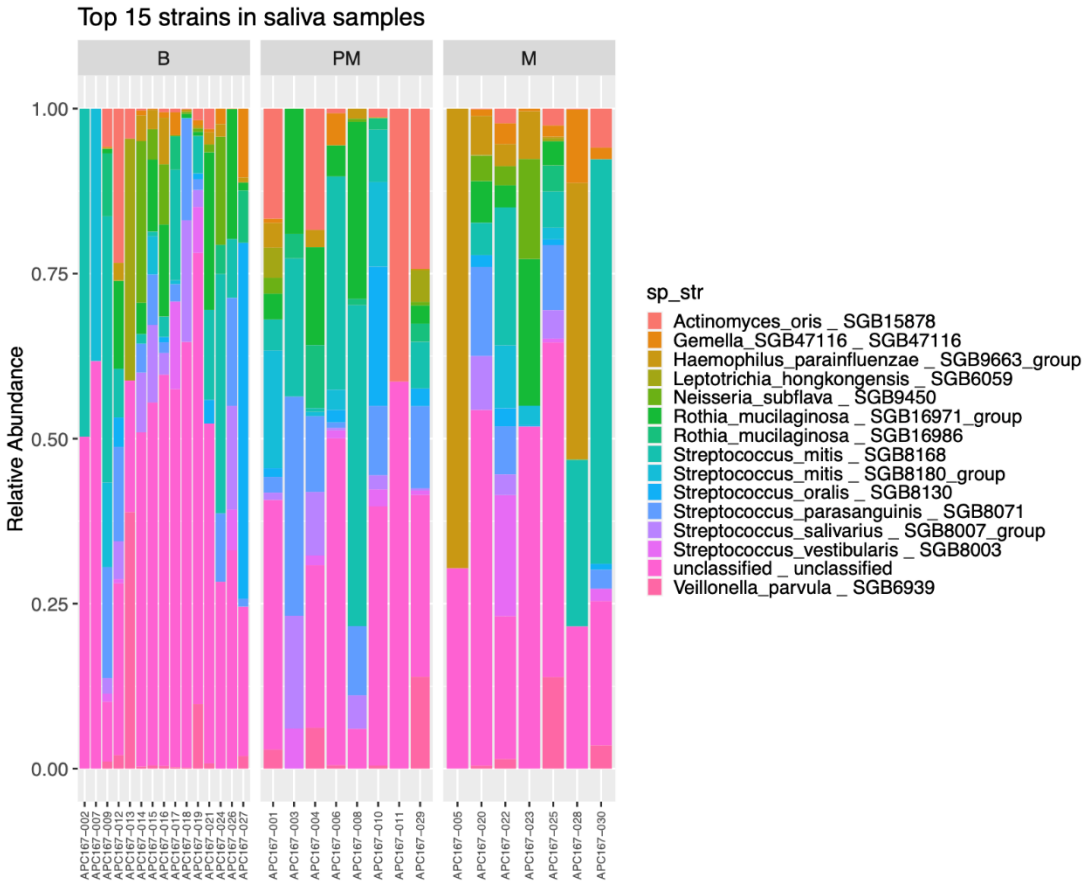



*Figure 21 Top 15 strains in Saliva*

In benign (B) samples, a diverse population is observed with relatively even representation among strains. The transition to potentially malignant (PM) samples is marked by a notable shift in microbial composition. Certain strains, such as *Streptococcus mitis* and *Rothia mucilaginosa*, appear to increase in relative abundance. In malignant (M) samples, the dominance of one particular strain is harder to distinguish. *Haemophilus parainfluenzae* was dominant in one sample while *Streptococcus mitis* was extremely abundant in another. The overall abundance is less evenly distributed in the malignant (M) group (Figure 21).

### 4.5.3 Taxonomic classification of saliva samples

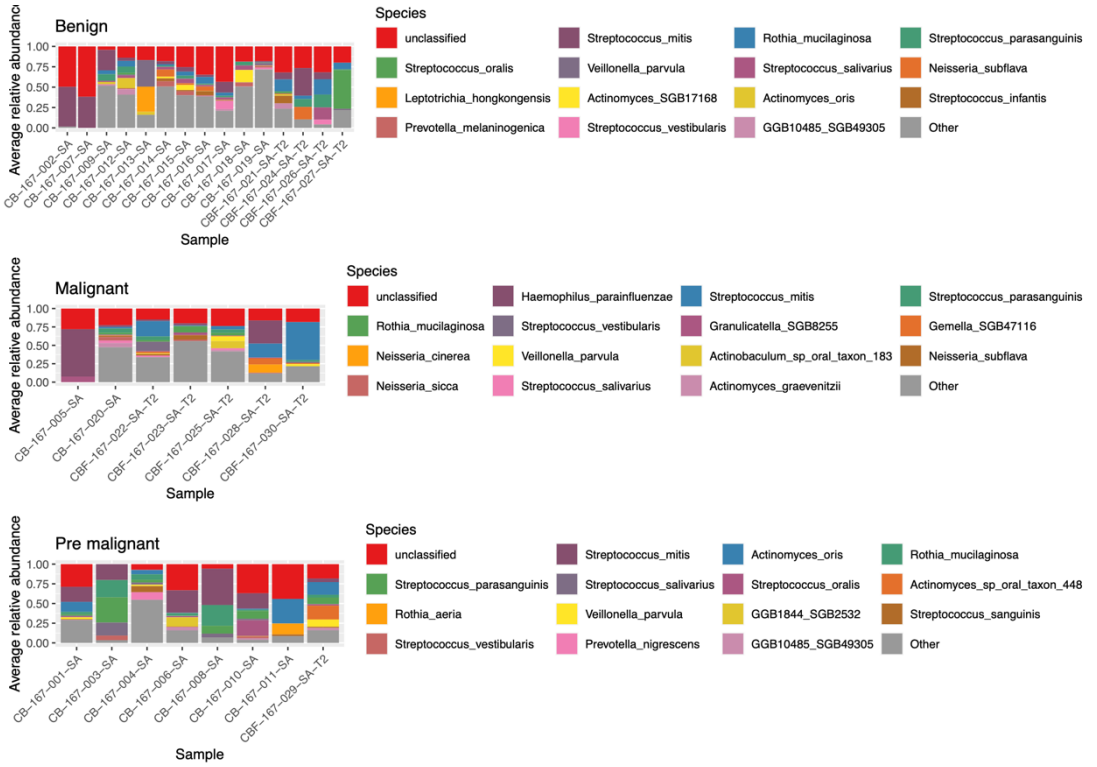



*Figure 22 Top 15 strains in Saliva organised for each histological group – the rest of strains were grouped as "Other"*

To get a deeper understanding of the taxonomic variations in each group, we examined the top 15 taxa present in each of the histological groups while grouping the rest of the taxa as "Other" (Figure 22). We observed the presence of various species associated with a healthy oral microbiome, including *Streptococcus mitis, S. parasanguinis, S. salivarius* and *Rothia mucilaginosa. Haemophilus parainfluenzae* was only observed in the malignant group and not in the top 15 taxa in the potentially malignant and benign groups. To further investigate the taxonomic profiles, we looked at the top 5 taxa per individual in each histological group for the saliva samples (Figure 23, Figure 24 and Figure 25). There was a lot of inter-individual variation observed for each taxa amongst the three histological groups. Certain species including *streptococcus mitis, oralis, infatis* were seen as the top 5 taxa in nearly all participants in the benign group. These species were also seen as the top 5 taxa in the potentially malignant and malignant groups but within fewer individuals. *Haemophilus parainfluenzae* and *Neisseara spp*. were present in some but not all in the malignant group.

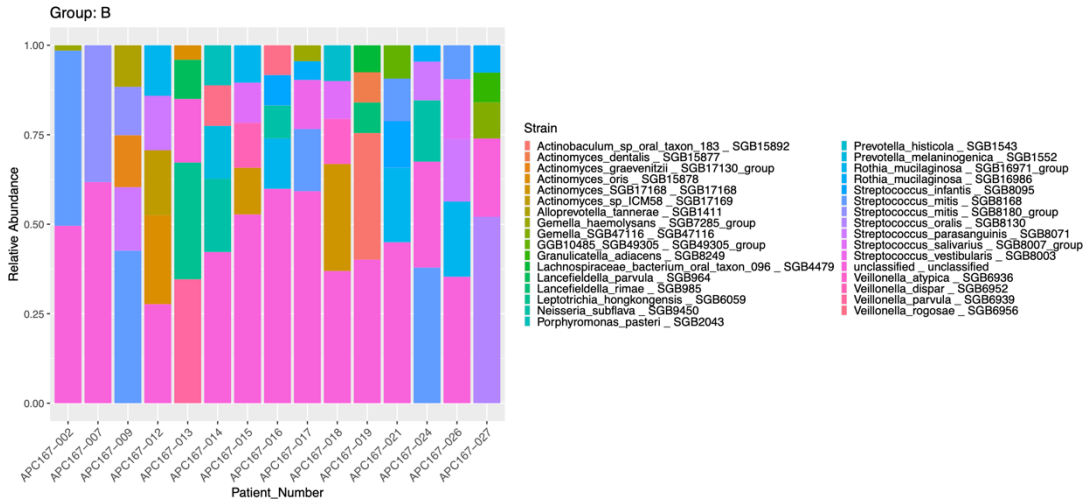



*Figure 23 Top 5 strains in Saliva –Benign group*

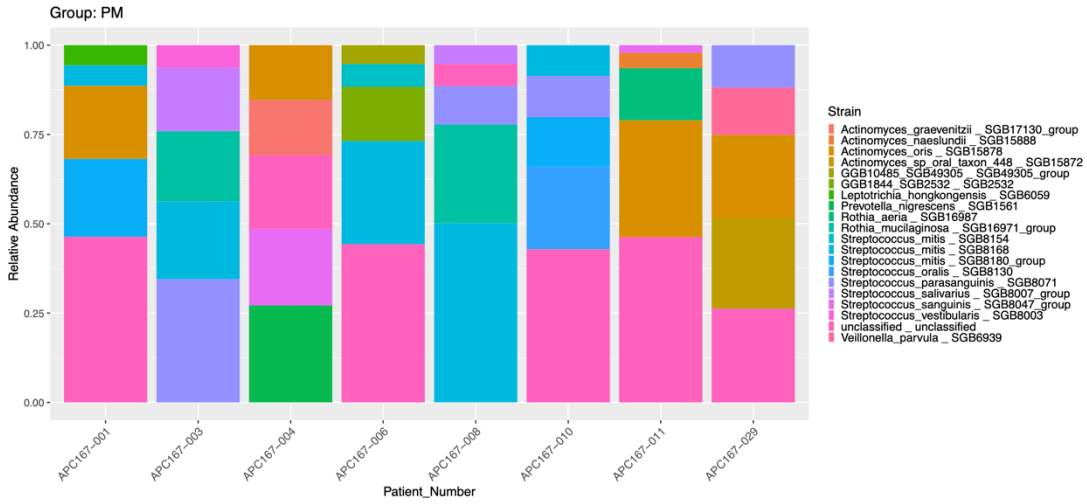



*Figure 24 Top 5 strains in Saliva – Potentially Malignant group*

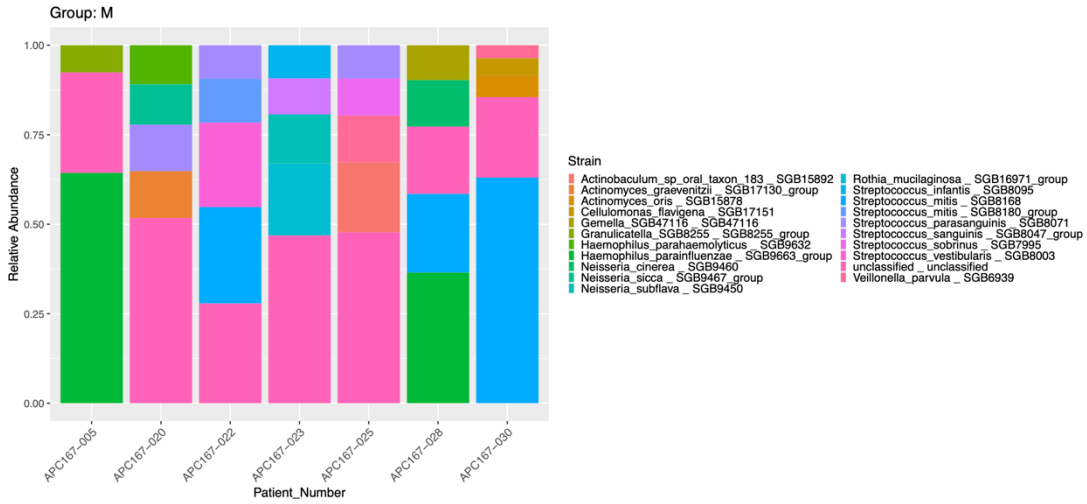



*Figure 25 Top 5 strains in Saliva –Malignant group*

### 4.5.4 Differential abundance analysis

Relative associations of genera to each group were estimated using songbird (Morton et al., 2019; "Songbird," 2024). This tool produces differentials which describe the log-fold change of microbes, with respect to a sample group, in this case the benign histological group. These differentials can then be ranked from lowest to highest giving information on the relative associations of microbes between the histological groups. Songbird differentials obtained using the formula C (Group, Treatment('Benign')) were sorted by ranks and the top 15 positive and negative features are plotted and depicted using bar charts. B (Benign) group is reference (negative here is more associated with reference group—Benign here) and (**a**) and (**b**) are top 15 positive and negative associations with treatment group malignant and potential malignant respectively – with stool and saliva samples used separately.

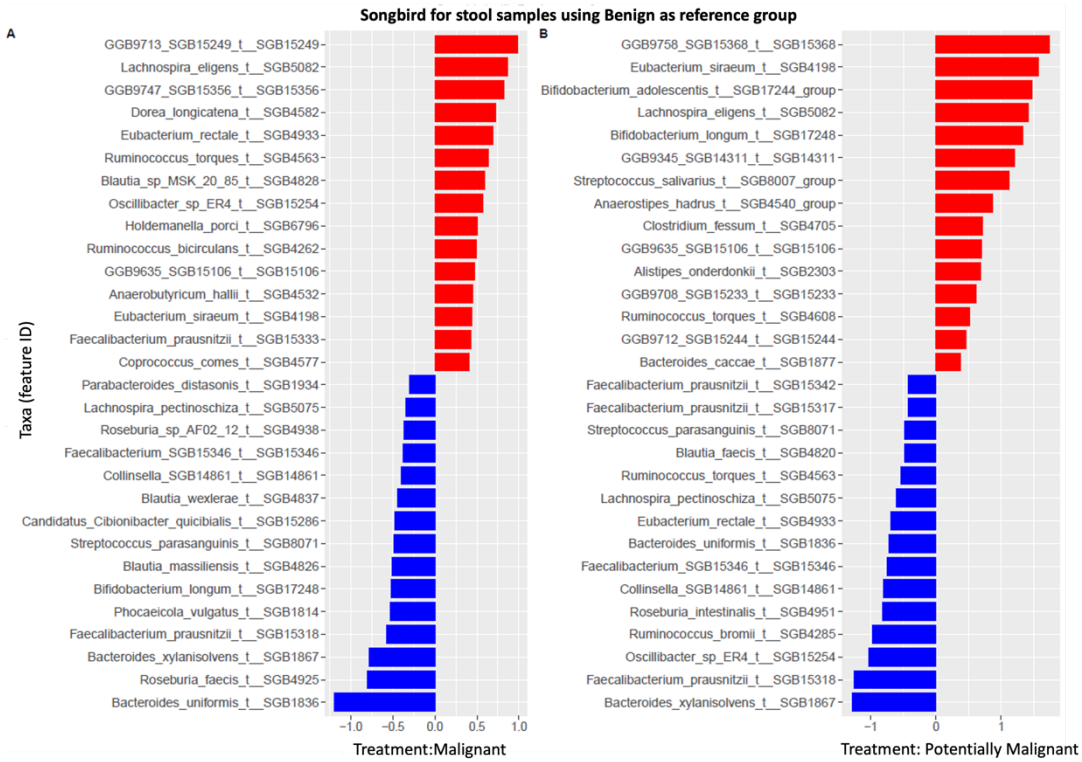



*Figure 26 Songbird differentials for Stool samples – Obtained using the formula C(Group, Treatment(Benign)) were sorted by ranks and the top 15 positive and negative features are plotted and depicted using bar charts with the benign group as reference (Negative is more associated to reference group – Benign here); A (Malignant), B (Potentially Malignant).*

We observed that for stool samples, when the benign group was compared with the malignant group, *Bacteroides spp*. and strain dependant *Faecalibacterium* were strongly associated with Benign group, while certain *Ruminococcus spp*., Eubacterium were associated to the Malignant and Potentially Malignant group (Figure 26).

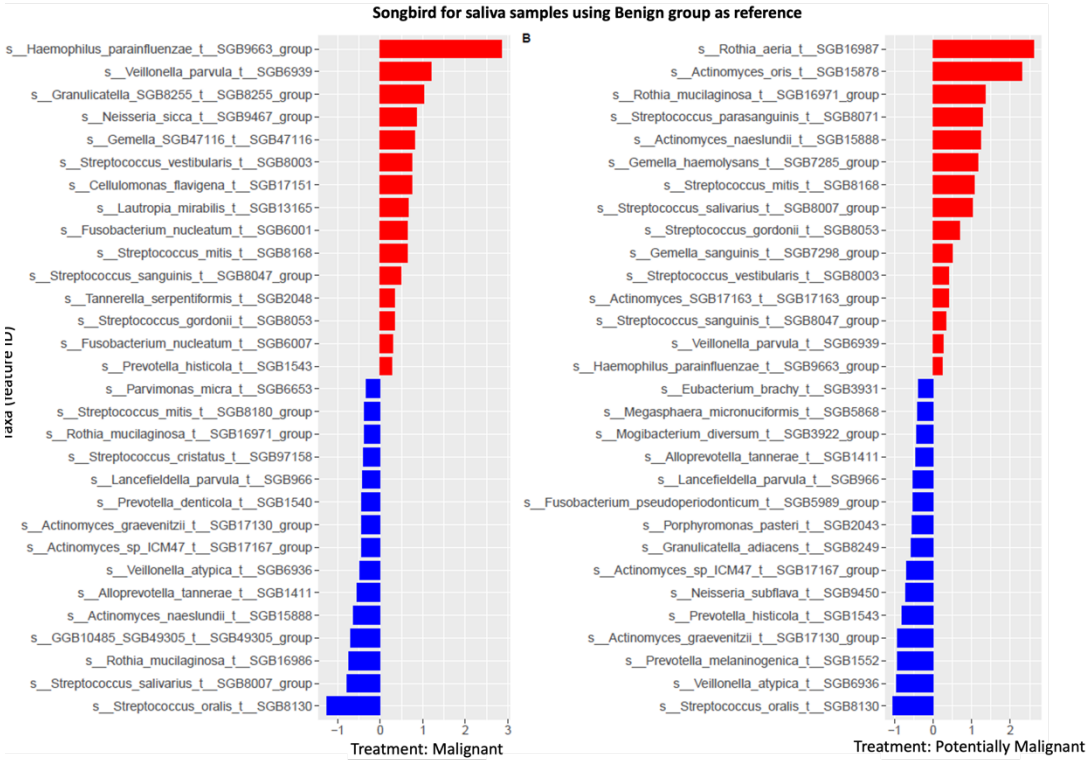



*Figure 27 Songbird differentials for Saliva samples – Obtained using the formula C(Group, Treatment(Benign)) were sorted by ranks and the top 15 positive and negative features are plotted and depicted using bar charts with the benign group as reference (Negative is more associated to reference group – Benign here); A (Malignant), B (Potentially Malignant).*

In the case of saliva samples, when the benign group was compared with the malignant group, Bacteroides spp. and strain dependant *Faecalibacterium* were strongly associated with Benign group, while certain *Ruminococcus spp*., Eubacterium were associated to the Malignant and Potentially Malignant group. *Streptococcus oralis, viellonella atypica* and *actinomyces spp.* in a strain dependant manner were more associated to the benign group in both cases. am *Haemophilus Parainfluenzae* was strongly associated with Malignant group and showed a positive though weaker association with the potentially malignant group. Similarly *Viellonella parvula*, *Streptococcus sanguinis*, *Streptococcus vestibularis* showed strong association to both malignant and potentially malignant groups. *Rothia spp. mucilaginosa*, *spp. aeria* showed strongest association to potentially malignant group and the benign group when compared to malignant group (Figure 27).

### 4.5.5 Functional profile analysis

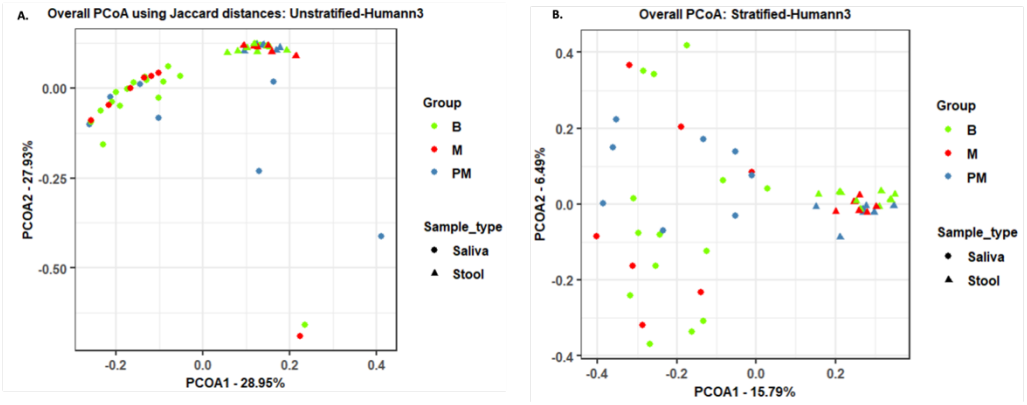



*Figure 28 - Beta diversity calculated with Jaccard distances visualised using PCoA showing functional profiling differences between histological groups and sample types; (A) unstratified pathway abundance data, (B) stratified pathway abundance data*

We carried out functional profiling analysis using Humann3 to look for differences in functionality between the histological groups. Humann3 output data was renormalized to copies per million by excluding unmapped and unintegrated pathways and the results visualized using Heatmap and PCoA in R. Using both unstratified and stratified pathway abundance data from Humann3 there was a profound effect on the functional profile of the microbiota depending on if the samples were from saliva or stool However there was no significant differences in functional profiling seen between the three histological groups for either stool or saliva samples (p=0.001) (Figure 28).

Further functional profiling analysis was done subdivided by sample type between histological groups and no statistically significant differences were seen (figure not shown).

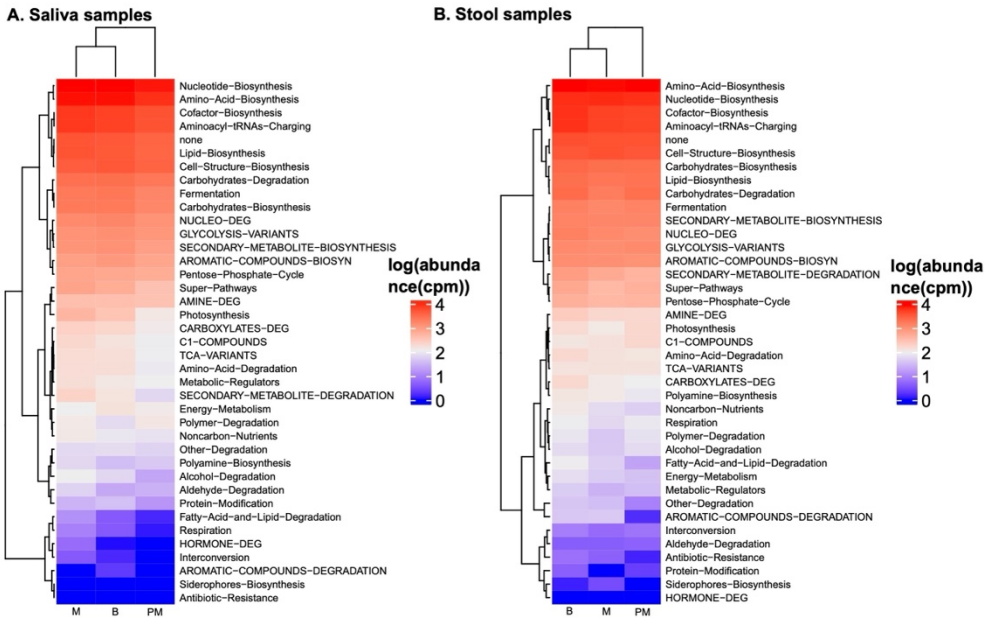



*Figure 29 - Heatmap showing functional profiling differences between histological groups (A) saliva samples, (B) stool samples*

Unstratified pathway abundance data was clustered into higher order metaCyc pathways for stool and saliva samples and visualised using a heatmap. No significance was observed for any higher order functional pathway between any two histological groups for saliva and stool samples. (Figure 29).

# Chapter 5: Discussion and Conclusions

Shotgun metagenomic sequencing methods were used in this study to determine if the development of oral malignancy could be related to the gut and oral microbiome. In addition, we set out to investigate if the gut and oral microbiomes could be a biomarker for malignant transformation of potentially malignant conditions. In order to achieve this, we observed the gut and oral microbiota of patients with benign, potentially malignant and malignant oral lesions to determine differences in alpha diversity, beta diversity, functional profiling and any distinct microbiome profiles. Further to this we also attempted to establish if there was a relationship between diet, BMI, smoking, alcohol intake and oral health on the gut and oral microbiota.

In our sample of 30 participants the ages ranged between 24 years and 78 years (mean age 56.37 years), with 12 (40%) female (F) and 18 (60%) male (M). The mean BMI of 27.13 kg/$m^2$ was slightly overweight.

Three (10%) participants had Good oral health, 8 (26.67%) had Moderate oral health, 13 (43.33%) had Poor oral health, and 4 (13.33%) had Very Poor oral health. Two (6.67%) participants had no teeth and could not have an oral health category assigned.

Antibiotics had been taken by 4 (13.3%) participants in the preceding two months. There were 13 (43.33%) non-smokers, 10 (33.33%) were former smokers, and 7 (23.33%) were current smokers. With regards alcohol consumption, 11 participants reported that they did not consume any alcohol. The remaining 19 participants who did report to consume alcohol, reported a range between 1 and 40 units, with the mean units consumed being 12.79 units. While it is not yet established how these background demographics impacts the microbiome they should be considered when comparing results across microbiome studies.

We observed differences in the alpha diversity of stool and saliva samples, though no significant differences in the composition of the oral microbiomes between histological groups was seen (Figure 11 and Figure 17). This highlights the challenge in the identification of an oral biomarker to distinguish oral lesions into their

histological groups. Nevertheless, significant differences were seen in the beta diversity of the gut microbiomes between the benign and malignant groups (Figure 16).

When comparing beta diversity between stool and saliva samples, we observed difference in microbial composition between the sample types. This is expected and is well documented in the literature that niche microbiomes residing in different habitats have a unique composition (Segata et al., 2012). Interestingly, there are some overlaps in microbial composition between these two environments, indicating shared taxa among these interconnected anatomical communities. These findings align with prior research, as demonstrated in a meta-analysis by (Valles-Colomer et al., 2023).

When comparing beta diversity between stool samples, we observed distinct clustering between the benign and malignant groups. While there was no significant difference in beta diversity between benign and potentially malignant groups, malignant transformation rate for OPMDs varies from 5% to 50%. Further research on potentially malignant lesions, accounting for the wide variation in malignant transformation rates is suggested. This could allow better understanding and detection of malignancy potential through measurable microbiome changes.

Oral hygiene is known to be an independent risk factor for oral cancer and also can affect the oral microbiome (Rosenquist et al., 2005). In our analysis we also observed that apart from sample type and histological groups, oral health explained the highest variance for the difference in microbial composition in both saliva and stool samples (Table 6 and Table 7), though this was not statistically significant. We found that taxa from *Veillonela* group were more associated in potentially malignant and malignant groups in a species dependant manner when observing the top 15 strains of microbes present. We also did not find a statistically significant difference in the alpha diversity of participants based on smoking status. However, it is seen in other studies as well that, even despite the lack of statistically significant differences there were trends towards an increased between-sample diversity (Kumar et al., 2011).

Amongst our three histological groups we found firmicutes to be in highest abundance in the benign group as has been demonstrated in previous studies (Z. Li et al., 2021). Our findings showed the highest abundance of proteobacteria was in the malignant group. While this is not seen in (Z. Li et al., 2021), it is corroborated in other studies (Radaic et al., 2023). The microbial abundance of leukoplakia (potentially malignant) was more similar to that of our study with Streptococcus species (*S. miitis, S. parasanguinis, S. salivarius*) present in our study and in (Herreros-Pomares et al., 2023). However with regards to proliferative verrucous leukoplakia (potentially malignant but with one of the highest malignant transformation rates), proteobacteria are seen in relatively low abundance in, which may indicate different processes at play (Herreros-Pomares et al., 2023).

In this study, with regards to saliva samples, *Haemophilus parainfluenzae* was strongly associated in the malignant group and to a lesser extent in the potentially malignant group as shown in the differential abundance analysis in Figure 27. This is confirmed in other studies (Chattopadhyay et al., 2019) where it has been suggested that *Haemophilus parainfluenzae* along with other microbes such as *Lactobacillus gasseri, Lactobacillus johnsonii*, *Lactobacillus fermentum*, and *Fusobacterium periodonticum* could be considered to be a biomarker for oral cancer. These other microbes were not seen in this study this could perhaps be due to the differences in sampling methods. In this study we collected oral saliva from the floor of the mouth using a sterile swab, while (Chattopadhyay et al., 2019) used and oral rinse and also included oropharyngeal cancer in addition to oral cavity cancer. It is known that the oral microbiome varies amongst the different oral cavity sites (Sami et al., 2023) and also depending on the sampling method (Roca et al., 2024). Additionally there was a lot of inter-individual variation in most histological groups. *Haemophilus parainfluenzae* and *Neisseara spp*. were present in some but not all in the malignant group. This could possibly be attributed to the stage and type of malignancy.

In our study *Veillonella parvula* was found to be differentially abundant in both malignant and potentially malignant groups, with a stronger association to the malignant group though the presence of this species may have different effects depending on the site and species (Baraniya et al., 2020; Chang et al., 2023; Li et al., 2022; Tsay et al., 2021). In some studies the presence of both *Haemophilus*

*parainfluenzae* and *Veillonella parvula* is down regulated though this is only seen in the latest stage, stage 4, of oral cancer (Yang et al., 2018). However our study did not categorise what stage of cancer the participants of the Malignant group were.

In our study *Rothia mucilaginosa* demonstrated differences in differential abundance when comparing the malignant to the benign group and when comparing the potentially malignant to the benign group. It showed a negative associating to the malignant but a positive association to the potentially malignant group. *Rothia mucilaginosa* has been seen in several studies to be present in potentially malignant conditions and predominantly appears in cases of oral leukoplakia (Amer et al., 2020, 2017).

While there is a great potential for the usefulness of biomarkers in the early identification and risk assessment of oral cancer, there are also challenges present in establishing and qualifying distinct microbiome markers in the oral and gut microbiomes. Currently in clinical practice, no qualified biomarkers are used despite their clinical value. We found distinct species associated with the three histological groups however it is difficult to establish a causal effect at this stage. The functional mechanisms of these microbiota could be further investigated to determine how the oral lesions develop into malignancy. A further complication of the inter-individual microbiome variation are confounding factors such as gender, alcohol, medications, and oral health. As seen in our study oral health explained the high variance in microbiome compositions. Oral health has been seen to be a risk factor in the several system diseases including malignancy (Maki et al., 2021).

## 5.1 Strengths, limitations and future research

Some of the strengths of this study include our analysis methods and the use of a reference control group. We used whole metagenomic sequencing (shotgun) instead of 16S amplicon sequencing (16S). While amplicon sequencing is relatively quick and inexpensive it can only yield taxonomy up to genera level. Shotgun metagenomic sequencing provides complete genomic information and higher taxonomic resolution than 16S sequencing. It can also generate multi-dimensional functional

profiles for understanding the specific functions of the microbiome. Shotgun sequencing also provides accurate taxonomy of established and previously known microbes as well as draft genomes of previously unreferenced microbes (Qian et al., 2020).

The use of three histological groups, with benign lesions acting as a reference control group allowed us to compare the microbiome changes in potentially malignant and malignant lesions more accurately. The use of a larger control group is required due to significant inter-individual variation in the oral and gut microbiomes. Results from studies that only look at the microbiome in potentially malignant and malignant lesions, can be difficult to compare as methodology differences are known to impact microbiome analysis.

To determine the sample size required in a study a power calculation is routinely used. Since this is an observational study and there is no hypothesis being tested so a strict sample size is not necessary to prove "sufficient effect". The number of samples required for a microbiome study depends on the effect size. However, because there is no standardised way of reporting effect sizes and similar studies may not exist, a pilot study may be necessary to define the effect size.

Obtaining participants and data for the malignant group presented some challenges. This group had the fewest samples, as thankfully the incidence or oral cancer is rare. In addition to this the presentation of oral cancer patients is urgently managed and recruitment into a research study is often not a clinical priority. The psychological impact and stress of a potential cancer diagnosis can affect participants ability to consent and willingness to provide data for research. In this study at least one participant (022) verbally reported being worried about providing a stool sample due to anxiety related constipation. This patient was informed at visit 1 that their definitive histopathology would most likely indicate malignancy. Five participants did not provide stool samples in this study and this was also the least collected sample.

With regards to the potentially malignant group, it is known that the wide variation in malignant transformation rates depends on the specific OPMD being considered. This study did not further classify the potentially malignant group by OPMDs as there were not enough participants to account for the wide subset variations that are to

be expected. A larger sample group or more targeted inclusion criteria could be used to aid in the analysis of specific microbiomes for each distinct OPMD being considered.

## 5.2 Conclusion

In this study we showed that distinct gut microbial profiles exist in benign and malignant oral lesions and that there is an association of specific microbes in oral saliva with malignant lesions. *Haemophilus parainfluenzae* and *Veillonella parvula* were seen to be strongly associated with malignancy. We also showed that while not statistically significant, oral hygiene was the leading factor in the variance of oral and gut microbiomes between individuals. Our study suggests that oral and gut microbiomes could act as possible biomarkers and aid in early detection and assessment of oral cancer risk. Further research is required to develop definitive biomarkers for malignant oral lesions. With regards to potentially malignant lesions future research could study individual OPMDs as distinct entities due to the wide variation in clinical and histological presentation, as well as malignant potential, and associated microbiome variations.

# References


1. Abdulkareem, A.A., Shelton, R.M., Landini, G., Cooper, P.R., Milward, M.R., 2018. Periodontal pathogens promote epithelial-mesenchymal transition in oral squamous carcinoma cells in vitro. Cell Adhes. Migr. 12, 127–137. https://doi.org/10.1080/19336918.2017.1322253
2. Amato, A., 2023. Periodontitis and Cancer: Beyond the Boundaries of Oral Cavity. Cancers 15, 1736. https://doi.org/10.3390/cancers15061736
3. Amer, A., Galvin, S., Healy, C.M., Moran, G.P., 2017. The Microbiome of Potentially Malignant Oral Leukoplakia Exhibits Enrichment for Fusobacterium, Leptotrichia, Campylobacter, and Rothia Species. Front. Microbiol. 8, 2391. https://doi.org/10.3389/fmicb.2017.02391
4. Amer, A., Whelan, A., Al-Hebshi, N.N., Healy, C.M., Moran, G.P., 2020. Acetaldehyde production by Rothia mucilaginosa isolates from patients with oral leukoplakia. J. Oral Microbiol. 12, 1743066. https://doi.org/10.1080/20002297.2020.1743066
5. Bäckhed, F., Ding, H., Wang, T., Hooper, L.V., Koh, G.Y., Nagy, A., Semenkovich, C.F., Gordon, J.I., 2004. The gut microbiota as an environmental factor that regulates fat storage. Proc. Natl. Acad. Sci. 101, 15718–15723. https://doi.org/10.1073/pnas.0407076101
6. Bagan, J., Sarrion, G., Jimenez, Y., 2010. Oral cancer: clinical features. Oral Oncol. 46, 414–417. https://doi.org/10.1016/j.oraloncology.2010.03.009
7. Bai, H., Yang, J., Meng, S., Liu, C., 2022. Oral Microbiota-Driven Cell Migration in Carcinogenesis and Metastasis. Front. Cell. Infect. Microbiol. 12. https://doi.org/10.3389/fcimb.2022.864479
8. Baker, J.L., Edlund, A., 2019. Exploiting the Oral Microbiome to Prevent Tooth Decay: Has Evolution Already Provided the Best Tools? Front. Microbiol. 9. https://doi.org/10.3389/fmicb.2018.03323
9. Baquero, F., Nombela, C., 2012. The microbiome as a human organ. Clin. Microbiol. Infect. Off. Publ. Eur. Soc. Clin. Microbiol. Infect. Dis. 18 Suppl 4, 2–4. https://doi.org/10.1111/j.1469-0691.2012.03916.x

10. Baraniya, D., Jain, V., Lucarelli, R., Tam, V., Vanderveer, L., Puri, S., Yang, M., Al-hebshi, N.N., 2020. Screening of Health-Associated Oral Bacteria for Anticancer Properties in vitro. Front. Cell. Infect. Microbiol. 10. https://doi.org/10.3389/fcimb.2020.575656
11. Biocore, 2023. Evident Tool [WWW Document]. URL https://github.com/biocore/evident (accessed 3.26.24).
12. Blanco-Míguez, A., Beghini, F., Cumbo, F., McIver, L.J., Thompson, K.N., Zolfo, M., Manghi, P., Dubois, L., Huang, K.D., Thomas, A.M., Nickols, W.A., Piccinno, G., Piperni, E., Punčochář, M., Valles-Colomer, M., Tett, A., Giordano, F., Davies, R., Wolf, J., Berry, S.E., Spector, T.D., Franzosa, E.A., Pasolli, E., Asnicar, F., Huttenhower, C., Segata, N., 2023. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. Nat. Biotechnol. 41, 1633–1644. https://doi.org/10.1038/s41587-023-01688-w
13. Bourgeois, D., Gonçalves, L.S., Lima-Junior, J. da C., Carrouel, F., 2022. Editorial: The Oral Microbiome Is a Key Factor in Oral and Systemic Health. Front. Microbiol. 13, 855668. https://doi.org/10.3389/fmicb.2022.855668
14. Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R.L., Soerjomataram, I., Jemal, A., 2024. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA. Cancer J. Clin. https://doi.org/10.3322/caac.21834
15. Califf, R.M., 2018. Biomarker definitions and their applications. Exp. Biol. Med. 243, 213–221. https://doi.org/10.1177/1535370217750088
16. Cani, P.D., Amar, J., Iglesias, M.A., Poggi, M., Knauf, C., Bastelica, D., Neyrinck, A.M., Fava, F., Tuohy, K.M., Chabo, C., Waget, A., Delmée, E., Cousin, B., Sulpice, T., Chamontin, B., Ferrières, J., Tanti, J.-F., Gibson, G.R., Casteilla, L., Delzenne, N.M., Alessi, M.C., Burcelin, R., 2007a. Metabolic endotoxemia initiates obesity and insulin resistance. Diabetes 56, 1761–1772. https://doi.org/10.2337/db06-1491
17. Cani, P.D., Neyrinck, A.M., Fava, F., Knauf, C., Burcelin, R.G., Tuohy, K.M., Gibson, G.R., Delzenne, N.M., 2007b. Selective increases of bifidobacteria in gut microflora improve high-fat-diet-induced diabetes in mice through a

mechanism associated with endotoxaemia. Diabetologia 50, 2374–2383. https://doi.org/10.1007/s00125-007-0791-0

18. Chang, X., Chen, Y., Cui, D., Wang, Y., Zhou, Y., Zhang, X., Tang, G., 2023. Propionate-producing Veillonella parvula regulates the malignant properties of tumor cells of OSCC. Med. Oncol. Northwood Lond. Engl. 40, 98. https://doi.org/10.1007/s12032-023-01962-6
19. Chattopadhyay, I., Verma, M., Panda, M., 2019. Role of Oral Microbiome Signatures in Diagnosis and Prognosis of Oral Cancer. Technol. Cancer Res. Treat. 18, 1533033819867354. https://doi.org/10.1177/1533033819867354
20. Chen, C., Hemme, C., Beleno, J., Shi, Z.J., Ning, D., Qin, Y., Tu, Q., Jorgensen, M., He, Z., Wu, L., Zhou, J., 2018. Oral microbiota of periodontal health and disease and their changes after nonsurgical periodontal therapy. ISME J. 12, 1210–1224. https://doi.org/10.1038/s41396-017-0037-1
21. Chen, J., Siliceo, S.L., Ni, Y., Nielsen, H.B., Xu, A., Panagiotou, G., 2023. Identification of robust and generalizable biomarkers for microbiome-based stratification in lifestyle interventions. Microbiome 11, 178. https://doi.org/10.1186/s40168-023-01604-z
22. Collins, H.L., Drazul-Schrader, D., Sulpizio, A.C., Koster, P.D., Williamson, Y., Adelman, S.J., Owen, K., Sanli, T., Bellamine, A., 2016. L-Carnitine intake and high trimethylamine N-oxide plasma levels correlate with low aortic lesions in ApoE(-/-) transgenic mice expressing CETP. Atherosclerosis 244, 29–37. https://doi.org/10.1016/j.atherosclerosis.2015.10.108
23. Colonia-García, A., Gutiérrez-Vélez, M., Duque-Duque, A., de Andrade, C.R., 2020. Possible association of periodontal disease with oral cancer and oral potentially malignant disorders: a systematic review. Acta Odontol. Scand. 78, 553–559. https://doi.org/10.1080/00016357.2020.1774076
24. Cryan, J.F., O'Riordan, K.J., Cowan, C.S.M., Sandhu, K.V., Bastiaanssen, T.F.S., Boehme, M., Codagnone, M.G., Cussotto, S., Fulling, C., Golubeva, A.V., Guzzetta, K.E., Jaggar, M., Long-Smith, C.M., Lyte, J.M., Martin, J.A., Molinero-Perez, A., Moloney, G., Morelli, E., Morillas, E., O'Connor, R., Cruz-Pereira, J.S., Peterson, V.L., Rea, K., Ritz, N.L., Sherwin, E., Spichak, S., Teichman, E.M., Van De Wouw, M., Ventura-Silva, A.P., Wallace-Fitzsimons, S.E., Hyland, N.,

Clarke, G., Dinan, T.G., 2019. The Microbiota-Gut-Brain Axis. Physiol. Rev. 99, 1877–2013. https://doi.org/10.1152/physrev.00018.2018

25. Cui, X., Ye, L., Li, J., Jin, L., Wang, W., Li, S., Bao, M., Wu, S., Li, L., Geng, B., Zhou, X., Zhang, J., Cai, J., 2018. Metagenomic and metabolomic analyses unveil dysbiosis of gut microbiota in chronic heart failure patients. Sci. Rep. 8, 635. https://doi.org/10.1038/s41598-017-18756-2
26. Dewhirst, F.E., Chen, T., Izard, J., Paster, B.J., Tanner, A.C.R., Yu, W.-H., Lakshmanan, A., Wade, W.G., 2010. The Human Oral Microbiome. J. Bacteriol. 192, 5002–5017. https://doi.org/10.1128/JB.00542-10
27. Di Stefano, M., Polizzi, A., Santonocito, S., Romano, A., Lombardi, T., Isola, G., 2022. Impact of Oral Microbiome in Periodontal Health and Periodontitis: A Critical Review on Prevention and Treatment. Int. J. Mol. Sci. 23, 5142. https://doi.org/10.3390/ijms23095142
28. Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S.R., Nelson, K.E., Relman, D.A., 2005. Diversity of the Human Intestinal Microbial Flora. Science 308, 1635–1638. https://doi.org/10.1126/science.1110591
29. European Medicines Agency, 2024. Biomarker [WWW Document]. URL https://www.ema.europa.eu/en/glossary/biomarker (accessed 4.14.24).
30. Everard, A., Lazarevic, V., Derrien, M., Girard, M., Muccioli, G.G., Neyrinck, A.M., Possemiers, S., Van Holle, A., François, P., de Vos, W.M., Delzenne, N.M., Schrenzel, J., Cani, P.D., 2011. Responses of gut microbiota and glucose and lipid metabolism to prebiotics in genetic obese and diet-induced leptin-resistant mice. Diabetes 60, 2775–2786. https://doi.org/10.2337/db11-0227
31. Fan, Y., Pedersen, O., 2021. Gut microbiota in human metabolic health and disease. Nat. Rev. Microbiol. 19, 55–71. https://doi.org/10.1038/s41579-020-0433-9
32. Farah, C.S., Woo, S., Zain, R.B., Sklavounou, A., McCullough, M.J., Lingen, M., 2014. Oral Cancer and Oral Potentially Malignant Disorders. Int. J. Dent. 2014, e853479. https://doi.org/10.1155/2014/853479
33. Flemer, B., Warren, R.D., Barrett, M.P., Cisek, K., Das, A., Jeffery, I.B., Hurley, E., O'Riordain, M., Shanahan, F., O'Toole, P.W., 2018. The oral microbiota in

colorectal cancer is distinctive and predictive. Gut 67, 1454–1463. https://doi.org/10.1136/gutjnl-2017-314814

34. Flint, H.J., Scott, K.P., Louis, P., Duncan, S.H., 2012. The role of the gut microbiota in nutrition and health. Nat. Rev. Gastroenterol. Hepatol. 9, 577–589. https://doi.org/10.1038/nrgastro.2012.156
35. Gaiser, R.A., Halimi, A., Alkharaan, H., Lu, L., Davanian, H., Healy, K., Hugerth, L.W., Ateeb, Z., Valente, R., Fernández Moro, C., Del Chiaro, M., Sällberg Chen, M., 2019. Enrichment of oral microbiota in early cystic precursors to invasive pancreatic cancer. Gut 68, 2186–2194. https://doi.org/10.1136/gutjnl-2018-317458
36. Gaundal, L., Myhrstad, M.C.W., Rud, I., Gjøvaag, T., Byfuglien, M.G., Retterstøl, K., Holven, K.B., Ulven, S.M., Telle-Hansen, V.H., 2022. Gut microbiota is associated with dietary intake and metabolic markers in healthy individuals. Food Nutr. Res. 66. https://doi.org/10.29219/fnr.v66.8580
37. Goodrich, J.K., Di Rienzi, S.C., Poole, A.C., Koren, O., Walters, W.A., Caporaso, J.G., Knight, R., Ley, R.E., 2014. Conducting a microbiome study. Cell 158, 250–262. https://doi.org/10.1016/j.cell.2014.06.037
38. Gopinath, D., Kunnath Menon, R., K. Veettil, S., George Botelho, M., Johnson, N.W., 2020. Periodontal Diseases as Putative Risk Factors for Head and Neck Cancer: Systematic Review and Meta-Analysis. Cancers 12, 1893. https://doi.org/10.3390/cancers12071893
39. Gorzelak, M.A., Gill, S.K., Tasnim, N., Ahmadi-Vand, Z., Jay, M., Gibson, D.L., 2015. Methods for Improving Human Gut Microbiome Data by Reducing Variability through Sample Processing and Storage of Stool. PloS One 10, e0134802. https://doi.org/10.1371/journal.pone.0134802
40. Guerrero-Preston, R., Godoy-Vitorino, F., Jedlicka, A., Rodríguez-Hilario, A., González, H., Bondy, J., Lawson, F., Folawiyo, O., Michailidi, C., Dziedzic, A., Thangavel, R., Hadar, T., Noordhuis, M.G., Westra, W., Koch, W., Sidransky, D., 2016. 16S rRNA amplicon sequencing identifies microbiota associated with oral cancer, human papilloma virus infection and surgical treatment. Oncotarget 7, 51320–51334. https://doi.org/10.18632/oncotarget.9710

41. Han, M., Hao, L., Lin, Y., Li, F., Wang, J., Yang, H., Xiao, L., Kristiansen, K., Jia, H., Li, J., 2018. A novel affordable reagent for room temperature storage and transport of fecal samples for metagenomic analyses. Microbiome 6, 43. https://doi.org/10.1186/s40168-018-0429-0
42. Herreros-Pomares, A., Hervás, D., Bagan-Debón, L., Jantus-Lewintre, E., Gimeno-Cardona, C., Bagan, J., 2023. On the Oral Microbiome of Oral Potentially Malignant and Malignant Disorders: Dysbiosis, Loss of Diversity, and Pathogens Enrichment. Int. J. Mol. Sci. 24, 3466. https://doi.org/10.3390/ijms24043466
43. Hou, K., Wu, Z.-X., Chen, X.-Y., Wang, J.-Q., Zhang, D., Xiao, C., Zhu, D., Koya, J.B., Wei, L., Li, J., Chen, Z.-S., 2022. Microbiota in health and diseases. Signal Transduct. Target. Ther. 7, 1–28. https://doi.org/10.1038/s41392-022-00974-4
44. Human Microbiome Project Consortium, 2012. Structure, function and diversity of the healthy human microbiome. Nature 486, 207–214. https://doi.org/10.1038/nature11234
45. Illumina, I., 2021. Illumina DNA Prep Reference Guide [WWW Document]. URL https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/illumina_prep/illumina-dna-prep-reference-guide-1000000025416-10.pdf (accessed 2.5.24).
46. Ishikawa, T., Terashima, J., Shimoyama, Y., Ohashi, Y., Mikami, T., Takeda, Y., Sasaki, M., 2020. Effects of butyric acid, a bacterial metabolite, on the migration of ameloblastoma mediated by laminin 332. J. Oral Sci. 62, 435–438. https://doi.org/10.2334/josnusd.19-0380
47. Kamada, N., Seo, S.-U., Chen, G.Y., Núñez, G., 2013. Role of the gut microbiota in immunity and inflammatory disease. Nat. Rev. Immunol. 13, 321–335. https://doi.org/10.1038/nri3430
48. Karpiński, T.M., 2019. Role of Oral Microbiota in Cancer Development. Microorganisms 7, 20. https://doi.org/10.3390/microorganisms7010020
49. KneadData, 2024.

50. Koay, Y.C., Chen, Y.-C., Wali, J.A., Luk, A.W.S., Li, M., Doma, H., Reimark, R., Zaldivia, M.T.K., Habtom, H.T., Franks, A.E., Fusco-Allison, G., Yang, J., Holmes, A., Simpson, S.J., Peter, K., O'Sullivan, J.F., 2021. Plasma levels of trimethylamine-N-oxide can be increased with "healthy" and "unhealthy" diets and do not correlate with the extent of atherosclerosis but with plaque instability. Cardiovasc. Res. 117, 435–449. https://doi.org/10.1093/cvr/cvaa094
51. Koren, O., Goodrich, J.K., Cullender, T.C., Spor, A., Laitinen, K., Bäckhed, H.K., Gonzalez, A., Werner, J.J., Angenent, L.T., Knight, R., Bäckhed, F., Isolauri, E., Salminen, S., Ley, R.E., 2012. Host remodeling of the gut microbiome and metabolic changes during pregnancy. Cell 150, 470–480. https://doi.org/10.1016/j.cell.2012.07.008
52. Kozak, M., Pawlik, A., 2023. The Role of the Oral Microbiome in the Development of Diseases. Int. J. Mol. Sci. 24, 5231. https://doi.org/10.3390/ijms24065231
53. Kumar, P.S., Matthews, C.R., Joshi, V., de Jager, M., Aspiras, M., 2011. Tobacco Smoking Affects Bacterial Acquisition and Colonization in Oral Biofilms ▿ . Infect. Immun. 79, 4730–4738. https://doi.org/10.1128/IAI.05371-11
54. Lauber, C.L., Zhou, N., Gordon, J.I., Knight, R., Fierer, N., 2010. Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples. FEMS Microbiol. Lett. 307, 80–86. https://doi.org/10.1111/j.1574-6968.2010.01965.x
55. Lederberg, J., McCray, A.T., 2001. Ome SweetOmics–A genealogical treasury of words. The scientist 15, 8–8.
56. Li, R.-J., Jie, Z.-Y., Feng, Q., Fang, R.-L., Li, F., Gao, Y., Xia, H.-H., Zhong, H.-Z., Tong, B., Madsen, L., Zhang, J.-H., Liu, C.-L., Xu, Z.-G., Wang, J., Yang, H.-M., Xu, X., Hou, Y., Brix, S., Kristiansen, K., Yu, X.-L., Jia, H.-J., He, K.-L., 2021. Network of Interactions Between Gut Microbiome, Host Biomarkers, and Urine Metabolome in Carotid Atherosclerosis. Front. Cell. Infect. Microbiol. 11, 708088. https://doi.org/10.3389/fcimb.2021.708088

57. Li, S., He, M., Lei, Y., Liu, Y., Li, X., Xiang, X., Wu, Q., Wang, Q., 2022. Oral Microbiota and Tumor—A New Perspective of Tumor Pathogenesis. Microorganisms 10, 2206. https://doi.org/10.3390/microorganisms10112206
58. Li, Z., Chen, G., Wang, P., Sun, M., Zhao, J., Li, A., Sun, Q., 2021. Alterations of the Oral Microbiota Profiles in Chinese Patient With Oral Cancer. Front. Cell. Infect. Microbiol. 11. https://doi.org/10.3389/fcimb.2021.780067
59. Lingen, M.W., Kalmar, J.R., Karrison, T., Speight, P.M., 2008. Critical Evaluation of Diagnostic Aids for the Detection of Oral Cancer. Oral Oncol. 44, 10–22. https://doi.org/10.1016/j.oraloncology.2007.06.011
60. Lloréns-Rico, V., Vieira-Silva, S., Gonçalves, P.J., Falony, G., Raes, J., 2021. Benchmarking microbiome transformations favors experimental quantitative approaches to address compositionality and sampling depth biases. Nat. Commun. 12, 3562. https://doi.org/10.1038/s41467-021-23821-6
61. Lukovac, S., Belzer, C., Pellis, L., Keijser, B.J., de Vos, W.M., Montijn, R.C., Roeselers, G., 2014. Differential modulation by Akkermansia muciniphila and Faecalibacterium prausnitzii of host peripheral lipid metabolism and histone acetylation in mouse gut organoids. mBio 5, e01438-14. https://doi.org/10.1128/mBio.01438-14
62. Maki, K.A., Kazmi, N., Barb, J.J., Ames, N., 2021. The Oral and Gut Bacterial Microbiomes: Similarities, Differences, and Connections. Biol. Res. Nurs. 23, 7–20. https://doi.org/10.1177/1099800420941606
63. Martel, C. de, Georges, D., Bray, F., Ferlay, J., Clifford, G.M., 2020. Global burden of cancer attributable to infections in 2018: a worldwide incidence analysis. Lancet Glob. Health 8, e180–e190. https://doi.org/10.1016/S2214-109X(19)30488-7
64. Mello, F.W., Miguel, A.F.P., Dutra, K.L., Porporatti, A.L., Warnakulasuriya, S., Guerra, E.N.S., Rivero, E.R.C., 2018. Prevalence of oral potentially malignant disorders: A systematic review and meta-analysis. J. Oral Pathol. Med. Off. Publ. Int. Assoc. Oral Pathol. Am. Acad. Oral Pathol. 47, 633–640. https://doi.org/10.1111/jop.12726

65. MetaHIT Consortium, Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., Mende, D.R., Li, J., Xu, J., Li, Shaochuan, Li, D., Cao, J., Wang, B., Liang, H., Zheng, H., Xie, Y., Tap, J., Lepage, P., Bertalan, M., Batto, J.-M., Hansen, T., Le Paslier, D., Linneberg, A., Nielsen, H.B., Pelletier, E., Renault, P., Sicheritz-Ponten, T., Turner, K., Zhu, H., Yu, C., Li, Shengting, Jian, M., Zhou, Y., Li, Y., Zhang, X., Li, Songgang, Qin, N., Yang, H., Wang, Jian, Brunak, S., Doré, J., Guarner, F., Kristiansen, K., Pedersen, O., Parkhill, J., Weissenbach, J., Bork, P., Ehrlich, S.D., Wang, Jun, 2010. A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464, 59–65. https://doi.org/10.1038/nature08821
66. Moloney, R.D., Desbonnet, L., Clarke, G., Dinan, T.G., Cryan, J.F., 2014. The microbiome: stress, health and disease. Mamm. Genome Off. J. Int. Mamm. Genome Soc. 25, 49–74. https://doi.org/10.1007/s00335-013-9488-5
67. Morton, J.T., Marotz, C., Washburne, A., Silverman, J., Zaramela, L.S., Edlund, A., Zengler, K., Knight, R., 2019. Establishing microbial composition measurement standards with reference frames. Nat. Commun. 10, 2719. https://doi.org/10.1038/s41467-019-10656-5
68. Nagy, K.N., Sonkodi, I., Szöke, I., Nagy, E., Newman, H.N., 1998. The microflora associated with human oral carcinomas. Oral Oncol. 34, 304–308. https://doi.org/10.1016/S1368-8375(98)80012-2
69. Naseribafrouei, A., Hestad, K., Avershina, E., Sekelja, M., Linløkken, A., Wilson, R., Rudi, K., 2014. Correlation between the human fecal microbiota and depression. Neurogastroenterol. Motil. 26, 1155–1162. https://doi.org/10.1111/nmo.12378
70. Nearing, J.T., DeClercq, V., Van Limbergen, J., Langille, M.G.I., 2020. Assessing the Variation within the Oral Microbiome of Healthy Adults. mSphere 5, 10.1128/msphere.00451-20. https://doi.org/10.1128/msphere.00451-20
71. Nel Van Zyl, K., Matukane, S.R., Hamman, B.L., Whitelaw, A.C., Newton-Foot, M., 2022. Effect of antibiotics on the human microbiome: a systematic review. Int. J. Antimicrob. Agents 59, 106502. https://doi.org/10.1016/j.ijantimicag.2021.106502

72. NIH HMP Working Group, Peterson, J., Garges, S., Giovanni, M., McInnes, P., Wang, L., Schloss, J.A., Bonazzi, V., McEwen, J.E., Wetterstrand, K.A., Deal, C., Baker, C.C., Di Francesco, V., Howcroft, T.K., Karp, R.W., Lunsford, R.D., Wellington, C.R., Belachew, T., Wright, M., Giblin, C., David, H., Mills, M., Salomon, R., Mullins, C., Akolkar, B., Begg, L., Davis, C., Grandison, L., Humble, M., Khalsa, J., Little, A.R., Peavy, H., Pontzer, C., Portnoy, M., Sayre, M.H., Starke-Reed, P., Zakhari, S., Read, J., Watson, B., Guyer, M., 2009. The NIH Human Microbiome Project. Genome Res. 19, 2317–2323. https://doi.org/10.1101/gr.096651.109

73. Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A.V., Mikheenko, A., Vollger, M.R., Altemose, N., Uralsky, L., Gershman, A., Aganezov, S., Hoyt, S.J., Diekhans, M., Logsdon, G.A., Alonge, M., Antonarakis, S.E., Borchers, M., Bouffard, G.G., Brooks, S.Y., Caldas, G.V., Chen, N.-C., Cheng, H., Chin, C.-S., Chow, W., de Lima, L.G., Dishuck, P.C., Durbin, R., Dvorkina, T., Fiddes, I.T., Formenti, G., Fulton, R.S., Fungtammasan, A., Garrison, E., Grady, P.G.S., Graves-Lindsay, T.A., Hall, I.M., Hansen, N.F., Hartley, G.A., Haukness, M., Howe, K., Hunkapiller, M.W., Jain, C., Jain, M., Jarvis, E.D., Kerpedjiev, P., Kirsche, M., Kolmogorov, M., Korlach, J., Kremitzki, M., Li, H., Maduro, V.V., Marschall, T., McCartney, A.M., McDaniel, J., Miller, D.E., Mullikin, J.C., Myers, E.W., Olson, N.D., Paten, B., Peluso, P., Pevzner, P.A., Porubsky, D., Potapova, T., Rogaev, E.I., Rosenfeld, J.A., Salzberg, S.L., Schneider, V.A., Sedlazeck, F.J., Shafin, K., Shew, C.J., Shumate, A., Sims, Y., Smit, A.F.A., Soto, D.C., Sović, I., Storer, J.M., Streets, A., Sullivan, B.A., Thibaud-Nissen, F., Torrance, J., Wagner, J., Walenz, B.P., Wenger, A., Wood, J.M.D., Xiao, C., Yan, S.M., Young, A.C., Zarate, S., Surti, U., McCoy, R.C., Dennis, M.Y., Alexandrov, I.A., Gerton, J.L., O'Neill, R.J., Timp, W., Zook, J.M., Schatz, M.C., Eichler, E.E., Miga, K.H., Phillippy, A.M., 2022. The complete sequence of a human genome. Science 376, 44–53. https://doi.org/10.1126/science.abj6987

74. O'Hara, A.M., Shanahan, F., 2006. The gut flora as a forgotten organ. EMBO Rep. 7, 688–693. https://doi.org/10.1038/sj.embor.7400731

75. Olsen, I., Yamazaki, K., 2019. Can oral bacteria affect the microbiome of the gut? J. Oral Microbiol. 11, 1586422. https://doi.org/10.1080/20002297.2019.1586422
76. O'Mahony, L., McCarthy, J., Kelly, P., Hurley, G., Luo, F., Chen, K., O'Sullivan, G.C., Kiely, B., Collins, J.K., Shanahan, F., Quigley, E.M.M., 2005. Lactobacillus and bifidobacterium in irritable bowel syndrome: symptom responses and relationship to cytokine profiles. Gastroenterology 128, 541–551. https://doi.org/10.1053/j.gastro.2004.11.050
77. Paster, B.J., Olsen, I., Aas, J.A., Dewhirst, F.E., 2006. The breadth of bacterial diversity in the human periodontal pocket and other oral sites. Periodontol. 2000 42, 80–87. https://doi.org/10.1111/j.1600-0757.2006.00174.x
78. Patterson, E., Ryan, P.M., Cryan, J.F., Dinan, T.G., Ross, R.P., Fitzgerald, G.F., Stanton, C., 2016. Gut microbiota, obesity and diabetes. Postgrad. Med. J. 92, 286–300. https://doi.org/10.1136/postgradmedj-2015-133285
79. Polk, D.B., Peek, R.M., 2010. Helicobacter pylori: gastric cancer and beyond. Nat. Rev. Cancer 10, 403–414. https://doi.org/10.1038/nrc2857
80. Pushalkar, S., Ji, X., Li, Y., Estilo, C., Yegnanarayana, R., Singh, B., Li, X., Saxena, D., 2012. Comparison of oral microbiota in tumor and non-tumor tissues of patients with oral squamous cell carcinoma. BMC Microbiol. 12, 144. https://doi.org/10.1186/1471-2180-12-144
81. Qian, X.-B., Chen, T., Xu, Y.-P., Chen, L., Sun, F.-X., Lu, M.-P., Liu, Y.-X., 2020. A guide to human microbiome research: study design, sample collection, and bioinformatics analysis. Chin. Med. J. (Engl.) 133, 1844–1855. https://doi.org/10.1097/CM9.0000000000000871
82. Radaic, A., Shamir, E.R., Jones, K., Villa, A., Garud, N.R., Tward, A.D., Kamarajan, P., Kapila, Y.L., 2023. Specific Oral Microbial Differences in Proteobacteria and Bacteroidetes Are Associated with Distinct Sites When Moving from Healthy Mucosa to Oral Dysplasia—A Microbiome and Gene Profiling Study and Focused Review. Microorganisms 11, 2250. https://doi.org/10.3390/microorganisms11092250
83. Rhee, S.H., Pothoulakis, C., Mayer, E.A., 2009. Principles and clinical implications of the brain–gut–enteric microbiota axis. Nat. Rev.

Gastroenterol. Hepatol. 6, 10.1038/nrgastro.2009.35. https://doi.org/10.1038/nrgastro.2009.35

84. Rivera, C., 2015. Essentials of oral cancer. Int. J. Clin. Exp. Pathol. 8, 11884–11894.
85. Roca, C., Alkhateeb, A.A., Deanhardt, B.K., Macdonald, J.K., Chi, D.L., Wang, J.R., Wolfgang, M.C., 2024. Saliva sampling method influences oral microbiome composition and taxa distribution associated with oral diseases. PloS One 19, e0301016. https://doi.org/10.1371/journal.pone.0301016
86. Rosenquist, K., Wennerberg, J., Schildt, E.-B., Bladström, A., Göran Hansson, B., Andersson, G., 2005. Oral status, oral infections and some lifestyle factors as risk factors for oral and oropharyngeal squamous cell carcinoma. A population-based case-control study in southern Sweden. Acta Otolaryngol. (Stockh.) 125, 1327–1336. https://doi.org/10.1080/00016480510012273
87. Sami, A., Elimairi, I., Ryan, C.A., Stanton, C., Patangia, D., Ross, R.P., 2023. Altered oral microbiome in Sudanese Toombak smokeless tobacco users carries a newly emerging risk of squamous cell carcinoma development and progression. Sci. Rep. 13, 6645. https://doi.org/10.1038/s41598-023-32892-y
88. Scharlau, D., Borowicki, A., Habermann, N., Hofmann, T., Klenow, S., Miene, C., Munjal, U., Stein, K., Glei, M., 2009. Mechanisms of primary cancer prevention by butyrate and other products formed during gut flora-mediated fermentation of dietary fibre. Mutat. Res. Mutat. Res. 682, 39–53. https://doi.org/10.1016/j.mrrev.2009.04.001
89. Schmidt, B.L., Kuczynski, J., Bhattacharya, A., Huey, B., Corby, P.M., Queiroz, E.L.S., Nightingale, K., Kerr, A.R., DeLacure, M.D., Veeramachaneni, R., Olshen, A.B., Albertson, D.G., 2014. Changes in Abundance of Oral Microbiota Associated with Oral Cancer. PLoS ONE 9, e98741. https://doi.org/10.1371/journal.pone.0098741
90. Schroeder, B.O., Bäckhed, F., 2016. Signals from the gut microbiota to distant organs in physiology and disease. Nat. Med. 22, 1079–1089. https://doi.org/10.1038/nm.4185

91. Sedghi, L., DiMassa, V., Harrington, A., Lynch, S.V., Kapila, Y.L., 2021. The oral microbiome: Role of key organisms and complex networks in oral health and disease. Periodontol. 2000 87, 107–131. https://doi.org/10.1111/prd.12393
92. Segata, N., Haake, S.K., Mannon, P., Lemon, K.P., Waldron, L., Gevers, D., Huttenhower, C., Izard, J., 2012. Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples. Genome Biol. 13, R42. https://doi.org/10.1186/gb-2012-13-6-r42
93. Sergeant, M.J., Constantinidou, C., Cogan, T., Penn, C.W., Pallen, M.J., 2012. High-Throughput Sequencing of 16S rRNA Gene Amplicons: Effects of Extraction Procedure, Primer Length and Annealing Temperature. PLoS ONE 7, e38094. https://doi.org/10.1371/journal.pone.0038094
94. Songbird, 2024.
95. Speight, P.M., 2007. Update on Oral Epithelial Dysplasia and Progression to Cancer. Head Neck Pathol. 1, 61–66. https://doi.org/10.1007/s12105-007-0014-5
96. Stephens, R.W., Arhire, L., Covasa, M., 2018. Gut Microbiota: From Microorganisms to Metabolic Organ Influencing Obesity. Obesity 26, 801–809. https://doi.org/10.1002/oby.22179
97. Sudo, N., Chida, Y., Aiba, Y., Sonoda, J., Oyama, N., Yu, X.-N., Kubo, C., Koga, Y., 2004. Postnatal microbial colonization programs the hypothalamic–pituitary–adrenal system for stress response in mice. J. Physiol. 558, 263–275. https://doi.org/10.1113/jphysiol.2004.063388
98. Sun, J., Tang, Q., Yu, S., Xie, M., Xie, Y., Chen, G., Chen, L., 2020. Role of the oral microbiota in cancer evolution and progression. Cancer Med. 9, 6306–6321. https://doi.org/10.1002/cam4.3206
99. Tezal, M., Grossi, S.G., Genco, R.J., 2005. Is periodontitis associated with oral neoplasms? J. Periodontol. 76, 406–410. https://doi.org/10.1902/jop.2005.76.3.406
100. The British Society of Periodontology, 2019. Basic Periodontal Examination (BPE) [WWW Document]. URL

https://www.bsperio.org.uk/assets/downloads/BSP_BPE_Guidelines_2019.pdf (accessed 3.31.24).

101. Tsay, J.-C.J., Wu, B.G., Sulaiman, I., Gershner, K., Schluger, R., Li, Y., Yie, T.-A., Meyn, P., Olsen, E., Perez, L., Franca, B., Carpenito, J., Iizumi, T., El-Ashmawy, M., Badri, M., Morton, J.T., Shen, N., He, L., Michaud, G., Rafeq, S., Bessich, J.L., Smith, R.L., Sauthoff, H., Felner, K., Pillai, R., Zavitsanou, A.-M., Koralov, S.B., Mezzano, V., Loomis, C.A., Moreira, A.L., Moore, W., Tsirigos, A., Heguy, A., Rom, W.N., Sterman, D.H., Pass, H.I., Clemente, J.C., Li, H., Bonneau, R., Wong, K.-K., Papagiannakopoulos, T., Segal, L.N., 2021. Lower Airway Dysbiosis Affects Lung Cancer Progression. Cancer Discov. 11, 293–307. https://doi.org/10.1158/2159-8290.CD-20-0263

102. Turnbaugh, P.J., Ley, R.E., Mahowald, M.A., Magrini, V., Mardis, E.R., Gordon, J.I., 2006. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature 444, 1027–1031. https://doi.org/10.1038/nature05414

103. Valles-Colomer, M., Blanco-Míguez, A., Manghi, P., Asnicar, F., Dubois, L., Golzato, D., Armanini, F., Cumbo, F., Huang, K.D., Manara, S., Masetti, G., Pinto, F., Piperni, E., Punčochář, M., Ricci, L., Zolfo, M., Farrant, O., Goncalves, A., Selma-Royo, M., Binetti, A.G., Becerra, J.E., Han, B., Lusingu, J., Amuasi, J., Amoroso, L., Visconti, A., Steves, C.M., Falchi, M., Filosi, M., Tett, A., Last, A., Xu, Q., Qin, N., Qin, H., May, J., Eibach, D., Corrias, M.V., Ponzoni, M., Pasolli, E., Spector, T.D., Domenici, E., Collado, M.C., Segata, N., 2023. The person-to-person transmission landscape of the gut and oral microbiomes. Nature 614, 125–135. https://doi.org/10.1038/s41586-022-05620-1

104. Vergara, D., Simeone, P., Damato, M., Maffia, M., Lanuti, P., Trerotola, M., 2019. The Cancer Microbiota: EMT and Inflammation as Shared Molecular Mechanisms Associated with Plasticity and Progression. J. Oncol. 2019, 1253727. https://doi.org/10.1155/2019/1253727

105. Wang, H., Funchain, P., Bebek, G., Altemus, J., Zhang, H., Niazi, F., Peterson, C., Lee, W.T., Burkey, B.B., Eng, C., 2017. Microbiomic differences in tumor and paired-normal tissue in head and neck squamous cell

carcinomas. Genome Med. 9, 14. https://doi.org/10.1186/s13073-017-0405-5

106. Warnakulasuriya, S., Kujan, O., Aguirre-Urizar, J.M., Bagan, J.V., González-Moles, M.Á., Kerr, A.R., Lodi, G., Mello, F.W., Monteiro, L., Ogden, G.R., Sloan, P., Johnson, N.W., 2021. Oral potentially malignant disorders: A consensus report from an international seminar on nomenclature and classification, convened by the WHO Collaborating Centre for Oral Cancer. Oral Dis. 27, 1862–1880. https://doi.org/10.1111/odi.13704

107. World Health Organization, n.d. Health topics: Cancer [WWW Document]. URL https://www.who.int/health-topics/cancer (accessed 4.7.24).

108. Yang, C.-Y., Yeh, Y.-M., Yu, H.-Y., Chin, C.-Y., Hsu, C.-W., Liu, H., Huang, P.-J., Hu, S.-N., Liao, C.-T., Chang, K.-P., Chang, Y.-L., 2018. Oral Microbiota Community Dynamics Associated With Oral Squamous Cell Carcinoma Staging. Front. Microbiol. 9, 862. https://doi.org/10.3389/fmicb.2018.00862

109. Yu, Z., Morrison, M., 2004. Improved extraction of PCR-quality community DNA from digesta and fecal samples. BioTechniques 36, 808–812. https://doi.org/10.2144/04365ST04

110. Zhang, J.S., Chu, C.-H., Yu, O.Y., 2022. Oral Microbiome and Dental Caries Development. Dent. J. 10, 184. https://doi.org/10.3390/dj10100184

111. Zhang, Y., Chen, R., Zhang, D., Qi, S., Liu, Y., 2023. Metabolite interactions between host and microbiota during health and disease: Which feeds the other? Biomed. Pharmacother. 160, 114295. https://doi.org/10.1016/j.biopha.2023.114295

# Appendices

*Appendix A*

Ethical approval

*Appendix B*

Patient Information Leaflet

*Appendix C*

Data Protection Notice

*Appendix D*

Consent form

*Appendix E*

Food Frequency Questionnaire

*Appendix F*

Stool collection instructions

*Appendix G*

Case report form – patient demographics

*Appendix H*

Laboratory Steps

*Appendix I*

Case report form – results

# Appendix A

COISTE EITICE UM THAIGHDE CLINICIÚIL
**Clinical Research Ethics Committee of the Cork Teaching Hospitals**

Tel: +353-21-4901901
Email: crec@ucc.ie

University College Cork
Lancaster Hall
6 Little Hanover Street
Cork
Ireland

---

**CREC Review Reference Number:** ECM 4 (k) 28/06/2022

**Date:** 4th July 2022

Mr Paul Brady
Consultant Oral Surgeon
Cork University Dental School and Hospital
Wilton
Cork

**Study Title: APC167 - Benign, potentially malignant & malignant oral lesions; an analysis of oral and gut microbiota.**

Approval is granted to carry out the above study at:

| |
|---|
| Cork University Dental School and Hospital, Teagasc Food Research Centre and Cork University Hospital. |

The following documents have been approved:

| Document | Approved | Version | Date |
|---|---|---|---|
| Cover Letter | Yes | | 10th June 2022 |
| Application Form | Yes | | Signed 10th June 2022 |
| Evidence of Insurance | Yes | | 10th June 2022 |
| CV for Chief Investigator | Yes | | |
| Instructions of Collecting an Aerobic Stool Sample | Yes | 1.0 | 10th June 2022 |
| Study Protocol | Yes | 1 | 10th June 2022 |
| Participant Information Leaflet/Informed Consent Form | Yes | 1 | 10th June 2022 |
| Food Frequency Questionnaire Instructions | Yes | 1 | 10th June 2022 |
| Food Frequency Questionnaire | Yes | 1 | 10th June 2022. |

We note that the co-investigator(s) involved in this project will be:

| Name | Appointment Details |
|---|---|
| Dr Junaid Nayyar | Doctorate of Clinical Dentistry Postgraduate Student |
| Dr Eimear Hurley | Lecturer in Paediatrics |
| Professor Catherine Stanton | Senior Research Officer |
| Professor Paul Ross | Director & Principal Investigator |
| Dr Linda Feeley | Consultant Histopathologist |
| Dr Cassandre Bedu | Post Doctoral Researcher |
| Dr Amel Sami | PhD Student. |

The date of this letter is the date of authorization of the study.

Please keep a copy of this signed approval letter in your study master file for audit purposes. The study must be carried out in accordance with General Data Protection Regulation and Health Research Regulation 2018/2021.

You should note that ethical approval will lapse if you do not adhere to the following conditions:

1. Submission of an Annual Progress Report/Annual Renewal Survey (due annually from the date of this approval letter). **We would encourage you to keep note of this date as the CREC will not issue a reminder.**

2. Report unexpected adverse events, serious adverse events or any event that may affect ethical acceptability of the study

3. Submit any change to study documentation (minor or major) to CREC for review and approval. Amendments must be submitted on an amendment application form and revised study documents must

clearly highlight the changes and contain a new version number and date. Amendments cannot be implemented without written approval from CREC.

4. Notify CREC of discontinuation of the study

5. Submit an End of Trial Declaration Form and Final Study Report/Study Synopsis when the study has been completed.

Yours sincerely

Professor Brendan Buckley
Vice Chairman
Clinical Research Ethics Committee
of the Cork Teaching Hospitals

# APC167 - Annual renewal - Acknowledgement

MN
APC 11/7/2023

Carmel Burke <CarmelBurke@ucc.ie>
Tue 11/07/2023 12:37
To:Michelle Nugent <michelle.nugent@ucc.ie>
Cc:Leona Heaphy <l.heaphy@ucc.ie>

Dear Michelle

**Re: Study Title: APC167 – Benign, potentially malignant & malignant oral lesions; an analysis of oral and gut microbiota**

We acknowledge receipt of Cover letter and Annual Renewal Form signed 4th July 2023 for the above study.

The CREC will not send a hard copy response to this submission and therefore we would ask that you print this email for your study file.

Kind regards
Carmel

*Carmel Burke*
*Administrative Assistant*
*Clinical Research Ethics Committee*
*Of the Cork Teaching Hospitals*
*University College Cork*

# Appendix B

## PARTICIPANT INFORMATION LEAFLET

**Study Title:** Benign, potentially malignant & malignant oral lesions; an analysis of oral and gut microbiota

**Study number:** APC167

**Chief Investigator:** Mr Paul Brady

**Site of Investigation:** Cork University Dental School and Hospital

**Sites of Analysis:** Histopathology Department, Cork University Hospital.
APC Microbiome Ireland, University College Cork.
Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork.

You are being invited to take part in a research study. Before you decide whether or not you wish to take part, it is important for you to understand why this research is being done and what it will involve. You should read this information leaflet carefully and, if you wish, discuss it with your family or friends. Once you understand it fully, you will be asked to sign the consent form if you are happy to take part. This process is known as "informed consent". If anything is unclear, or if you would like more information, please feel free to ask us any questions – we will be happy to answer them.

You do not have to take part in this study, participation is voluntary. You can change your mind about taking part in the study any time you like. Even if the study has started, you can still opt out. You do not have to give us a reason for withdrawing.

**Why is this study being carried out?**

This study is being carried out to analyse the normal microorganisms (bacteria, fungi, viruses) that are present in the gut and in the mouth. We are trying to investigate if there is a link between the types of microorganisms that are present in the gut and the mouth

and if this link has an impact on conditions that affect the mouth such as oral cancer and other diseases.

The results of this study will provide more knowledge and data to this emerging area of research and add to the body of knowledge that already exists.

**Why have I been asked to take part in this study?**

You have been asked to take part in this study because you are attending for a biopsy of a lesion in your mouth. We aim to recruit up to 100 participants to this study.

**What does my participation involve?**

If you decide to take part in the study after reading this information leaflet, you will be asked to read and sign the consent form and will be given a signed copy to keep. Your eligibility to participate in the study will be decided using inclusion and exclusion criteria.

This study involves you attending as normal for your biopsy visits at Cork University Dental School & Hospital. This can take up to three visits; one consultation visit, one biopsy visit and one visit to deliver results. You will not be required to attend for anything other than your usual visits. The oral biopsy will be performed by your dental surgeon and your chart will be reviewed if needed for your biopsy results.

You will be asked to provide:

- your height and weight (BMI)
- your smoking and alcohol usage
- any medication and antibiotic use
- periodontal status
- your food and diet consumption via a questionnaire
- a fecal sample via a provided kit and instructions that you can bring to your next visit
- an oral biopsy as scheduled, part of your biopsy will be used for this research
- an oral saliva sample

You will be instructed to not brush your teeth or use mouthwash for at least 3 hours prior to collecting the saliva sample.

**Study Visit Schedule**

| | Visit 1 (screening visit) | Visit 2 | Visit 3 |
|---|---|---|---|
| **Informed Consent** | X | | |
| **Inclusion/Exclusion Criteria** | X | | |
| **Body Mass Index (BMI)** | X | | |
| **Medications/Antibiotics** | X | | |
| **Alcohol/Smoking** | X | | |
| **Food Frequency Questionnaire** | X | | |
| **Receive Stool Kit** | X | | |
| **Stool Sample Collection** | | X | or X |
| **Oral Biopsy** | | X | |
| **Oral Saliva** | | X | |
| **Biopsy Results** | | | X |

**Will I benefit from taking part in this study?**

There will be no personal benefit to you from the study, however your participation and the results of the study will hopefully benefit people/society in the future.

**Will I experience any risks or unpleasant side effects?**

There are no side effects with providing a stool sample or an oral saliva sample. The oral biopsy will be done as part of your standard of care and will be performed by the dental surgeon.

**What happens if I start the study and change my mind later?**

You do not have to take part in this study, participation is entirely voluntary. Refusal to participate, or discontinuing participation at any time, will not affect the health care you normally receive and will not involve any penalty or loss of benefits. You can withdraw from the study at any time by contacting a member of the research team listed below.

**Confidentiality**

All information will be stored securely and will be treated with the strictest confidence as per General Data Protection Regulation (GDPR). All the information gathered from this study will be stored on a computer, paper files and will be treated confidentially. You will be identified only by a unique identification code. In the event of any publication regarding this study, your identity will not be disclosed. Please see attached Data Protection Notice for more information on your rights and how your data will be treated. The investigator will go through this data protection notice and how relevant it is to you. This study will comply with the Clinical Research Ethics Committee guidelines as well as the General Data Protection Regulation and the Health Research Regulation.

**Will I be informed of the results of the study?**

The results of the study will be published via UCC's open access research repository – CORA and will be freely available to view.

**What happens if there is anything I do not understand?**

If there is anything you are not sure about, the Investigator(s) listed below will be happy to explain in more detail to yourself or your relatives (or legal representative if required). The study will be fully explained to you before you decide if you want to take part.

**Contact Information**

Dr Junaid Nayyar or Mr Paul Brady
Oral Surgery Department, Cork University Dental School & Hospital
Phone: 021 490 1127

# Appendix C

**CONSENT BY SUBJECT FOR PARTICIPATION IN A HUMAN OBSERVATIONAL STUDY**

**Protocol Number:** APC167

**Participant Name:** _____

**Study Title:** Benign, potentially malignant & malignant oral lesions; an analysis of oral and gut microbiota

**Chief Investigator:** Mr Paul Brady

**Participation in this study is voluntary and you may withdraw at any time for any reason**

This research study and sampling procedure associated with it have been fully explained to me. All experimental procedures have been identified and no guarantees have been given about the possible results. I have had the opportunity to ask questions concerning any and all aspects of the study and the sampling procedures involved. I am aware that participation is voluntary, and I may withdraw consent at any time. I am aware that my decision not to participate or to withdraw will not restrict my access to health care services normally available to me. Confidentiality of records concerning my involvement in this study will be maintained in an appropriate manner. When required by law, the records may be reviewed by government agencies/sponsors of the research/University College Cork and the Clinical Research Ethics Committee of the Cork Teaching Hospitals.

I understand that the Sponsors and investigators have such insurance as is required by law in the event of injury as a direct result of my participation in the study. I understand that if I have any questions concerning this research, I can contact the Investigator listed above. If I have further queries concerning my rights in connection with the research, I can contact the Clinical Research Ethics Committee of the Cork Teaching Hospitals, Lancaster Hall, 6 Little Hanover Street, Cork, who have approved this research study.

I, the undersigned, hereby consent to participate in the above described study conducted at Cork University Dental School & Hospital. I have received a copy of this consent form for my records.

Analyses of all samples and information collected will be conducted at:

- APC Microbiome Ireland, University College Cork
- Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork
- Histopathology Department, Cork University Hospital

On some occasions the analyses may be done in collaboration with third parties, including commercial partners, which may require samples to be shipped to these organisations. In all cases, samples and data will be coded with unique identification codes.

Access to samples and/or data will require approval from the Director and/or Executive Management Group of APC Microbiome Ireland, University College Cork.

I agree to allow the clinical researchers access to my medical chart to obtain information for the purpose of this research study.

I agree that results and microbial isolates may be used to design a commercial strategy or product for human health.

After reading the entire consent form, if you have no further questions about giving consent, please sign where indicated.

Participant's Signature: _____ Date: _____
dd mon yyyy

NAME (BLOCK LETTERS): _____ Time: _____

Investigator's Signature: _____ Date:_____
dd mon yyyy

NAME (BLOCK LETTERS): _____ Time: _____

# Appendix D

## UCC Clinical Studies Data Protection Notice

At University College Cork, we treat your privacy seriously. Any personal data which you provide to the University will be treated with the highest standards of security and confidentiality, in accordance with Irish and European Data Protection legislation. This notice sets out details of the information that we collect, how we process it and who we share it with. It also explains your rights under data protection law in relation to our processing of your data.

### Who we are

Throughout this Notice, "we", "us" and "our" refers to University College Cork, as study sponsor. For more information about us, please refer to our website:

### How we will use your personal data

By participating in this study, information about you (also called "personal data") will be accessed, collected and stored for the purposes mentioned in the Participant Information Leaflet relating to the study. This personal data includes

- Information that directly identifies you (such as your name and your year of birth);
- Your gender, ethnic and racial background;
- Information on your health and medical condition including your medical history;
- Your treatments and your response to treatments;
- Information contained in your samples and the results after analysis;
- Information from your medical records

Personal data collected at any time during the study will be kept strictly confidential. The data will be held securely at the hospital, referred to in this notice as "the clinical site". To ensure confidentiality, the data generated during the study is **coded** with a number that will identify you in the study. Any information that leaves the clinical site will be labelled with your code instead of your name. Every person that has access to your uncoded data at the clinical site is subject to professional secrecy and confidentiality.

Data that directly identifies you (uncoded data) is stored in your medical files at the clinical site in which the samples and data have been obtained. A list or 'key' linking your study number to your name will also be kept securely (in a locked cupboard in a room with restricted access) by the researcher.

### Who will access my personal data?

Your uncoded data will only be accessible to clinical site employees, the study researcher and site staff, quality representatives from the study Sponsor (Monitors and Auditors), and the Clinical Research Ethics Committee of the Cork Teaching Hospitals (CREC) so that they can check if the study is being conducted to the best standards.

Results of the study will be provided to the ethics committee approving the study in compliance with national and international regulations on clinical studies.

## Cross-border data transfers

*Data may be transferred outside of the European Economic Area on the understanding we rely on legally approved mechanisms to lawfully transfer data across borders including the Standard Contractual Clauses approved by the European Commission however no personal data will be transferred.*

## The purpose and legal basis for collecting your data

Any personal data you provide to us during the course of this study will be processed fairly and lawfully.

Signing the Informed Consent Form means that your personal data and biological samples will be used for the purposes outlined in the Participant information leaflet (PIL). You are entitled to withdraw your consent at any time.

Personal data collected during this study and the results of the study may be presented for scientific purposes. However, you will never be identified individually during these presentations. Your identity will not be revealed in any reports or publications.

The clinical site staff, the study researcher and the members of the study's team will use your personal data within the scope defined above. If the clinical site staff, study researcher or study team wish to use your data for a purpose other than the purpose specified, the researcher must contact you again to give you more information and ask your permission to use your data for the new purpose.

The General Data Protection Regulation (GDPR) allows us to process your data because the research is of substantial public interest (Articles 6(1) (e) and 9(2) (j) of the GDPR). If you require further information on the legal basis for processing your personal data, please contact UCC's Data Protection Officer – details below.

## How long we will keep your data

The personal data collected in the study will be kept for a period of 15 years as per UCC Code of Research Conduct after the end of the study. Thereafter, they may be stored for a further period of time for legal reasons (e.g. revised retention obligations), or more if required by law.

## Your rights

You have various rights under data protection law, subject to certain exemptions, in connection with our processing of your personal data, including the right:

- to find out if we use your personal data, access your personal data and receive copies of your personal data;
- to have inaccurate/incomplete information corrected and updated;
- in certain circumstances, to have your details deleted from systems that we use to process your personal data or have the use of your personal data restricted in certain ways;

- to object to certain processing of your data by UCC;
- to exercise your right to data portability where applicable (i.e. obtain a copy of your personal data in a commonly used electronic form);
- to withdraw your consent to the processing of your data at any time without giving a reason by notifying your decision to the study researcher. If you withdraw your consent for data processing, your participation in the study stops and no further data will be collected from you. Your study Researcher will present you the options you have concerning your personal data.
- If, for any reason, you stop attending the study visits, your study Researcher may decide to withdraw you from the study. If this happens, we will continue to hold your data for research purposes unless you inform us that you do not want us to hold your data any longer.
- Along with study withdrawal, you have the right to request the deletion of data about you if your data are no longer necessary for the purposes of processing or there is no other legal ground for their further processing.

If you wish to exercise any of these rights, please address your request to the study researcher or the Data Protection Officer, University College Cork (details below).

## Questions or Complaints

If you have any questions in relation to this study please contact the study Chief Investigator Mr Paul Brady. Email: paul.brady@ucc.ie or phone 021 490 1127

University College Cork (UCC) is the Data Controller for this research project.

If you have any complaints in connection with our processing of your personal data, you can contact UCC's Data Protection Officer (DPO): **DPO, Information Compliance Section, Office of Corporate & Legal Affairs, University College Cork, Western Road, Cork E: gdpr@ucc.ie**

You also have the right to lodge a complaint with the Data Protection Commission if you are unhappy with our processing of your personal data. Details of how to lodge a complaint can be found on the Data Protection Commission's website (www.dataprotection.ie), or by telephoning 1890 252 231.

**Participant Study Number:**

# Appendix E

| **Date of visit: Encounter:** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Meat, fish, poultry - Medium serving (size deck of cards)** | **Never** | **<1/M** | **1/M** | **1/W** | **2/W** | **5/W** | **1/D** | **2/D** | **4/D** | **6+/D** | **Comment** |
| Beef (roast / steak) | | | | | | | | | | | |
| Beef: stew | | | | | | | | | | | |
| Beef burger (1 burger) | | | | | | | | | | | |
| Pork (roast / chops / escalopes) | | | | | | | | | | | |
| Lamb (roast / chops / stew) | | | | | | | | | | | |
| Chicken or other poultry e.g. turkey: roast | | | | | | | | | | | |
| Breaded chicken, chicken nuggets, chicken burger | | | | | | | | | | | |
| Bacon / Ham | | | | | | | | | | | |
| Processed meat (Corned beef, Luncheon meats, sausages) | | | | | | | | | | | |
| Savoury pies (e.g. meat, pork, steak & kidney, sausage roll) | | | | | | | | | | | |
| Liver, heart, kidney, paté | | | | | | | | | | | |
| Fish fried (batter/breadcrumbs), baked, grilled, fingers/cakes | | | | | | | | | | | |
| White fish, fresh or frozen (e.g. cod, haddock, plaice, sole) | | | | | | | | | | | |
| Oily fish, fresh or canned (e.g. mackerel, kippers, tuna, salmon, sardines, herring) | | | | | | | | | | | |
| Shellfish (e.g. crab, prawns, mussels) | | | | | | | | | | | |
| **Bread and savoury biscuits - One slice or one biscuit** | **Never** | **<1/M** | **1/M** | **1/W** | **2/W** | **5/W** | **1/D** | **2/D** | **4/D** | **6+/D** | **Comment** |
| White bread and rolls (including ciabatta & pannini) | | | | | | | | | | | |
| Brown bread and rolls | | | | | | | | | | | |
| Wholemeal bread and rolls | | | | | | | | | | | |
| Wheat-free; Rye brad; spelt bread (specify) | | | | | | | | | | | |
| Cream crackers, cheese biscuits | | | | | | | | | | | |
| Crisp bread, e.g. Ryvita | | | | | | | | | | | |
| Pancakes, muffins, oatcakes | | | | | | | | | | | |
| Scone (white) | | | | | | | | | | | |
| Scone (brown) | | | | | | | | | | | |
| **Cereals - One medium sized bowl** | **Never** | **<1/M** | **1/M** | **1/W** | **2/W** | **5/W** | **1/D** | **2/D** | **4/D** | **6+/D** | **Comment** |
| Non-ready to eat - Porridge, Readybrek | | | | | | | | | | | |
| High fibre (All Bran/flakes, Weetabix, Shredded wheat) | | | | | | | | | | | |
| Cornflakes, Rice Krispies | | | | | | | | | | | |
| Muesli (e.g. Country Store, Aplen) | | | | | | | | | | | |
| Sugar-coated cereals (e.g. Frosties, Crunchy nut cornflakes) | | | | | | | | | | | |
| **Potatoes, rice and pasta - Medium serving = a cupful** | **Never** | **<1/M** | **1/M** | **1/W** | **2/W** | **5/W** | **1/D** | **2/D** | **4/D** | **6+/D** | **Comment** |
| Boiled, instant or jacket potatoes | | | | | | | | | | | |
| Mashed potatoes | | | | | | | | | | | |
| Chips / Roast potatoes | | | | | | | | | | | |
| Potato Salad | | | | | | | | | | | |
| White Rice | | | | | | | | | | | |
| Brown Rice | | | | | | | | | | | |
| White/yellow or green pastas (e.g. spagetti, noodles) | | | | | | | | | | | |
| Wholemeal pasta | | | | | | | | | | | |
| Lasagne (meat based) | | | | | | | | | | | |
| Lasagne (vegetarian) | | | | | | | | | | | |
| Pizza | | | | | | | | | | | |
| **Dairy products and fat** | **Never** | **<1/M** | **1/M** | **1/W** | **2/W** | **5/W** | **1/D** | **2/D** | **4/D** | **6+/D** | **Comment** |
| Cream (tablespoon) | | | | | | | | | | | |
| Full-fat yoghurt or Greek style yoghurt (125g carton) | | | | | | | | | | | |
| Dairy desserts (125g carton) | | | | | | | | | | | |
| Cheddar cheese (medium serving) | | | | | | | | | | | |
| Low-fat cheddar cheese (medium serving) | | | | | | | | | | | |
| Cottage cheese | | | | | | | | | | | |
| Eggs as boiled, fried, scrambled, poached (one) | | | | | | | | | | | |
| Quiche (medium serving) | | | | | | | | | | | |
| Light salad cream or light mayonnaise (tablespoon) | | | | | | | | | | | |
| Salad cream, mayonnaise (tablespoon) | | | | | | | | | | | |
| French dressing (tablespoon) | | | | | | | | | | | |
| Other salad dressing (tablespoon) | | | | | | | | | | | |
| **The following on bread or vegetables (teaspoon):** | **Never** | **<1/M** | **1/M** | **1/W** | **2/W** | **5/W** | **1/D** | **2/D** | **4/D** | **6+/D** | **Comment** |
| Butter | | | | | | | | | | | |
| Lite Butter e.g. Dawn Lite, Connacht Gold | | | | | | | | | | | |
| Sunflower margarine e.g. Flora | | | | | | | | | | | |
| Low-fat margarine (e.g. lowlow) | | | | | | | | | | | |
| Cholesterol Lowering e.g. Flora Pro-Active, DairyGold Heart | | | | | | | | | | | |
| Cream & Veg. Oil spread e.g. Kerrymaid, Dairy Gold | | | | | | | | | | | |
| Olive oil spread e.g. Golden Olive | | | | | | | | | | | |
| **Fruit - 1 fruit or a medium serving (a cupful)** | **Never** | **<1/M** | **1/M** | **1/W** | **2/W** | **5/W** | **1/D** | **2/D** | **4/D** | **6+/D** | **Comment** |
| Apples | | | | | | | | | | | |
| Pears | | | | | | | | | | | |
| Oranges, satsumas, mandarins | | | | | | | | | | | |
| Grapefruit | | | | | | | | | | | |
| Bananas | | | | | | | | | | | |
| Grapes | | | | | | | | | | | |
| Melon | | | | | | | | | | | |
| Peaches, plums | | | | | | | | | | | |
| Apricots | | | | | | | | | | | |
| Strawberries, raspberries, kiwi fruit | | | | | | | | | | | |
| Blueberries | | | | | | | | | | | |
| Tinned fruit (specify) | | | | | | | | | | | |
| Dried fruit e.g. raisins | | | | | | | | | | | |

**Participant Study Number:**

| **Date of visit: Encounter:** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Frozen fruit (specify) | | | | | | | | | | | |
| **Vegetable - A medium serving (a cupful)** | **Never** | **<1/M** | **1/M** | **1/W** | **2/W** | **5/W** | **1/D** | **2/D** | **4/D** | **6+/D** | **Comment** |
| Carrots | | | | | | | | | | | |
| Spinach | | | | | | | | | | | |
| Broccoli, spring greens, kale | | | | | | | | | | | |
| Brussel sprouts | | | | | | | | | | | |
| Cabbage | | | | | | | | | | | |
| Peas | | | | | | | | | | | |
| Green beans, broad beans, runner beans | | | | | | | | | | | |
| Courgettes | | | | | | | | | | | |
| Cauliflower | | | | | | | | | | | |
| Parsnips, turnips | | | | | | | | | | | |
| Leeks | | | | | | | | | | | |
| Onions | | | | | | | | | | | |
| Garlic | | | | | | | | | | | |
| Mushrooms | | | | | | | | | | | |
| Sweet peppers | | | | | | | | | | | |
| Green salad, lettuce | | | | | | | | | | | |
| Cucumber, celery | | | | | | | | | | | |
| Tomatoes | | | | | | | | | | | |
| Sweetcorn | | | | | | | | | | | |
| Beetroot | | | | | | | | | | | |
| Coleslaw | | | | | | | | | | | |
| Baked beans | | | | | | | | | | | |
| Dried lentils, beans, peas | | | | | | | | | | | |
| Tofu, soya meat, TVP, vegeburger | | | | | | | | | | | |
| **Sweets and snacks - medium serving** | **Never** | **<1/M** | **1/M** | **1/W** | **2/W** | **5/W** | **1/D** | **2/D** | **4/D** | **6+/D** | **Comment** |
| Chocolate coated sweet biscuits e.g. digestive (one) | | | | | | | | | | | |
| Plain biscuit e.g. marietta, digestives, rich tea (one) | | | | | | | | | | | |
| Cakes e.g. fruit, sponge | | | | | | | | | | | |
| Buns, pastries e.g. croissants, doughnuts | | | | | | | | | | | |
| Fruit pies, tarts, crumbles | | | | | | | | | | | |
| Sponge puddings | | | | | | | | | | | |
| Milk puddings e.g. rice, custard, trifle | | | | | | | | | | | |
| Ice cream, choc ices, frozen desserts | | | | | | | | | | | |
| Chocolates, singles or squares | | | | | | | | | | | |
| Sweets, toffees, mints | | | | | | | | | | | |
| Sugar added to tea coffee, cereals (teaspoons) | | | | | | | | | | | |
| Sugar substitute e.g. canderel (teaspoon) | | | | | | | | | | | |
| Crisps or other packet snacks | | | | | | | | | | | |
| Peanuts or other nuts | | | | | | | | | | | |
| **Soups, sauces and spreads** | **Never** | **<1/M** | **1/M** | **1/W** | **2/W** | **5/W** | **1/D** | **2/D** | **4/D** | **6+/D** | **Comment** |
| Vegetable soups: homemade/fresh (1 bowl) | | | | | | | | | | | |
| Vegetable soups: tinned/packet (1 bowl) | | | | | | | | | | | |
| Meat or cream soups: homemade / fresh (1 bowl) | | | | | | | | | | | |
| Meat or cream soups: tinned / packet (1 bowl) | | | | | | | | | | | |
| Sauces e.g. white, cheese, gravy (tablespoon) | | | | | | | | | | | |
| Tomato based sauces e.g. pasta sauces | | | | | | | | | | | |
| Curry-type sauces | | | | | | | | | | | |
| Pickles, chutney (tablespoon) | | | | | | | | | | | |
| Marmite, Bovril (tablespoon) | | | | | | | | | | | |
| Jam, marmalade, honey, syrup (teaspoon) | | | | | | | | | | | |
| Peanut butter (teaspoon) | | | | | | | | | | | |
| **Drinks** | **Never** | **<1/M** | **1/M** | **1/W** | **2/W** | **5/W** | **1/D** | **2/D** | **4/D** | **6+/D** | **Comment** |
| Tea (cup) | | | | | | | | | | | |
| Herbal tea (cup) | | | | | | | | | | | |
| Whole milk (cup) - cow, goat, soya, rice milk - specify | | | | | | | | | | | |
| Semi-skimmed milk (cup) | | | | | | | | | | | |
| Coffee instant (cup) | | | | | | | | | | | |
| Coffee ground (cup) | | | | | | | | | | | |
| Coffee, decaffeinated (cup) | | | | | | | | | | | |
| Coffee whitener e.g. Coffee-mate (teaspoon) | | | | | | | | | | | |
| Cocoa, Hot Chocolate (cup) | | | | | | | | | | | |
| Horlicks, Ovaltine (cup) | | | | | | | | | | | |
| Wine (glass) | | | | | | | | | | | |
| Beer, Larger or Cider (half pint) | | | | | | | | | | | |
| Low alcohol / alcohol free beer / larger (half pint) | | | | | | | | | | | |
| Port, Sherry, Vermouth, liqueurs (glass) | | | | | | | | | | | |
| Spirits e.g. Gin, Whiskey (Single measure) | | | | | | | | | | | |
| Low calorie or diet soft fizzy drink (glass) | | | | | | | | | | | |
| Fizzy Soft drinks e.g. (Coca Cola - Glass) | | | | | | | | | | | |
| Pure fruit drinks e.g. orange juice (small glass) | | | | | | | | | | | |
| Fruit squash (small glass) | | | | | | | | | | | |
| **Functional Foods / Supplements** | **Never** | **<1/M** | **1/M** | **1/W** | **2/W** | **5/W** | **1/D** | **2/D** | **4/D** | **6+/D** | **Comment** |
| Probiotic Yoghurts / cheese / milk | | | | | | | | | | | |
| Vitamin supplements (details) | | | | | | | | | | | |
| Mineral supplements (details) | | | | | | | | | | | |
| **Ready meals** | **Never** | **<1/M** | **1/M** | **1/W** | **2/W** | **5/W** | **1/D** | **2/D** | **4/D** | **6+/D** | **Comment** |
| Ready meal (specify) | | | | | | | | | | | |
| Takeaway (specify) | | | | | | | | | | | |

**167_001_ST_T2**

Stool sample should be as fresh as possible.

Ideally on the morning of or evening before you attend.

1. Place and rest the container in the toilet bowl.

   You can save the whole bowel motion or a part of it.

   Only a very small amount is used for analysis.

2. Use disposable gloves provided to secure lid FIRMLY.

   Place box in red bag first and then inside zip-lock bag second.

3. Store as cold as possible. Use provided cooler bag and icepack.

4. Bring with you to you next appointment.

**Appendix G**

# Case Report Form

APC167 Participant number:_____

| Patient Demographic Information | | | |
|---|---|---|---|
| Study ID Number | APC167 | | |
| Institution | Cork University Hospital | | |
| Inclusion/exclusion criteria<br>*Patient must meet all criteria to eligible for the study | Met all ☐1 | Not met* ☐2 | |
| Date of Enrolment & Consent | D D M M Y Y Y Y | | |
| Date of Birth | D D M M Y Y Y Y | | |
| Gender | ☐ Male | ☐ Female | |
| Smoking Status | ☐ Current | ☐ Never | ☐ Former |
| If positive smoking history, _____ cigarettes/day _____no of years | | | |
| C2H5OH | Yes _____ units/week | No | |
| Weight _____ (kg) Height _____(m) | BMI _____ (kg/m$^2$) | | |
| BPE Score | | | |

| | |
|---|---|
| Antibiotic in last 2 months? | ☐ Yes ☐ No |
| If yes, which antibiotic? | Antibiotic _____<br>Dose _____<br>Route _____<br>Duration _____ |
| Probiotic use? | Yes No |

# Case Report Form

APC167 Participant number:_____

## Current Medication

| Drug | Dose | Frequency |
|---|---|---|
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

## Comorbidities

| |
|---|
| |
| |
| |
| |
| |
| |
| |

## Checklist (amend as appropriate)

| Sample | Comments | Yes | No |
|---|---|---|---|
| Informed Consent | | | |
| Demographics | | | |
| Oral saliva 1 (SA_T1) | | | |
| Oral biopsy (TS_T1) | | | |
| Oral saliva 2 (SA_T2) | | | |
| Stool sample (ST_T2) | | | |
| Food Frequency Questionnaire (FFQ) | | | |

# Appendix H

**30/09/2022**
**Materials for samples collection transferred to Junaid**

- Saliva swab tubes
- Biohazard bags
- Zip lock bags
- Cooler zipper bags
- Ice packs

For biopsy samples: preparation of sterile 2 mL Eppendorf screw-cap tubes + 600 µL RNA later (n=56) (to be stored at RT)

**11/11/2022**
**Reception of samples**

Transported into ice packs by Junaid to Teagasc
Stored in -80°C hydrofish freezer (A.2)

List of samples: **cf. APC167_List_Samples_CBF.xlsx**
**This file is updated with the extraction day and collection day to follow the progression.**

**02/12/2022**
**Reception of samples**

Transported into ice packs by Junaid to Teagasc
Stored in -80°C hydrofish freezer (A.2)
Update of the list of samples: **cf. APC167_List_Samples_CBF.xlsx**

**03/02/2023**
**Reception of samples**

Transported into ice packs by Junaid to Teagasc
Stored in -80°C hydrofish freezer (A.2)
Update of the list of samples: **cf. APC167_List_Samples_CBF.xlsx**

**14/02/2023**
**Extraction of DNA**

Selection of 10 samples of saliva / swabs

| Participant Number | Sample Name | Reception Day | Sample Number |
|---|---|---|---|
| 167-001 | 167-001_SA_T1 | 11/11/2022 | 1 |
| 167-002 | 167-002_SA_T1 | 11/11/2022 | 2 |
| 167-003 | 167-003_SA_T1 | 11/11/2022 | 3 |
| 167-004 | 167-004_SA_T1 | 11/11/2022 | 4 |
| 167-005 | 167-005_SA_T1 | 11/11/2022 | 5 |
| 167-006 | 167-006_SA_T1 | 11/11/2022 | 6 |
| 167-007 | 167-007_SA_T1 | 11/11/2022 | 7 |
| 167-008 | 167-008_SA_T1 | 02/12/2022 | 8 |
| 167-009 | 167-009_SA_T1 | 02/12/2022 | 9 |
| 167-010 | 167-010_SA_T1 | 02/12/2022 | 10 |

Preparation of the DNAeasy® PowerSoil® Pro Kit for 50 samples

- CD1: 40 mL
- CD2: 10 mL **/!\ Stored at 4°C**
- CD3: 30 mL
- EA: 25 mL
- C5: 25 mL
- C6: 2.5 mL

Modifications of the manufacturer's instructions

*Step 2:* Use beadbeater for 5 min at 3x1000 oscillations/min

*Step 3:* Centrifuge for 1 min at 15000 g

*Step 4:* Transfer supernatant

About 400 µL for all samples, except < 200 µL for sample n°10

HOWEVER 500 to 600 µL were expected

*Additional step:*

Add 200 µL of lysis buffer, except 600 µL for sample n°10

Repeat steps 2 – 4

*Step 6:* Transfer supernatant

About 600 - 650 µL for all samples, except 550 µL for sample n°10

**15/02/2023**
**Nanodrop (by Jack)**

| Sample Number | Extraction Day | Nanodrop (ng/uL) | Nanodrop 260/280 | Nanodrop 260/230 |
|---|---|---|---|---|
| 1 | 14/02/2023 | 84.08 | 1.87 | 1.22 |
| 2 | 14/02/2023 | 52.1 | 1.91 | 0.19 |
| 3 | 14/02/2023 | 10.0 | 2.25 | 0.05 |
| 4 | 14/02/2023 | 37.1 | 1.97 | 0.10 |
| 5 | 14/02/2023 | 9.57 | 2.33 | 0.03 |
| 6 | 14/02/2023 | 43.1 | 1.87 | 0.23 |
| 7 | 14/02/2023 | 22.4 | 1.97 | 0.16 |
| 8 | 14/02/2023 | 40.2 | 1.97 | 0.13 |
| 9 | 14/02/2023 | 18.9 | 2.29 | 0.06 |
| 10 | 14/02/2023 | 3.09 | 1.14 | 0.03 |

Conclusion: the DNA extracts are of good quality, except for 4 samples (in red).

**17/02/2023**
**QUBIT BR Assay**

| Sample Number | Extraction Day | QUBIT (ng/uL) |
|---|---|---|
| 1 | 14/02/2023 | 40.5 |
| 2 | 14/02/2023 | 30.0 |
| 3 | 14/02/2023 | 6.35 |
| 4 | 14/02/2023 | 23.1 |
| 5 | 14/02/2023 | 4.00 |
| 6 | 14/02/2023 | 23.8 |
| 7 | 14/02/2023 | 12.9 |
| 8 | 14/02/2023 | 23.0 |
| 9 | 14/02/2023 | 10.0 |
| 10 | 14/02/2023 | 1.45 |

Conclusion: the quantities of DNA extracts measured using Nanodrop and QUBIT are consistent between the two methods.

**21/02/2023**

**Preparation of stool samples**

Samples were defrosted for *i)* stool stocks transferred into Sarstedt tubes and *ii)* DNA extractions. Samples were stored at -80C in hydrofish freezer until dowmstream experiments.

| Sample Name | Sample Weight (g) | Volume Lysis Buffer (uL) |
|---|---|---|
| 167.001.ST.T2 | 0.40 | 1500 |
| 167.004.ST.T2 | 0.35 | 1400 |
| 167.005.ST.T2 | 0.34 | 1360 |
| 167.006.ST.T2 | 0.28 | 1120 |
| 167.007.ST.T2 | 0.42 | 1500 |
| 167.008.ST.T2 | 0.42 | 1500 |
| 167.009.ST.T2 | 0.36 | 1440 |
| 167.011.ST.T2 | 0.42 | 1500 |
| 167.012.ST.T2 | 0.41 | 1500 |

**07/03/2023**

**DNA extraction of stool samples** **cf. Protocols APC123-PreFer**

**QUBIT BR Assay**

| Sample Number | Extraction Day | QUBIT (ng/uL) |
|---|---|---|
| 1 | 07/03/2023 | too high |
| 4 | 07/03/2023 | 910.00 |
| 5 | 07/03/2023 | too high |
| 6 | 07/03/2023 | 281.00 |
| 7 | 07/03/2023 | too high |
| 8 | 07/03/2023 | 384.00 |
| 9 | 07/03/2023 | 421.00 |
| 11 | 07/03/2023 | 308.00 |
| 12 | 07/03/2023 | 306.00 |

**27/03/2023**

**Reception of samples**

Transported with ice packs to Teagasc
Stored in -80°C hydrofish freezer (A.2)
Update of the list of samples: **cf. APC167_List_Samples_CBF.xlsx**

**11/04/2023**

**Preparation of stool samples**

Samples were defrosted for *i)* stool stocks transferred into Sarstedt tubes and *ii)* DNA extractions. Samples were stored at -80C in hydrofish freezer until dowmstream experiments.

| Sample Name | Sample Weight (g) | Volume Lysis Buffer (µL) |
|---|---|---|
| 167.010.ST.T2 | 0.37 | 1480 |
| 167.015.ST.T2 | 0.24 | 960 |
| 167.016.ST.T2 | 0.39 | 1500 MAX |

| | | |
|---|---|---|
| 167.017.ST.T2 | 0.33 | 1320 |
| 167.018.ST.T2 | 0.35 | 1400 |
| 167.020.ST.T2 | 0.32 | 1280 |

**13/04/2023**
**Extraction of DNA**

Selection of 10 samples of saliva / swabs

| | | | |
|---|---|---|---|
| 167-011 | 167-011_SA_T1 | 02/12/2022 | 11 |
| 167-012 | 167-012_SA_T1 | 03/02/2023 | 12 |
| 167-013 | 167-013_SA_T1 | 03/02/2023 | 13 |
| 167-014 | 167-014_SA_T1 | 03/02/2023 | 14 |
| 167-015 | 167-015_SA_T1 | 03/02/2023 | 15 |
| 167-016 | 167-016_SA_T1 | 03/02/2023 | 16 |
| 167-017 | 167-017_SA_T1 | 03/02/2023 | 17 |
| 167-018 | 167-018_SA_T2 | 27/03/2023 | 18 |
| 167-019 | 167-019_SA_T2 | 27/03/2023 | 19 |
| 167-020 | 167-020_SA_T2 | 27/03/2023 | 20 |

Modifications of the manufacturer's instructions
*Step 2:* Use beadbeater for 5 min at 3x1000 oscillations/min
*Step 3:* Centrifuge for 1 min at 15000 g
*Step 4:* Transfer supernatant
About 400 µL for all samples
HOWEVER 500 to 600 µL were expected
*Step 6:* Transfer supernatant
About 600 - 650 µL for all samples

**26/04/2023**
**Illumina DNA Prep cf. Illumina DNA Prep Reference Guide**

DNA Dilution **cf. APC167_Shotgun_Run1.xlsx**
The DNA purity of the 72 samples was assessed using Nanodrop.

Starting material of the 72 samples was normalised to equal DNA concentration. At least 1 ng of total DNA in 30 µL is the minimum amount required, which corresponds to a concentration of 0.034 ng/µL. Dilutions of 10-fold range were prepared in 1.5 mL Eppendorf tubes, as indicated in APC167_Shotgun_Run1.xlsx. Concentrated 10X samples were normalised to 0.34 ng/µL in a 96-well plate N1, as indicated in APC167_Shotgun_Run1.xlsx. The quantity of DNA is 10.2 ng in 30 µL.

In a new 96-well plate N2, samples were normalised to 0.034 ng/µL, adding 3 µL of 10X samples and 27 µL biomolecular water. The quantity of DNA is 1.2 ng in 30 µL.

The plates were kept at -30°C overnight.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 10 | 18 | 26 | 34 | 42 | 51 | 59 | 67 | 37 | **7_SA** | **15_SA** |
| 2 | 11 | 19 | 27 | 35 | 43 | 52 | 60 | 68 | H2O | **8_SA** | **16_SA** |
| 3 | 12 | 20 | 28 | 36 | 45 | 53 | 61 | 69 | **1_SA** | **9_SA** | **17_SA** |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 13 | 21 | 29 | MIX | 46 | 54 | 62 | 73 | **2_SA** | **10_SA** | **18_SA** |
| 6 | 14 | 22 | 30 | 38 | 47 | 55 | 63 | 77 | **3_SA** | **11_SA** | **19_SA** |
| 7 | 15 | 23 | 31 | 39 | 48 | 56 | 64 | 78 | **4_SA** | **12_SA** | **20_SA** |
| 8 | 16 | 24 | 32 | 40 | 49 | 57 | 65 | 79 | **5_SA** | **13_SA** | **H2O** |
| 9 | 17 | 25 | 33 | 41 | 50 | 58 | 66 | 80 | **6_SA** | **14_SA** | **H2O** |

Pure DNA extract |**10X diluted DNA extract** | **100X diluted DNA extract**

**27/04/2023**
**Illumina DNA Prep** **cf. Illumina DNA Prep Reference Guide**

Reagents:

- Illumina DNA Prep Tagmentation (M) Beads REF 20015880 LOT 20679826

/!\ DLC 2023/04/11

- IDT for Illumina DNA/RNA UD indexes Set A REF 20026121 LOT 20658104
- Illumina DNA Prep IPB + Buffers REF 20049006 LOT 20654085
- Illumina DNA Prep PCR + Buffers REF 20015829 LOT 20670024

Tagment Genomic DNA

**Bring BLT and TB1 to RT.**

The 96-well PCR plate N1 with a DNA concentration at 0.34 ng/µL was used. The total volume is 27 µL. The total DNA quantity was 9.18 ng.

**The lid preheat of the thermal cycler is automatic.**

The final reaction volume of the thermal cycler was set at 47 µL.

Post Tagmentation Clean-up

**Bring TWB to RT.**

**Heat TSB at 37°C for 10 minutes.**

**The lid preheat of the thermal cycler is automatic.**

The final reaction volume of the thermal cycler was set at 57 µL.

Amplify Tagmented DNA

**Thaw EPM on ice.**

**Thaw index adapters at RT.**

The PCR MM was prepared in a 50 mL Falcon tube and 175 µL of MM PCR were transferred into three 8-tip strips. The total DNA input is 9.18 ng. The number of PCR cycles was set at 12. The plate was kept at 2-8°C overnight.

**28/04/2023**
**Illumina DNA Prep** **cf. Illumina DNA Prep Reference Guide**

Clean Up Libraries

**Bring IPB & RSB to RT.**

**Prepare fresh 80% ethanol** (40 mL 100% ethanol + 10 mL biomolecular water)

At step 3, the samples 11H and 7E were containing < 45 µL.

At sept 5-g, the samples 7C, 8C, and 10C were containing < 125 µL.

At step 14, 33 µL RSB were used to resuspend the beads.

At step 18, the sample 2E was containing < 30 µL. A volume of 5 µL RSB was added.

/!\ Some beads may still be present in the plate.

The plates were kept at -20°C.

**01/05/2023**
**Illumina DNA Prep** **cf. Illumina DNA Prep Reference Guide**

Library Quantification **cf. APC167_Shotgun_Run1.xlsx**
Each library were individually quantified using QUBIT dsDNA HS Assay kit.

**/!\ Temperature fluctuations are critical for the accuracy of the assay. All solutions are brought to RT. /!\ The machine can raise the temperature of the assay solution significantly!**

In 50 mL Falcon tube: mix 120 µL of Reagent + 23880 µL of Buffer

In QUBIT strips: 199 µL Working Solution + 1 µL Sample

Transfer 1.5 µL of samples into 8-well PCR strips.

Incubate for 2 minutes in the dark.

NB: The tubes were read twice. However, the measurements were very different. It would be better to run another QUBIT dsDNA HS Assay to confirm the values.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 62.4 | 0.372 | 52.0 | 48.6 | 81.4 | 56.8 | 22.8 | 56.6 | 73.8 | 48.2 | **47.2** | **20.0** |
| 53.0 | 48.8 | 55.4 | 92.8 | 72.2 | 21.4 | 40.2 | 32.8 | 39.8 | 0.942 | **40.6** | **26.4** |
| 50.0 | 12.3 | 53.2 | 80.2 | 77.8 | 52.0 | 49.0 | 80.4 | 40.4 | **20.2** | **25.4** | **50.6** |
| 27.0 | 72.8 | 65.2 | 56.0 | 98.4 | 66.0 | 49.0 | 53.8 | 45.2 | **77.0** | **30.0** | **50.8** |
| 10.9 | 65.0 | 59.2 | 65.8 | 101 | 96.6 | 53.8 | 39.8 | 17.2 | **86.6** | **30.0** | **55.4** |
| 40.8 | 34.0 | 38.0 | <120 | <120 | 40.4 | 21.0 | 41.4 | 40.2 | **45.2** | **31.2** | **61.8** |
| 31.8 | 68.2 | 38.6 | 56.2 | 47.0 | 67.2 | 41.4 | 24.2 | 40.6 | **61.8** | **15.5** | **0.728** |
| 33.4 | 48.8 | 64.6 | 51.0 | 45.0 | 41.8 | 32.4 | 56.2 | 16.3 | **31.8** | **32.4** | **0.718** |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 10 | 18 | 26 | 34 | 42 | 51 | 59 | 67 | 37 | **7_SA** | **15_SA** |
| 2 | 11 | 19 | 27 | 35 | 43 | 52 | 60 | 68 | H2O | **8_SA** | **16_SA** |
| 3 | 12 | 20 | 28 | 36 | 45 | 53 | 61 | 69 | **1_SA** | **9_SA** | **17_SA** |
| 5 | 13 | 21 | 29 | MIX | 46 | 54 | 62 | 73 | **2_SA** | **10_SA** | **18_SA** |
| 6 | 14 | 22 | 30 | 38 | 47 | 55 | 63 | 77 | **3_SA** | **11_SA** | **19_SA** |
| 7 | 15 | 23 | 31 | 39 | 48 | 56 | 64 | 78 | **4_SA** | **12_SA** | **20_SA** |
| 8 | 16 | 24 | 32 | 40 | 49 | 57 | 65 | 79 | **5_SA** | **13_SA** | **H2O** |
| 9 | 17 | 25 | 33 | 41 | 50 | 58 | 66 | 80 | **6_SA** | **14_SA** | **H2O** |

**02/05/2023**
**Illumina DNA Prep** **cf. Illumina DNA Prep Reference Guide**

Library Quality **cf. APC167_Shotgun_Run1.xlsx**

**Bring Reagents to RT.**

The molarity of each samples was quantified to normalise the library at 4 nM final concentration.

**04/05/2023**
**Illumina DNA Prep cf. APC167_Shotgun_Run1.xlsx**

Step 1: Pooling samples at 16 Nm
To allow accurate pipetting, highly concentrated samples were pooled at final concentration 16 nM. A volume of 1154.4 µL of RSB was transferred into a new 1.5 mL Eppendorf. Appropriate volumes of each DNA samples were added to the 1.5 mL Eppendorf.

Step 2: Pooling samples at 4 nM
Lower concentrated DNA samples were pooled at final concentration 4 nM. A volume of 131.9 µL of RSB (116.9 + 15 µL) was transferred into a new 1.5 mL Eppendorf. A dilution to 4 nM of the 16nM pool was prepared by adding 5 µL. Appropriate volumes of each DNA samples were added to the 1.5 mL Eppendorf.

Step 3: QUBIT HS dsDNA Assay

We look for a final concentration of the library pool between 5-10 ng/µL. **This value will give an indication of the volume of RSB to resuspend the beads during the clean-up.**

**Before_clean_16nM_pool: 9.26 ng/µL**

Before_clean_4nM_pool_ERROR: 20.2 ng/µL

Step 4: Removing the small PCR amplicons
Following the Illumina DNA Prep Protocol, 180 µL of the 4 nM pool was clean up.

- ✓ Transfer 180 µL of the pool into a new 1.5 mL Eppendorf.
- ✓ Vortex and invert IPB multiple times to resuspend.
- ✓ Add 126 µL of IPB to respect the ratio 0.7:1 (70% of the volume).
- ✓ Pipette 10 times to mix.
- ✓ Incubate at RT for 5 min.
- ✓ Place on magnetic stand and wait until the liquid is clear (~5 min).
- ✓ Without disturbing the beads, remove and discard supernatant.
- ✓ Add 200 µL of 80% EtOH without mixing.
- ✓ Incubate 30 seconds.
- ✓ Without disturbing the beads, remove and discard supernatant.
- ✓ Add 200 µL of 80% EtOH without mixing.
- ✓ Incubate 30 seconds.
- ✓ Without disturbing the beads, remove and discard supernatant.
- ✓ Use a 20 µL pipette to remove and discard residual EtOH.
- ✓ Air-dry on the magnetic stand for 5 min.
- ✓ Remove from the magnetic stand.
- ✓ Add between **182 µL** of RSB to the beads.
- ✓ Pipette to resuspend.
- ✓ Incubate at RT for 2 min.
- ✓ Place on magnetic stand and wait until the liquid is clear (~2 min).
- ✓ Transfer between **180 µL** of supernatant to a new 1.5 mL Eppendorf tube.

Step 5: Agilent 2100 Bioanalyzer HS DNA

The following 1:5 dilutions were prepared: Before_clean_16nM_pool: 2.5 µL + 10 µL RSB
Before_clean_4nM_pool: 2.5 µL + 10 µL RSB

**Average sizes (bp):** f12_sample (dil-1:7.5): 550 bp
Before_clean_16nM_pool (dil-1:5): 600 bp
Before_clean_4nM_pool (dil-1:5): >1000 bp

**/!\ERROR OF VOLUME for the preparation of 4 nM pool: an additional volume of 54.8 µL of RSB is needed to adjust the molarity to 4nM.**

**FIXING THE ERROR OF MOLARITY**

Step 2: Pooling samples at 4 nM

Lower concentrated DNA samples were pooled at final concentration 4 nM. Appropriate volumes of **f12 sample** was added into the 1.5 mL Eppendorf. A volume of 54.8 µL of RSB was transferred into the 1.5 mL Eppendorf.

Step 3: QUBIT HS dsDNA Assay

**Before_clean_4nM_pool_ADJUSTED: 0.190 ng/µL**

Step 4: Removing the small PCR amplicons

- ✓ Transfer 120 µL of the pool into a new 1.5 mL Eppendorf.
- ✓ Vortex and invert IPB multiple times to resuspend.
- ✓ Add 84 µL of IPB to respect the ratio 0.7:1 (70% of the volume).
- ✓ Pipette 10 times to mix.
- ✓ Incubate at RT for 5 min.
- ✓ Place on magnetic stand and wait until the liquid is clear (~5 min).
- ✓ Without disturbing the beads, remove and discard supernatant.
- ✓ Add 200 µL of 80% EtOH without mixing.
- ✓ Incubate 30 seconds.
- ✓ Without disturbing the beads, remove and discard supernatant.
- ✓ Add 200 µL of 80% EtOH without mixing.
- ✓ Incubate 30 seconds.
- ✓ Without disturbing the beads, remove and discard supernatant.
- ✓ Use a 20 µL pipette to remove and discard residual EtOH.
- ✓ Air-dry on the magnetic stand for 5 min.
- ✓ Remove from the magnetic stand.
- ✓ Add between **42 µL** of RSB to the beads.
- ✓ Pipette to resuspend.
- ✓ Incubate at RT for 2 min.
- ✓ Place on magnetic stand and wait until the liquid is clear (~2 min).
- ✓ Transfer between **40 µL** of supernatant to a new 1.5 mL Eppendorf tube.

Step 5: Agilent 2100 Bioanalyzer HS DNA

The following 1:5 dilutions were prepared: After_clean_4nM_pool_ADJUSTED: 1 + 4 µL RSB

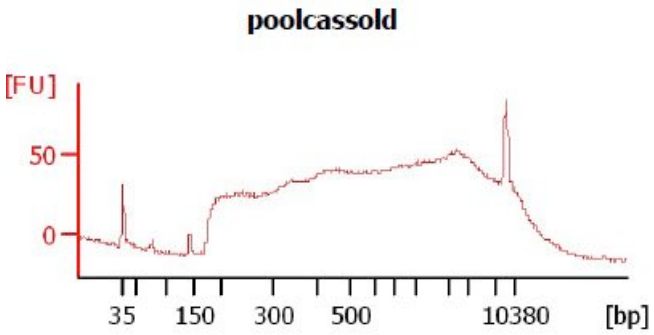
**Average sizes (bp):** After_clean_4nM_pool_ADJUSTED:
After_clean_4nM_pool_ADJUSTED (dil-1:5):

Step 6: QUBIT HS dsDNA Assay

**After_clean_4nM_pool_ADJUSTED: 0.650 ng/µL**

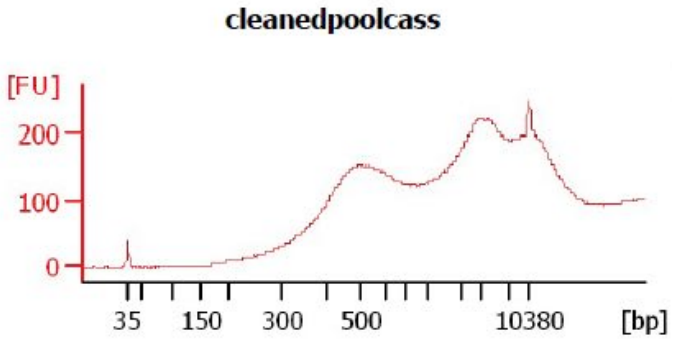
Resume:

1. **Starting from this pool of 4nM samples**

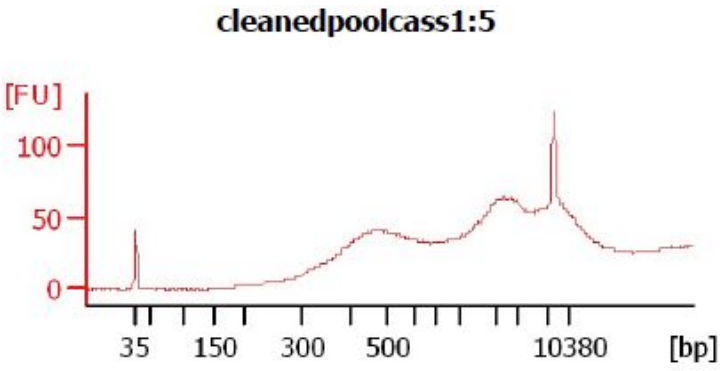



Pool 04/05/2023 (13 ng/µL)

2. **Clean-up to remove small PCR amplicons**




Pool 05/05/2023 (29 ng/µL)




Pool 05/05/2023 dil1:5 (6.4 ng/µL)

3. **Clean-up to remove larger PCR amplicons**

Pool 05/05/2023 (3.9 ng/µL)

**02/06/2023**
**Reception of samples**

Transported with ice packs to Teagasc
Stored in -80°C hydrofish freezer (A.2)
Update of the list of samples: **cf. APC167_List_Samples_CBF.xlsx**

**16/06/2023**
**Preparation of stool samples**

Samples were defrosted for *i)* stool stocks transferred into Sarstedt tubes and *ii)* DNA extractions. Samples were stored at -80C in hydrofish freezer until dowmstream experiments.

| Sample Name | Sample Weight (g) | Volume Lysis Buffer (µL) |
|---|---|---|
| 167.019.ST.T2 | 0.38 | 1500 MAX |
| 167.021.ST.T2 | 0.31 | 1240 |
| 167.022.ST.T2 | 0.29 | 1160 |

**12/07/2023**
**Reception of samples**

Transported with ice packs to Teagasc
Stored in -80°C hydrofish freezer (A.2)
Update of the list of samples: **cf. APC167_List_Samples_CBF.xlsx**

**15/08/2023**
**Preparation of stool samples**

Samples were defrosted for *i)* stool stocks transferred into Sarstedt tubes and *ii)* DNA extractions. Samples were stored at -80C in hydrofish freezer until dowmstream experiments.

| Sample Name | Sample Weight (g) | Volume Lysis Buffer (µL) |
|---|---|---|
| 167.023.ST.T2 | 0.42 | 1500 MAX |
| 167.025.ST.T2 | 0.22 | 880 |

**18/10/2023**
**DNA extraction of stool samples** **cf. Protocols APC123-PreFer**

Selection of samples: 167010, 167015, 167016, 167017, 167018, 167019, 167020, 167021, 167022, 167023, 167025
First part of the protocol (stop at the isopropanol step).

**19/10/2023**
**DNA extraction of stool samples** **cf. Protocols APC123-PreFer**

Second part of the protocol.

**27/10/2023**
**Reception of the last samples (by Aoife)**

Transported with ice packs to Teagasc
Stored in -80°C hydrofish freezer (A.2)
Update of the list of samples: **cf. APC167_List_Samples_CBF.xlsx**

**01/11/2023**
**Preparation of stool samples**

Samples were defrosted for *i)* stool stocks transferred into Sarstedt tubes and *ii)* DNA extractions. Samples were stored at -80C in hydrofish freezer until dowmstream experiments.

| Sample Name | Sample Weight (g) | Volume Lysis Buffer (µL) |
|---|---|---|
| 167.027.ST.T2 | 0.20 | 700 |
| 167.028.ST.T2 | 0.36 | 1440 |
| 167.029.ST.T2 | 0.28 | 1120 |

**03/11/2023**
**DNA extraction of stool sample N°27** **cf. Protocols APC123-PreFer**

**08/11/2023**
**DNA extraction of stool sample N°28 & N°29** **cf. Protocols APC123-PreFer**

**09/11/2023**
**Extraction of DNA**

Selection of 10 samples of saliva / swabs

| | | | |
|---|---|---|---|
| 167-021 | 167-021_SA_T1 | 02/06/2023 | 21 |
| 167-022 | 167-022_SA_T1 | 02/06/2023 | 22 |
| 167-023 | 167-023_SA_T1 | 12/07/2023 | 23 |
| 167-024 | 167-024_SA_T1 | 12/07/2023 | 24 |
| 167-025 | 167-025_SA_T1 | 12/07/2023 | 25 |
| 167-026 | 167-026_SA_T1 | 12/07/2023 | 26 |
| 167-027 | 167-027_SA_T1 | 27/10/2023 | 27 |
| 167-028 | 167-028_SA_T1 | 27/10/2023 | 28 |
| 167-029 | 167-029_SA_T1 | 27/10/2023 | 29 |
| 167-030 | 167-030_SA_T1 | 27/10/2023 | 30 |

Modifications of the manufacturer's instructions
*Step 2:* Use beadbeater for 5 min at 3x1000 oscillations/min
*Step 3:* Centrifuge for 1 min at 15000 g
*Step 4:* Transfer supernatant
About 400 µL for all samples
HOWEVER 500 to 600 µL were expected
*Step 6:* Transfer supernatant
About 600 - 650 µL for all samples

**24/11/2023**
**Illumina DNA Prep** **cf. Illumina DNA Prep Reference Guide**

DNA Dilution **cf. APC167-Shotgun_Run2.xlsx**
The DNA purity of the samples was assessed using Nanodrop.

Starting material of the samples was normalised to equal DNA concentration. At least 1 ng of total DNA in 30 µL is the minimum amount required, which corresponds to a concentration of 0.034 ng/µL. Dilutions of 10-fold range were prepared in 1.5 mL Eppendorf tubes, as indicated in APC167-Shotgun_Run2.xlsx. The concentrations were normalised to 0.34 ng/µL in a 96-well plate 1, as indicated in APC167-Shotgun_Run2.xlsx. The quantity of DNA is 10.2 ng in 30 µL.

The plates were kept at -30°C overnight.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1-ST | 11-ST | 21-ST | 22-SA | 30-SA | 88 | 97 | 105 | 113 | D6 | D14 | D22 |
| 4-ST | 12-ST | 22-ST | 23-SA | 81 | 89 | 98 | 106 | 114 | D7 | D15 | D23 |
| 5-ST | 15-ST | 23-ST | 24-SA | 82 | 90 | 99 | 107 | 115 | D8 | D16 | D24 |
| 6-ST | 16-ST | 25-ST | 25-SA | MIX | 91 | 100 | 108 | D1 | D9 | D17 | D25 |
| 7-ST | 17-ST | 27-ST | 26-SA | 84 | 93 | 101 | 109 | D2 | D10 | D18 | D26 |
| 8-ST | 18-ST | 28-ST | 27-SA | 85 | 94 | 102 | 110 | D3 | D11 | D19 | D27 |
| 9-ST | 19-ST | 29-ST | 28-SA | 86 | 95 | 103 | 111 | D4 | D12 | D20 | D28 |
| 10-ST | 20-ST | 21-SA | 29-SA | 87 | 96 | 104 | 112 | D5 | D13 | D21 | H2O |

Pure DNA extract |**10X diluted DNA extract** | **100X diluted DNA extract** | **1000X diluted DNA extract**

**27/11/2023**
**Illumina DNA Prep** **cf. Illumina DNA Prep Reference Guide**

Reagents:

- Illumina DNA Prep Tagmentation (M) Beads REF 20015880 LOT 20663581
- Illumina DNA Prep UD indexes Set A REF 20026121 LOT 20658104
- Illumina DNA Prep IPB + Buffers REF 20049006 LOT 20654085
- Illumina DNA Prep PCR + Buffers REF 20015829 LOT 20682834

Tagment Genomic DNA
Post Tagmentation Clean-up
Amplify Tagmented DNA

The plate was kept at 2-8°C overnight.

**28/11/2023**
**Illumina DNA Prep** **cf. Illumina DNA Prep Reference Guide**

Clean Up Libraries
/!\ The following samples were mixed together:

- D6 + D14
- D7 + D15
- D8 + D16
- D9 + D17
- D10 + D18
- D11 + D19
- D12 + D20
- D13 + D21

Library Quantification **cf. APC167_Shotgun_Run2.xlsx**
Each library were individually quantified using QUBIT dsDNA HS Assay kit.

**/!\ Temperature fluctuations are critical for the accuracy of the assay. All solutions are brought to RT. /!\ The machine can raise the temperature of the assay solution significantly!**

**01/12/2023**
**Bioanalyser** **cf Agilent High Sensitivity DNA Kit Guide**

**RESUME**

**Stool samples transferred into Sarstedt tubes are stored in -80C hydrofish freezer. There are no saliva samples stored. Tissue samples are stored in -80C hydrofish freezer.**

**All samples are stored in Hydrofish -80C Freezer in A2 compartment:**

- **2 boxes Oral Microbiome with stool samples**
- **1 box Oral Microbiome with saliva samples**
- **1 box Oral Microbiome with tissue samples in RNA later**

**DNA extracts sent to Shotgun_Run1 & Shotgun_Run2 are stored in -80C Freezer Freezer in A2 compartment.**

# Appendix I

| Patient_Number | Group | Consent | Date_Birth | Age | Sex | Height (cm) | Weight (kg) | BMI (kg/m2) | BMI_Category |
|---|---|---|---|---|---|---|---|---|---|
| APC167-001 | PM | 03/10/2022 | 31/10/1990 | 31 | Male | 171 | 84 | 28.7 | Overweight |
| APC167-002 | B | 11/10/2022 | 22/05/1983 | 39 | Male | 172 | 100 | 33.8 | Obesity |
| APC167-003 | PM | 10/10/2022 | 05/08/1962 | 60 | Female | 158 | 72.3 | 28.9 | Overweight |
| APC167-004 | PM | 17/10/2022 | 28/09/1951 | 71 | Female | 160 | 69.0 | 26.9 | Overweight |
| APC167-005 | M | 18/10/2022 | 09/05/1961 | 61 | Male | 175 | 90 | 29.4 | Overweight |
| APC167-006 | PM | 24/10/2022 | 25/01/1977 | 45 | Female | 160 | 98 | 38.2 | Obesity |
| APC167-007 | B | 01/11/2022 | 28/01/1971 | 51 | Male | 172 | 89 | 30.1 | Obesity |
| APC167-008 | PM | 16/11/2022 | 07/06/1944 | 78 | Male | 168 | 70 | 24.8 | Normal_Weight |
| APC167-009 | B | 22/11/2022 | 18/09/1973 | 49 | Female | 165 | 49 | 18.0 | Underweight |
| APC167-010 | PM | 28/10/2022 | 27/09/1954 | 68 | Female | 160 | 87 | 34.0 | Obesity |
| APC167-011 | PM | 28/11/2022 | 29/05/1962 | 60 | Male | 170 | 88 | 30.4 | Obesity |
| APC167-012 | B | 12/12/2022 | 11/01/1960 | 62 | Male | 181 | 92.1 | 28.1 | Overweight |
| APC167-013 | B | 13/12/2022 | 13/04/1979 | 43 | Female | 166 | 90 | 32.7 | Obesity |
| APC167-014 | B | 19/01/2023 | 12/03/1978 | 44 | Male | 178 | 95 | 29.9 | Overweight |
| APC167-015 | B | 19/01/2023 | 07/04/1998 | 24 | Male | 180 | 70 | 21.6 | Normal_Weight |
| APC167-016 | B | 19/01/2023 | 13/01/1980 | 43 | Male | 175 | 81 | 26.4 | Overweight |
| APC167-017 | B | 02/02/2023 | 22/12/1965 | 57 | Female | 162 | 87.9 | 33.5 | Obesity |
| APC167-018 | B | 14/02/2023 | 11/01/1959 | 64 | Female | 157 | 45 | 18.3 | Underweight |
| APC167-019 | B | 28/02/2023 | 14/04/1960 | 62 | Male | 181 | 85.2 | 26.0 | Overweight |
| APC167-020 | M | 28/02/2023 | 05/02/1960 | 63 | Female | 157 | 68 | 27.6 | Overweight |
| APC167-021 | B | 30/03/2023 | 20/09/1971 | 51 | Male | 185 | 101 | 29.5 | Overweight |
| APC167-022 | M | 24/04/2023 | 06/11/1963 | 59 | Female | 166 | 59 | 21.4 | Normal_Weight |
| APC167-023 | M | 15/06/2023 | 28/02/1978 | 45 | Female | 165 | 59.8 | 21.9 | Normal_Weight |
| APC167-024 | B | 22/06/2023 | 09/04/1947 | 76 | Female | 164 | 50 | 18.5 | Normal_Weight |
| APC167-025 | M | 27/06/2023 | 05/12/1958 | 64 | Male | 172 | 81 | 27.4 | Overweight |
| APC167-026 | B | 06/07/2023 | 07/05/1964 | 59 | Male | 190 | 107 | 29.6 | Overweight |
| APC167-027 | B | 03/08/2023 | 03/08/1970 | 53 | Male | 182 | 19.1 | 19.1 | Normal_Weight |
| APC167-028 | M | 20/09/2023 | 02/02/1948 | 75 | Male | 170 | 85.2 | 29.5 | Overweight |
| APC167-029 | PM | 26/09/2023 | 23/11/1957 | 68 | Male | 170 | 76 | 26.3 | Overweight |
| APC167-030 | M | 11/10/2023 | 13/07/1954 | 69 | Male | 171 | 67 | 22.9 | Normal_Weight |

| BPE_Score_Up | BPE_Score_Down | Oral_Health | Antibiotic_Use | Probiotic_Use | Probiotics | Current_Medication | Smoker |
|---|---|---|---|---|---|---|---|
| 2-1-2 | 1-2-1 | Moderate_Oral_Health | No | No | No | None | Former |
| 3-1-3 | 2-2-/ | Poor_Oral_Health | No | Yes | Zenflora (irregularly) | Daily | Current |
| 2-2-2 | /-2-/ | Moderate_Oral_Health | No | No | No | Daily | Current |
| 1-0-1 | 1-2-1 | Good_Oral_Health | No | No | No | Daily | Never |
| 1-0-0 | 0-2-1 | Good_Oral_Health | No | No | No | Daily | Never |
| 3-2-3 | 3-2-3 | Poor_Oral_Health | No | No | No | Daily | Former |
| 3-2-3 | 3-2-3 | Poor_Oral_Health | No | No | No | Daily | Current |
| /-/-/ | /-/-/ | Very_Poor_Oral_Health | No | No | No | Daily | Former |
| 3-2-3 | 3-3-3 | Poor_Oral_Health | No | No | No | None | Current |
| /-/-/ | /-2-2 | Poor_Oral_Health | No | No | No | Daily | Current |
| 2-2-2 | /-2-2 | Moderate_Oral_Health | No | No | No | None | Never |
| 2-1-2 | 2-2-2 | Moderate_Oral_Health | No | No | No | Daily | Never |
| 3-3-3 | 3-4-3 | Very_Poor_Oral_Health | Yes | No | No | Daily | Former |
| 3-3-3 | 3-3-3 | Poor_Oral_Health | No | No | No | None | Current |
| 2-1-2 | 2-1-2 | Moderate_Oral_Health | No | No | No | None | Never |
| 2-2-2 | 2-3-2 | Poor_Oral_Health | No | No | No | None | Never |
| 2-2-3 | 1-2-2 | Moderate_Oral_Health | No | No | No | Daily | Never |
| 2-1-2 | 1-1-1 | Good_Oral_Health | No | No | No | Weekly | Never |
| 2-/-2 | 1-2-1 | Moderate_Oral_Health | No | No | No | Daily | Former |
|  |  | Moderate_Oral_Health | No | No | No | Daily | Never |
| 3-2-3 | 2-3-2 | Poor_Oral_Health | No | No | No | None | Never |
| 3-0-3 | 3-2-3 | Moderate_Oral_Health | No | Yes | Kefir (occasionnal) | None | Former |
| 3-3-3 | 3-3-3 | Poor_Oral_Health | No | No | No | None | Former |
| /-/-/ | /-3-/ | Very_Poor_Oral_Health | No | No | No | Daily | Never |
| 3-2-3 | 2-2-3 | Poor_Oral_Health | No | No | No | Daily | Former |
| /-/-/_no teeth | /-/-/_no teeth | Very_Poor_Oral_Health | Yes(amoxicillin) | No | No | Daily | Former |
| 4-3-3 | 4-3-4 | Very_Poor_Oral_Health | Yes(amoxicillin) | No | No | Daily | Former |
| 3-3-3 | 3-3-3 | Poor_Oral_Health | Yes(amoxicillin) | No | No | As_Needed | Never |
| 2-/-2 | 2-3-2 | Very_Poor_Oral_Health | No | No | No | Daily | Never |
| 3-3-3 | 3-3-3 | Poor_Oral_Health | No | No | No | None | Current |

| Cigarettes/Day | Years_Smoking | Alcohol_Consumption | Units_EtOH/Week | Co-morbidities |
|---|---|---|---|---|
| >20 | 17 | No | 0 | 3 years sober (alcoholic + drug user) |
| 20 | 25 | Yes | 12 | asthma |
| 20 | 25 | Yes | 2 | underactive thyroid, stomach acid reflux, chronic pain |
| 0 | 0 | Yes | 10 | hypertension |
| 0 | 0 | Yes | 2 | high cholesterol |
| 40 | 20 | Yes | 2 | gastric reflux, back pain |
| 20 | 30 | Yes | 30 | / |
| / | / | Yes | 16 | / |
| 20 | 20 | Yes | 15 | / |
| 10 | 40 | Yes | 5 | osteoarthritis, genital lichen planus, acid reflux |
| 0 | 0 | No | 0 | / |
| 0 | 0 | Yes | 20 | / |
| 5 | 20 | Yes | >40 | chronic liver disease, ex-alcoholic |
| 0.25 | 10 | Yes | 10 | / |
| 0 | 0 | Yes | 1 | / |
| 0 | 0 | Yes | 12 | / |
| 0 | 0 | No | 0 | history of ovarian cancer 2015, depression |
| 0 | 0 | No | 0 | / |
| 20 | 10 | No | 0 | / |
| 0 | 0 | No | 0 | ulcerative colitis, osteoporosis |
| 0 | 0 | Yes | 10 | vit B12 deficiency |
| 15 | 20 | No | 0 | / |
| 2 | 15 | No | 0 | / |
| 0 | 0 | No | 0 | / |
| 20 | 30 | No | 0 | hypertension |
| / | / | Yes | 4 | rheumatoid arthritis, high cholesterol |
| / | / | No | 0 | ex-alcoholic |
| 0 | 0 | Yes | 8 | asthma |
| 0 | 0 | Yes | 4 | / |
| 10 | 30 | Yes | >40 | poor health attender |