Characterisation of the skin and oral microbiome of younger and older adults in a UK population.

Duncan James Newman

Doctor of Philosophy

Aston University

May 2025

© Duncan James Newman 2025

Duncan James Newman asserts their moral right to be identified as the author of this thesis.

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright belongs to its author and that no quotation from the thesis and no information derived from it may be published without appropriate permission or acknowledgement.

Aston University

Characterisation of the skin and oral microbiome of younger and older adults in a UK population.

Duncan James Newman

May 2025

Thesis Abstract

The human microbiome has been implicated in contributing to various aspects of human health and disease. One aspect is the ageing process, which in humans is exhibited through physical and functional deterioration. During ageing the microbiome undergoes taxonomic shifts, some contribute to various chronic comorbidities in ageing, while others are due to ageing changes in the host. Studies of the human skin and oral microbiomes in the UK were revealed to account for a combined 4% of UK microbiome research in 2021, therefore there is a lack of knowledge about these microbiomes in UK populations.

A suitable sampling strategy for obtaining skin and oral microbiome samples in non-clinical settings was implemented on a participant population of 30 young adults (aged 20-40 years old), and 29 older adults (aged 60-80 years old). Two skin samples (from the right antecubital fossa and the right face cheek) and one oral (an oral rinse) were taken from each participant, during sampling metadata about participant health and hygiene habits were also obtained. An optimised bioinformatics pipeline was developed to enable analysis of 16S rRNA sequencing of each sample.

Analysis of the skin microbiomes revealed significant differences in beta diversity associated with age and with females in age whereas the male populations tended to be similar, specifically a reduction in *Cutibacterium acnes* population was demonstrated. With increasing age, the oral microbiome was characterised by an increase in various *Prevotella* species and an increase in the pathogenic potential of the microbiome metabolome, however participant health rather than gender was identified to have an influence on these differences. Across all microbiomes a reduction in community resilience in age was seen. Overall, this study identified changes in all microbiomes investigated associated with ageing in a UK population, and that these changes are different across the skin and oral microbiomes.

Keywords: Human Microbiome, Oral Microbiome, Skin Microbiome, Human Ageing, UK Population, Microbial diversity, Co-occurrence Networks, Network analysis, Microbiome Metabolome

To all those teachers who have got me to where I am now.

GNU TERRY PRATCHETT

Acknowledgements

My opportunity to carry out and complete this thesis would not have been possible without the work and experience of my supervisors who saw the capabilities I had and would develop. Firstly, I would like to thank James Brown for his work in setting up the project, his assistance in obtaining ethical approval and guidance on working with participants. The continued personal support and encouragement from Tony Worthington has been invaluable to throughout this project, along with Anthony Hilton whose input into microbiome and statistical processes has been greatly appreciated. Additionally, they together have provided important contacts both within and outside the university who have assisted with various aspects of administration and study design throughout. Finally, both Peter Lambert and Ann Vernallis who have provided constant guidance during the final stages of data analysis and thesis construction, their insights and encouragement have become indebted to over these past few months.

I would also like to extend my thanks and appreciation to my industry partners at Colgate-Palmolive, who have provided guidance and assistance through the project but particularly with the processing of sequence dataset creation and 16S analysis, these include: Lyndsay Schaeffer and Min Li who have been present throughout, along with Fatima Saravia and Mary Hannah Swaney who've provided technical assistance and suggestions. Additionally, I would like to thank Matthew Marples and John Simms for their work in designing and creating the mould used for the creation of the skin sampling templates.

To the friends and colleagues who I've worked alongside and enjoyed the company of in MB327 and MB436, I thank you for your kind support and endurance of my constant complaining. To Antonia, Gemma, Emily, Tom, Jaspreet, Calum, and Ali, I thank you for many memorable conversations and support over breaks, lunches, pub visits, and badminton games, they were a much-needed way to let of steam and relax throughout the process.

As for my closest friends I thank you for your support and stamina with me throughout the years. To Matt, Ben, Sam, Phil, and Kurt, I deeply appreciate your understanding, emotional support, and encouragement throughout the difficult and good times, they mean more to me than I may necessarily let on. I especially thank you for providing me with opportunities to physically and psychologically escape from the drudgery, stress, and strain, if only briefly.

This leaves only my family who have supported me ceaselessly throughout my entire academic journey. To my mum for constant emotional support and communication, along with financial and practical assistance right back to paying for the train when I did my A-levels. To my dad for his various assistance but particularly driving me to and from various universities, even all the way to devon and back just for an interview that went nowhere, but also right back to taking me to Cambridge science festival for all those years. And my brother for providing an example of determination in a career path and supporting me directly while I have finished this thesis. Beyond this to my grandparents whose advice and wisdom I paid attention to and have tried to implement as I my understanding develops in my own age; I am only regretful that none of you were able to see this achievement and the future it will help bring for me.

While I am deeply thankful to all the above, I also feel that I should apologise unreservedly to most of you for my attitude, abrasiveness, and rudeness at various times over the past few years as I have put this thesis above nearly all other considerations. I am sorry for anything I've put you through and hopefully I can now allow more of the person you knew to return.

Finally, I wish to thank the participants who willingly gave their time up to contribute to this project, it would not have been possible without any of them.

Table of Contents

List of Abbreviations

%	percent
°C	degrees celsius
Ancombc	analysis of compositions of microbiomes with bias correction
ANOVA	analysis of variance
AUDIT-C	alcohol use disorders identification test consumption
bp	base pair
CBA	columbia blood agar
CFU	colony forming unit
cm	centimetre
CXCA	colorex™ candida
CXSA	colorex™ staphylococcus aureus
DADA2	divisive amplicon denoising algorithm 2
DeSeq	differential expression analysis for sequence count data
DH2O	distilled water
DNA	deoxyribonucleic acid
DNAse	deoxyribonuclease
Edger	empirical analysis of digital gene expression data in r
FAA	fastidious anaerobe agar
FI	frailty index
g	gram
HLS	health and life sciences
HOMD	human oral microbiome database
hrs	hours
IMDD	index of multiple deprivation decile
LCC	largest connected component
Lefse	linear discriminant analysis *ef*fect *size*
LPS	lipopolysaccharide
MAC3	macconkey agar no3
Metacyc	metabolic pathway database
ml	millilitre
MSA	mannitol salt agar
ng	nanogram
nm	nanometre
OA	older adult
PBS	phosphate buffered saline
PCoA	principal coordinates analysis
PCR	polymerase chain reaction
PERMANOVA	permutational analysis of variance
Qiime2	quantitative insights into microbial ecology
R1	forward sequencing read
R2	reverse sequencing read
RDP	ribosomal database project
RefSeq	reference sequence

RESCRIPt	reference sequence annotation and curation pipeline
RNA	ribonucleic acid
SCBA	streptococcal selective columbia blood agar
SDA	Sabouraud dextrose agar
SDAC	Sabouraud dextrose agar + clotrimazole
SPIEC-EASI	sparse inverse covariance estimation for ecological association and statistical inference
spp.	species
TSA	tryptone soya agar
UK	united kingdom
UniFrac	unique fraction metric
USA	united states of america
YA	younger adult
µl	microlitre

List of Figures

List of Tables

Chapter 1. Introduction

1.1. The Human Microbiome

The human microbiome is a complex ecosystem unique to each individual, composed of bacteria, fungi, viruses, and protozoa; which in recent years has been understood in terms of its content, growth, and molecular interactions with the body (Hayes and Sahu, 2020). In humans the microbiome is either partially or completely inhabiting in a non-infectious manner the skin, eye, respiratory tract, urinary and reproductive systems, oral cavity, and gastrointestinal tract; microbes found in other systems such as the circulatory or nervous are considered infections. Both the human host and the microbiome are dynamic systems that are capable of responding to alterations to their direct environment caused by each other or external influences, and therefore could be considered a meta-organism (Santoro *et al.*, 2020). Recently, research has shown how the microbiome influences the host body, be it through altering hormonal levels, potentially contributing to diseases such as Alzheimer's, or involvement with the nitric oxide cycle and how this contribute towards atherosclerosis (Shoemark and Allen, 2014; Martin *et al.*, 2018; Walker *et al.*, 2018). These influences can even include effects on the mitochondria via the regulation of mitochondrial biogenesis factors contributing to potentially improved cellular energy production or increased cancer incidences (Clark and Mach, 2017; Jackson and Theiss, 2020). Conversely, the host's actions such as the immune response or dietary changes will alter the environment inhabited by the microbiome, leading to changes in the microbial populations. While both systems are robust and capable of adapting to change, they can still have a significant impact. One such change that has an impact upon both systems is the ageing process, throughout a host's lifetime the microbiome changes in response to growth and development; for example changes in hormone levels during puberty (Ottman *et al.*, 2012). Similarly, towards the end of a host's lifetime the body's functions begin to deteriorate, allowing for different microbes to either colonise, increase or decrease in population, or cause infection. Conversely courses of medical treatments such as antibiotics, host diet, hygiene routine will all impact the microbiome (Bana and Cabreiro, 2019).

1.1.1. Skin Microbiome

The skin microbiome is defined mostly by its environment. The skin is the largest organ in the human body allowing several artificial and natural niches to exist for the microbiome to express its diversity and change or repopulate, in response to changing physiological and external factors. Areas of the skin are either covered or often left uncovered for example the hands or head allowing transient microbes to be acquired from the external environment encountered throughout a person's daily routine and lifestyle choices (Vandegrift *et al.*, 2017). Whereas the other areas of the body are usually covered with clothes (which can vary dependent upon cultural choices/values), which limits contact with potential transient and pathogenic microorganisms allowing the transfer of microorganisms to the clothes; washing these clothes does little to reduce the microbial load due to the modern need for energy efficiency leading to lower wash temperatures coupled with a reduction in bleaching agents being used in detergents lessening their antimicrobial effect (Bockmühl, 2017; Sterndorff *et al.*, 2020). This does allow unique environments to be created for example if the feet are often enclosed in both socks and shoes then this will allow for sweat to build up creating a moist environment; but this can also create the conditions for colonisation with unique specific microorganisms that subsequently cause infections, in this case with the feet dermatophytes caused most often by *Trichophyton rubrum* (Nenoff *et al.*, 2014).

Skin can be physiologically categorised into three types which inform the structure of the microbiome: sebaceous e.g. inside the ear or the middle of the chest, moist e.g. inside the nostril or under the armpit, and dry e.g. the forearm or the hands. Aside from these categories and away from the internal structure of the hair follicles and glands the surface skin is an acidic environment with minimal nutrients (Cundell, 2018). This acidic environment is called the acid mantle and is constructed from a mixture of different metabolic pathways that release organic acids (such as urocanic acid) into their environment, this permeates the stratum corneum and helps contribute to normal skin function along with having shown to affect bacterial metabolic growth (Fluhr and Elias, 2002; Elias, 2017; Li *et al.*, 2023; Brooks *et al.*, 2025).

Skin is not just a single layered barrier that prevents penetration beyond the surface of the human body, as it is often considered by the average lay person. The skin is composed of two main layers (Fig.1.1), the epidermis which is the upper layer and cut off from vascular flow, and the dermis which is the lower layer and not exposed to the environment under normal conditions; members of the skin's microbiome have been found in layers of the dermis (Nakatsuji *et al.*, 2013). The top of the epidermis is a layer called the Stratum Corneum composed of dead keratinocytes filled with keratin, and glycolipids, together these form a strong waterproof barrier that protects the body from damage and potential infection risks. Lower within the epidermis is a layer called the Stratum Spinosum which is in close proximity to the blood vessels to receive nutrients via diffusion and where dendritic and other immune cells can be found along with immune active molecules such as antimicrobial peptides; the other non-immune skin cells within this layer are capable of assisting with the immune response (Nguyen and Soulika, 2019; Mansfield and Naik, 2020; Brito, Baek and Bin, 2024). Within the dermis is a layer called the papillary dermis, cells from the immune system can also be found here. Together these regions support the skin in acting as an active defensive mechanism to the passive defence provided by the skin barrier.

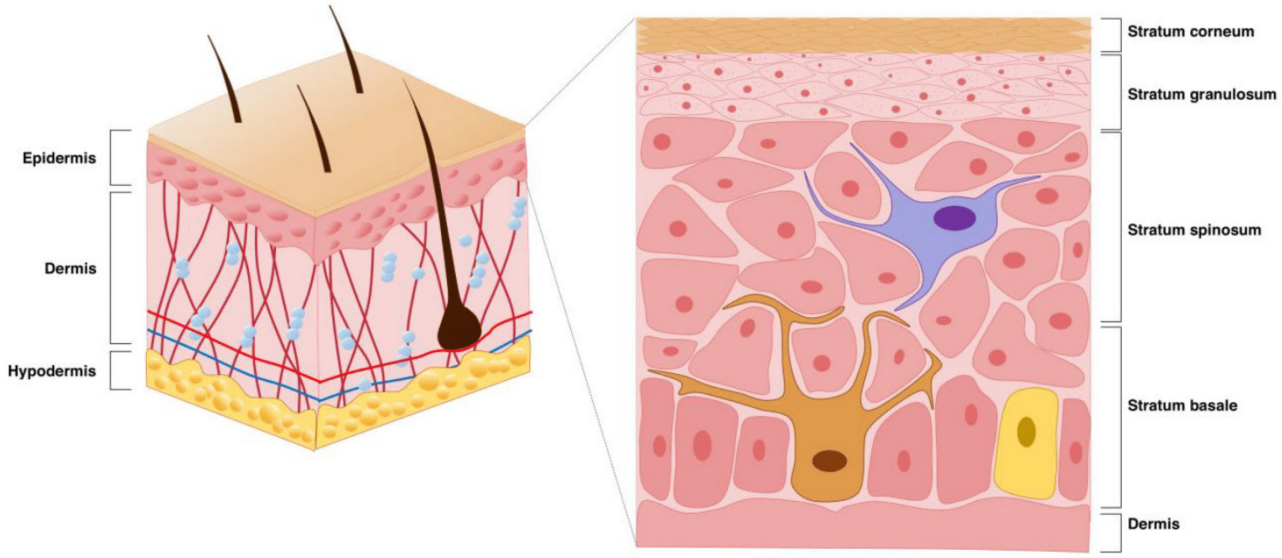
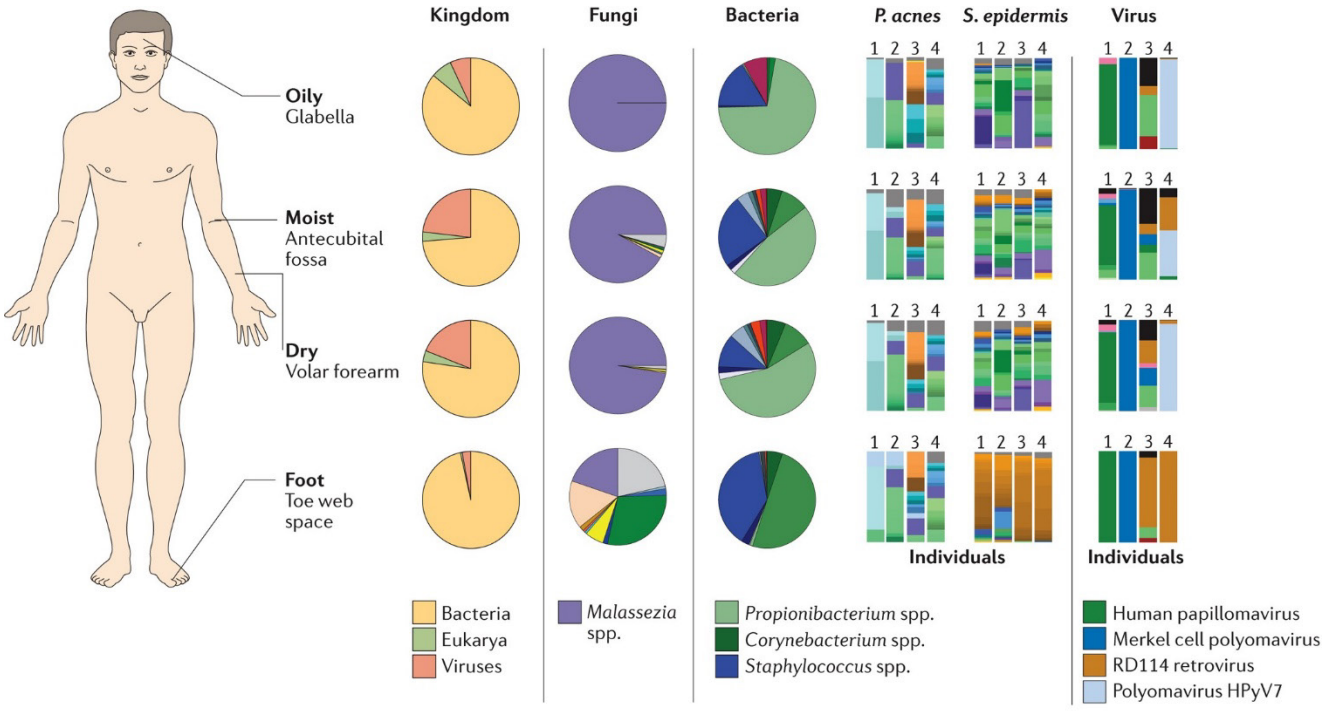


Figure 1.1 Diagram showing the layers and structures of the skin with a focus on the layers of the epidermis. *Sourced from Brito, Baek and Bin, 2024.*

Within the skin there are several appendages with their own defined roles; the glands and their secretions are of importance when considering the microbiology of the skin. There are three types of glands within the skin, the content of the secretions is different between these types. The eccrine produce sweat which is mostly water but does contain some trace ions, while the apocrine also produces sweat it includes lipids and proteins these are decomposed by microbes (Baker, 2019). Then the sebaceous glands which are usually found associated with a hair follicle produce sebum that contains antimicrobial peptides (Lousada *et al.*, 2021). These glands provide different niches and thus different bacterial populations can be found within them or their associated pores, ducts, and the hair follicle, for example *Cutibacterium acnes* (formerly *Propionibacterium*

(Scholz and Kilian, 2016) is associated with the sebaceous glands. As these glands also appear in different concentrations over the skin their secretions contribute to the different categories previously mentioned (Hwang and Baik, 1997; Best, Lieberman and Kamilar, 2019).

Previous studies have identified the different microorganisms that make up the skin microbiome (Fig. 1.2). Here bacteria are seen to dominate most areas, however in moist and dry skin sites around 20% are fungi, among these bacteria *Cutibacterium spp.* and *Staphylococcus spp.* are found as the main members, comparatively the fungi are mostly *Malassezia spp.* with the exception of the foot where more diverse fungal community is seen (Byrd, Belkaid and Segre, 2018). More generally the skin microbiome includes members from the *Actinobacteria, Bacteroidetes, Firmicutes,* and *Proteobacteria* phyla; along with eukaryotic microorganisms from the *Malassezia,* and *Aspergillus* genera, and a selection of bacteriophages (e.g. *Staphylococcus* phage's) and other viruses that target humans (e.g. Papillomavirus) either as a vector or an active infection (Wilson, 2008; Hannigan *et al.*, 2015; Graham *et al.*, 2022). Archaea such as *Thaumarchaeota* have also been found on the skin (Skowron *et al.*, 2021). This makes the skin microbiome a surprisingly complex mixture of microorganisms with diversities and community structures based off the niches created in the local environment.



Nature Reviews | Microbiology

Figure 1.2 Distribution of the different species of the skin microbiome based on skin environment. *(From Byrd, Belkaid and Segre, 2018.) Here the dominance of bacteria across all environments is displayed, but an increase in viruses occurs on dry and moist skin. Aside from the foot Malassezia spp. are the dominant fungi on the skin, similarly Cutibacterium spp. are the dominant bacteria across all environments apart from the feet.*

All the species are distributed across the various niches present within the skin; the proportions of all species present within the skin microbiome at these niches are unique to everyone. Specifically, the sebaceous glands are optimised for facultative and obligatory anaerobes as the sebum contains a source of lipids, while the

breakdown of these lipids lowers the pH inhibiting the growth of microorganisms like *Streptococcus spp.* or *Corynebacterium bovis* as such areas of the skin with high concentrations of sebaceous glands (such as the head) commonly find *Cutibacterium spp.* and *Staphylococcus epidermis*. These acidic byproducts will also help contribute to the normal function of the skin as previously mentioned. Additionally, areas such as the armpit and toe interspace are occluded regions within the skin allowing for a proportional increase in *Corynebacterium spp.* compared to other areas of the skin. Individual *Staphylococcus spp.* distribution will also vary across skin area for example *Staphylococcus hominis* can have an increased population on the legs and arm when compared to *S. epidermidis* (Wilson, 2008; Skowron *et al.*, 2021). Organisms that live on the skin will bind using adhesins to human matrix proteins that are present within the skin's extracellular matrix such as keratin and collagen, this allows them to resist mechanical sheering off the skin, however many of these species are not known for motility thus suggesting that when the extra cellular matrix is remodelled these microorganisms either rebind or are moved to other areas through passively being carried via gland secretions or through attachment to cells migrating through the matrix or other as yet unknown means (Chiller, Selkin and Murakawa, 2001; Sfriso *et al.*, 2020; Pfisterer *et al.*, 2021).

Finally the main contributor to the skin microbiome are the personal hygiene habits employed by the individual, these are generally acquired during childhood but will also be influenced by changes in environment or socioeconomic status over time (Ross, Doxey and Neufeld, 2017; Ramos-Morcillo *et al.*, 2019). The most obvious way of this having an effect is the frequency and method of performing hygiene i.e. if a person has a bath or shower and if these occur daily or weekly; if access to hygiene is restricted then an increase in the number of some species such as *Malassezia spp.* will occur (Sugita and Cho, 2015). The different types of detergents, make-up, and other personal hygiene or beauty products such as moisturisers and deodorants used will have a different impact as the skin microbiome responds and adapts to the chemical ingredients (Mim *et al.*, 2024), some of which may have antimicrobial activities or be lipid based providing nutrients to some of the species; these chemicals are known to last on the skin for several weeks (Wallen-Russell, 2018; Sfriso *et al.*, 2020). Research has shown that the continuous application of deodorants and personal beauty products causes a stable microbial community to form and that changes in diversity occur when either no or a different product was used (Callewaert *et al.*, 2014; Bouslimani *et al.*, 2019).

1.1.2. Oral Microbiome

Like the skin, the content of the oral microbiome is a result of the environment the oral cavity creates. The oral cavity is the beginning of the digestive tract and is also used for talking and breathing where it serves as part of the respiratory system, it is separated from these roles by the epiglottis in the pharynx. Aside from this the oral cavity can be used to hold objects or for intimate contact with another human. The oral cavity is composed of five main areas, the teeth, the gingivae, the cheeks, the tongue, and the palate, with all these

surfaces coated in saliva. The cheeks, palate, and gingivae are soft tissue with the gingivae being tightly bound over the jaw bone, the tongue is a muscle covered with taste receptor cells; these all have tight junctions and active blood flow that aid in preventing microbial infection with the help of a range of immune cell's signalling pathways dependent upon the environment present (Hovav, 2014; Wu *et al.*, 2014; Dutzan *et al.*, 2016; Kay, Kramer and Visser, 2019). The outside layer of the teeth is enamel; this is mainly constructed of hydroxyapatite crystals and along with the dentin underneath is avascular, this prevents microbes from penetrating but does allow for bacteria to bind and create biofilms on the surface. The saliva is produced by a number of minor and major glands distributed throughout the oral cavity, it consists mostly of water with potassium and bicarbonate ions, proteins and glycoproteins (John E. Hall and Arthur C. Guyton, 2010). Key to its function is the digestive enzyme amylase, also included are three immune factors: IgA antibodies, lysozymes, and defensins (Crawford, Taubman and Smith, 1975; Zasloff, 2002).

The oral microbiome is the second most diverse human microbiome (Giordano-Kelhoffer *et al.*, 2022), it benefits from having a warm wet environment with regular nutrient deposits and its main feature are the biofilms that form on the surfaces of the oral cavity. However regular dental hygiene will also affect the composition of the oral microbiome. Generally, the oral microbiome consists of a wide range of microorganisms from almost every phyla of bacteria along with *Candida spp.* and a selection of bacteriophages (Deo and Deshmukh, 2019). Common bacteria belong to the *Lactobacilli, Neisseria,* and *Streptococci* genera. Some of the microorganisms found in oral samples also include extremophiles like *Hydrogenophilus thermoluteolus* and as yet uncultured microorganisms given the temporary *Candidatus* genus; it however remains to be seen what role, if any, these types of microorganisms play in the oral microbiome or if they are just some form of common transient microorganisms that will frequently get detected in samples (Caselli *et al.*, 2020). Outside of bacteria the archaea *Methanobrevibacter spp.* have been isolated from dental plaques and cavities and the amoebae *Entamoeba gingivalis* is linked with periodontitis (Trim *et al.*, 2011; Sogodogo *et al.*, 2019). Within the oral cavity some genera such as *Streptococcus* are associated with all areas (Fig. 1.3), whereas others are only commonly recovered from specific locations such as *Campylobacter* being found in the subgingival dental plaque (Giordano-Kelhoffer *et al.*, 2022), demonstrating that part of the increased diversity seen in the oral cavity is driven by environmental adaptation. More specifically *Streptococcus mitis* and *Gemella haemolysans* can be found on all areas of the oral cavity, whereas *Actinomyces spp.* and *Granulicatella adiacens* can only be found on the tooth surface and subgingival area, while on the tongue *Streptococcus salivarius* can be found in abundance (Aas *et al.*, 2005; Wilson, 2008; Zhou and Li, 2020).

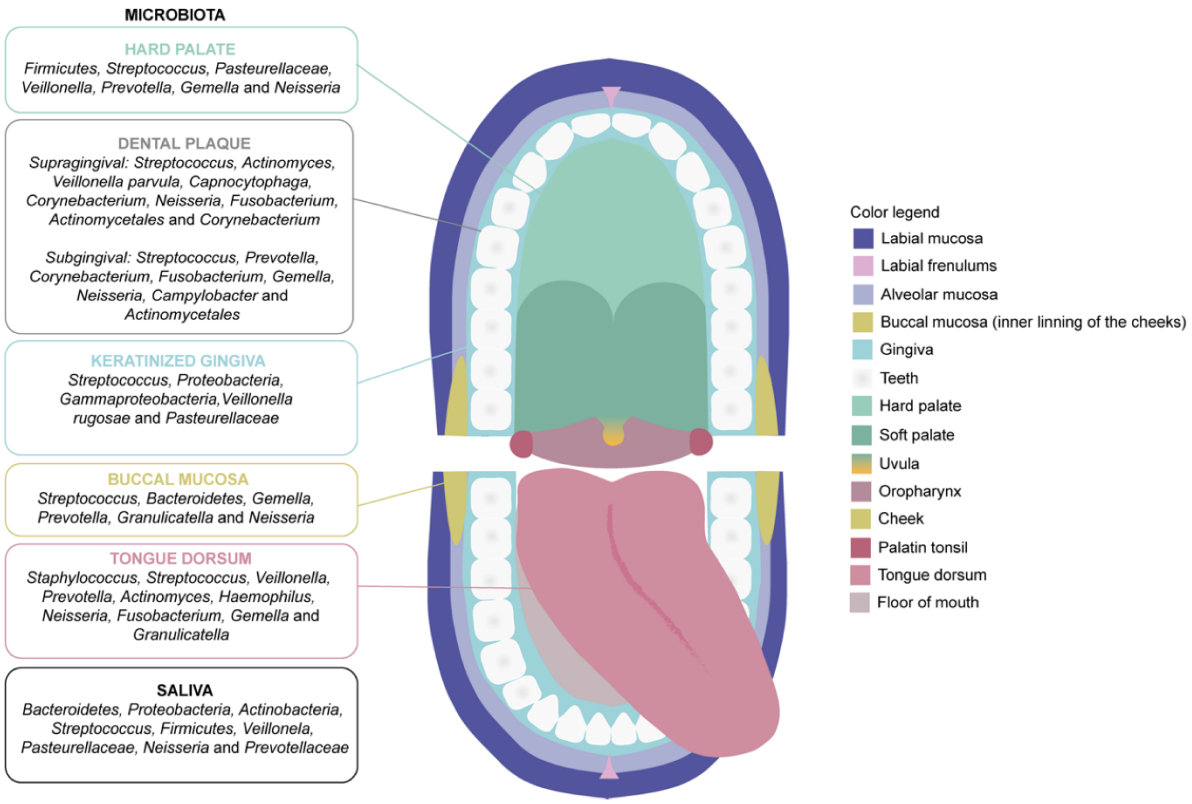


Figure 1.3. Distribution of the different species of the oral microbiome based on location in the oral cavity. *(From Giordano-Kelhoffer et al., 2022.) Highlighting the dominance of Streptococcus spp. across all environments.*

These microorganisms not in a planktonic mode of growth within the saliva are attached to surfaces as part of a biofilm. The bacteria form biofilms on available surfaces particularly the teeth, this is an ordered process with different species having different surface adhesins that bind with different effectiveness to different materials, some are helped by the molecules such as proteins present in the saliva pre coating the teeth or by the proteins present in other bacteria (Mosaddad *et al.*, 2019). Detailed research has shown that during the establishment of biofilms *Corynebacterium spp.* play a central role in the binding after *Streptococcus spp.* and *Actinomyces spp.*, acting as a structure for other species to bind to; this research also shows that microenvironments are generated within the biofilm where aerobic species such as *Neisseria spp.* that feed of the raw nutrients colonise at the external surface and an anoxic layer develops where species like *Leptotrichia spp.* and *Fusobacterium spp.* can be found (Welch *et al.*, 2016). In these forms the bacteria are capable of breaking down food and generating acidic molecules which degrade the enamel and dentin among other structures leading to dental caries.

Like the skin microbiome oral hygiene habits will influence the microbiome, oral hygiene typically features twice daily brushing of teeth with a fluoride toothpaste, this can then be accompanied by use of a mouthwash and/or the use of dental floss or interdental brushes, these mechanical actions dislodge bacteria from oral surfaces (Min *et al.*, 2024). The ingredients in toothpaste or mouthwash usually have some form of active antimicrobial action, for example the disinfectant chlorohexidine in mouthwash or the inclusion of

essential plant oils or triclosan in some toothpastes, these will further serve to disrupt the microbiome (Prasanth, 2011). Oral hygiene is also important in the prevention of dental caries and other periodontal diseases, long term this prevents teeth loss as the fluoride in toothpaste serves to catalyse the diffusion of calcium and phosphate into the tooth during the remineralisation process (Selwitz, Ismail and Pitts, 2007; Fink, 2018; Worthington *et al.*, 2019). Fluoride is also available in other forms such as water fluoridation schemes in countries such as the UK; fluoride is added to many of the water supplies or in some areas a natural level of fluoride exceeds the minimum deemed necessary to have an impact on teeth health (*Fluoride - NHS*, 2018). In recent years while the effectiveness of continued water fluoridation has been demonstrated, the effectiveness on the protection of teeth has lessened with the prevalence of fluoride toothpastes and awareness of oral hygiene (Belotti and Frazão, 2022; Iheozor-Ejiofor *et al.*, 2024).

Regular examinations by a dentist or other dental professional can be used to guide dental hygiene and act as interventions when dental caries arise. When performed correctly these procedures along with the natural forces generated while chewing and swallowing will break apart the biofilms and reduce the bacterial load, however the oral bacteria are able to recolonise the teeth quickly (Sedghi *et al.*, 2021). This helps to prevent pathogenic bacteria from colonising the mouth, evidence has shown that bacteria within the oral microbiome regularly share genes suggesting that they help each other to survive (Roberts and Kreth, 2014; Sukumar *et al.*, 2016). Socioeconomic factors and education will affect the ability to perform dental hygiene as a more affluent person will be able to purchase toothpaste and toothbrushes regularly and know how to use them appropriately (Oberoi, Sharma and Oberoi, 2016; Renson *et al.*, 2019). This will also affect the ability to attend dentist appointments and undergo treatments such as cleaning by a dental hygienist to help them maintain good oral health, this is not just reflected in oral health as several studies have identified professional oral cleaning as helping reduce the incidents of other diseases (Adachi *et al.*, 2007; Yamada *et al.*, 2018; Mitsutake *et al.*, 2023).

Research has shown that the oral microbiome plays a role in the body's nitrate cycle and thus involving the oral microbiome in cardiovascular health, the oral microbiome converts nitrate to nitrite; a key part of this process are *Veillonella spp.* Which have been found to be the most effective at nitrate reduction out of the normal oral flora (Mitsui, Saito and Harasawa, 2018). Several species have been implicated in processes that help prevent pathogenic bacteria from colonising or damage to the oral cavity. *S. salivarius* can produce a bacteriocin which inhibits Gram negative microorganisms, while *Lactococcus lactis* produces the bacteriocin nisin which has shown to be active against oral tumorigenesis (Wade, 2013; Kamarajan *et al.*, 2015; Shin *et al.*, 2016). *Porphyromonas gingivalis* has been shown to be inhibited by some *Streptococcus spp., Actinomyces spp.,* and *Bifidobacterium spp.* additionally *Streptococcus spp.* can decrease the ability of *P. gingivalis* to bind (Van Hoogmoed *et al.*, 2008; Essche *et al.*, 2013). The content of the oral microbiome may contribute to host taste responses due to the products of the bacteria altering the sensitivity to certain molecules or taste groups for example *Actinobacteria spp.* can produce bitter acids which will enhance the taste of bitter foods (Feng *et al.*, 2018; Cattaneo *et al.*,

2019). Conversely the host saliva helps control the environment and selects for certain species as the saliva contains antimicrobials such as histatins, as well as some of the macromolecules providing a nutrient source some species can work in synergy with others to break down these nutrients more completely favouring their presence in the oral microbiome; the unique composition of each person's saliva will contribute to the uniqueness of their oral microbiome due to the environment it creates (Marsh *et al.*, 2016).

1.1.3. Gut Microbiome

The human gut microbiome is the most diverse human microbiome and considered the most diverse microbiome known to exist. It covers the areas of the digestive tract from the stomach to the anus, this provides a warm environment with a constant supply of nutrients but little to no oxygen (Donaldson, Lee and Mazmanian, 2016). As such an incredible amount of research has been performed into characterising it and its role in various diseases (Abdill, Adamowicz and Blekhman, 2022). An overview is provided here to help contextualise this thesis and its work compared to the already established large body of literature, along with highlighting how there is cross-talk across the various human microbiomes (Sinha, Lin and Ferenczi, 2021; Xu *et al.*, 2025).

The acidic nature of the stomach prevents all but the most specialist organisms from inhabiting it, this allows the stomach to act as a passive part of the immune system as it will kill microorganisms caught in the mucociliary clearance system in the lungs when they are deposited by swallowing along with other transient microorganisms that enter the oral cavity. However, when food enters this raises the pH of the stomach acid making it more neutral, which enables more species to survive. Along with this pH change the ability of microorganisms including *Clostridioides spp.* and B*acillus spp.* to form spores enables these microorganisms to enter the small intestine (Waterman and Small, 1998; Clavel *et al.*, 2004; Wetzel and McBride, 2020). The small intestine is lined with intestinal villi which allow for increased surface area enabling food absorption, the motions and high concentrations of digestive factors limit population by microbes. While the large intestine does not have any permanent protrusions it is lined with absorptive cells which perform water absorption which helps convert the remaining digestive material into faeces, these conditions allow for the majority of the colonisation of the digestive tract to occur (Wilson, 2008). Unlike the skin and oral cavity, hygiene is not a factor in the composition of the gut microbiome, instead diet and any antibiotic treatments have direct impacts on composition (Guarner and Malagelada, 2003).

Common microorganisms found in the intestines belong to the *Actinobacteria, Firmicutes, Bacteroidetes, Proteobacteria, Fusobacteria,* and *Verrucomicrobia* phyla; their distribution is affected by the various niches present within the intestines, the most dominant members of this microbiome will change during age. In the stomach acid tolerant *Lactobacilli spp.* and *Streptococci spp.* are often found, along with the specialist *Helicobacter pylori* which has adapted to burrow into the stomach's mucous lining to create a more tolerable micro-environment (Amieva and El-Omar, 2008; Wilson, 2008). The appendix is believed to function as a

reservoir for the gut microbiome as its removal has been linked to changes in the microbiome content and linked to increased incidence of certain conditions (Killinger and Labrie, 2019; Vitetta, Chen and Clarke, 2019; Chen *et al.*, 2020).

Current research focusing on the interactions between the gut microbiome and the host have revealed an interplay between both covering areas such as the immune system to others like neural disorders. In relation to the immune system evidence has shown that there is decreased immune function in animals with sterile digestive tracts, along with microbial products including anti-inflammatory molecules (Rooks and Garrett, 2016; Killinger and Labrie, 2019). Research has also shown that the gut microbiome plays a role in irritable bowel syndrome as suffers often share a similar microbial components (Sabo and Dumitrascu, 2021). The gut brain axis between the host and the gut microbiome is a complex bidirectional system. Briefly, gut bacteria can produce several neuroactive molecules which can either influence local neurons or travel through to the brain where a direct effect can occur, conversely dietary choices will affect the available nutrients within the gut altering the abundance of species or the molecules they produce. These interactions within the gut brain axis are being researched in connection with long term and complex neurological disorders such as Alzheimer's disease or Depression (Zhong *et al.*, 2021). This can extend beyond the gut brain axis as there is the gut-brain-skin axis where skin diseases can be detected by the alteration of the gut microbiome from its normal state and these conditions can further affect the brain. Some of these interactions between the gut and the skin are a result of both the intestines and skin being made of epithelial cells and thus sharing communication molecules, for example inflammation in the skin system can cause inflammation of the gut (Pessemier *et al.*, 2021). Psoriasis is a skin condition caused by the interactions of the immune system, this is linked to the gut-brain-skin axis by the cross talk between the immune and nervous system; in this condition there is evidence to suggest that disruptions to the skin and gut microbiome can effect depression which is a comorbidity of psoriasis and the psoriasis itself, similarly these conditions can affect the gut and skin microbiomes, a vicious cycle of depression and psoriasis can be created by dysregulation of the gut-brain-skin axis (Chen *et al.*, 2021).

1.2. Ageing in Humans

The ageing process in humans is a continuous process that exhibits itself through a mixture of physical and cellular changes, leading to a decline in functionality. Physically, in humans surface ageing is associated with the acquisition of wrinkles along with other sagging skin, and an increase in frailty which includes the increase in exhaustion, overall weakness, and physical inactivity from exhaustion and decreased muscle mass (Kim and Rockwood, 2024). On a cellular level ageing is defined by nine hallmarks, genomic instability, telomere attrition, epigenetic alterations, loss of proteostasis, deregulated nutrient sensing, mitochondrial dysfunction, cellular senescence, stem cell exhaustion, and altered intercellular communication (López-Otín *et al.*, 2013). The causes can be categorised into extrinsic and intrinsic factors, the extrinsic are generally caused by lifestyle choices made

during a lifetime, while the intrinsic are genetic in nature; however, both intrinsic and extrinsic factors can cause physical and cellular changes.

1.2.1. Skin Ageing

The skin is where some of the most obvious signs of physical ageing occur, this is a multifactorial process that leads to the characteristic signs observed; including wrinkles, thinning of the epidermis, acquisition of age spots, becoming dryer (xerosis), reduction in wound healing, reduced functionality as a barrier, and an increase in acquired infections. Skin cell senescence contributes to some of the reduction in function experienced by several of the systems of the skin. Decreasing keratinocyte migration, function, and communication with fibroblasts contributes to a thinning of the epidermis, along with the reduction in wound healing; this also leads to less immune cells within the skin causing an impaired innate immune response (Choi, 2019; Fernandez-Flores and Saeb-Lima, 2019; Ho and Dreesen, 2021). Reduction in all varieties of lipid found within the skin occurs during ageing along with the natural fluctuation in lipid level that occurs within the seasons of the year; this reduction in lipids can be linked to the changes in sebum production during ageing, along with changes in the structures created by these lipids altering the feel of the skin during ageing. The reduction in the blood vessels and deterioration in neuromuscular junctions, cause skin sagging, lack of nutrient supply and limited repair functions leading to increasing skin frailty. The reduction in oestrogen experienced by women undergoing the menopause can accelerate ageing within the skin. Reduced cell function causes the synthesis of collagen and other fibres to become impaired also contributing to skin frailty and thinning. As the number and function of the glands is reduced this leads to an impact on thermoregulation and will alter the niches present within the skin (Rittie and Fisher, 2015; Ho and Dreesen, 2021; Lee, Hong and Kim, 2021; Ahlers *et al.*, 2022; Franco, Aveleira and Cavadas, 2022).

Extrinsically skin ageing is affected by decisions and activities of each person, these extrinsic factors can compound onto the intrinsic factors, making the ageing more pronounced. Prolonged UV radiation exposure leads to an increase in cell number however these are abnormal in form due to it causing DNA damage, it also triggers a higher turnover in skin fibres and a reduction in collagen; in ageing individuals the reduced vascular presence within the skin exacerbates these elements. UV exposure also initiates some of the pigmentation changes that occur. Exposure to air pollutants leads to an increase in oxidative stressors within skin cells, increasing the damage to the cells, leading to diseases such as atopic dermatitis and cancer (Puri *et al.*, 2017). In smokers' skin similar responses occur to UV damaged skin, except fibres have altered shapes which impairs skin structure (Makrantonaki *et al.*, 2013; Kuehne *et al.*, 2017; Fernandez-Flores and Saeb-Lima, 2019). The effects of these extrinsic factors can also be affected by socioeconomic status and education, for example populations that are educated about UV radiation and have access to sun cream may use it to protect themselves against harmful UV exposure thus experiencing less of an impact in older age (Zell *et al.*, 2008; Hausauer *et al.*, 2011).

These changes will affect the structure of the skin microbiome. Research has shown an increase in microorganisms such as *Pseudomonas aeruginosa* and yeasts occur within the elderly, along with a general increase in microorganism species diversity (Laube, 2004). The increase in frailty caused by ageing leads to an inability of the host to perform hygiene properly and remaining immobile for prolonged periods will increase the development of skin infections through pressure sores. Reductions in *Cutibacterium spp.* and *Staphylococcus aureus* in comparison to younger adults have been found, it has been theorised that this reduction in *Cutibacterium spp.* is due to the decrease in sebum that occurs in older age (Luna, 2020). Microbes which are more commonly found within the oral cavity have been found on the skin of elderly Japanese women, possibly due to translocation this could be due to poorer hygiene or other age linked difficulties, the exact source is not clear and requires further investigations (Shibagaki *et al.*, 2017). A study into the skin microbiome of Korean women has demonstrated that there is a reduction in interactions and metabolic pathways that help maintain the stability of the skin (Kim *et al.*, 2022). An increase in Archaea abundance on the skin has been noted in those over 60 when compared to middle aged, this suggests that the reduction in sebum production can influence more than the bacteria in the human microbiome (Moissl-Eichinger *et al.*, 2017). The reduction in the immune system function enables different bacteria to inhabit the lower layers of the skin altering the population of the microbiome; possibly leading to an increased risk of infection when coupled with skin frailty making it more prone to damage such as pressure wounds and thus provides more opportunities (Nagase *et al.*, 2020). Some reports suggest that the skin microbiome may exacerbate ageing by producing cell stressors and causing DNA damage to cells (Bana and Cabreiro, 2019). Along with studies that suggest the microbiome may help reduce the effects of UV exposure or trigger immune responses when pathogenic species are present (Nguyen and Soulika, 2019).

1.2.2. Oral Ageing

While the skin shows the most obvious signs of physical ageing, the oral cavity can experience some of the most drastic changes during a lifetime particularly during old age. A mixture of intrinsic and extrinsic factors lead to these changes occurring. Intrinsically the factors of ageing are similar to the skin, with a loss of epithelial thickness this effects the gingiva, the mucous membranes, and the tongue. The output of the salivary glands reduces, along with chemical content changes can lead to a dry mouth (xerostomia), or difficulty swallowing in older age. Additionally, tooth enamel becomes brittle, along with a reduction in dental pulp which is accompanied with dental avascularity, outside of the tooth the cementum between the teeth and gingiva begins to experience irregularities. Chewing problems arise from a loss of thickness and performance from the masticatory muscles; similarly, a reduction in function of throat muscles can combine with the reduction in saliva production to make swallowing more difficult. Finally, the changes to the salivary content and general avascularity that occurs with old age reduces the effectiveness of the immune system in the oral cavity and leads to a delayed healing response (Thompson and Chen, 2021; Houge and Ruiz, 2022).

The extrinsic factors that contribute to an ageing mouth are a mixture of personal lifestyle choices and decisions made due to dental treatments. The biggest factors are hygiene and diet, a diet with a high proportion of sugary foods and an improper hygiene routine may lead to an increase in incidents of dental caries and periodontitis. The standard treatment for caries is the removal of the damaged dental material and replacement with a dental amalgam or similar material. When dental caries become severe enough then other treatments such as crowns, or root canals are employed; each of these methods involves the replacement of significant portions of dental material with an artificial material. The final treatment option is extraction which can see the tooth replaced with an implant or bridge work, in cases of complete or near complete edentulism a denture can be used (Selwitz, Ismail and Pitts, 2007; Houge and Ruiz, 2022). Aside from caries diet will also contribute to a general wearing of the enamel over time, particularly the occlusal surface of the tooth which is subject to the most grinding against other teeth (Lussi, Schaffner and Jaeggi, 2007). Drinking alcohol will contribute to increased dental erosion, oral cancer risk, and oral lesions, while smoking will see a similar increase in oral cancer risk and gum disease (Agnihotri and Gaur, 2014; Priyanka, 2017).

In an ageing oral cavity, the microbiome can undergo changes representative of its more diverse nature. The changes within the saliva will lead to the structure of oral biofilms changing; along with this the reduction in immune molecules and immune cell function will also enable pathogenic bacteria to colonise the mouth more often, particularly as infections can be more frequent and last for longer. The inhabitants of the subgingival tissues have been found to spread to the gut in older people (Iwauchi *et al.*, 2019). There has been research suggesting that in Alzheimer's the inflammation response is triggered by oral microbiome changes, but the mechanisms for this are still being elucidated (Sureda *et al.*, 2020; Maitre *et al.*, 2021; Yang *et al.*, 2021). Similarly links between the oral microbiome and cardiovascular disease are also being explored, which is expanding upon the knowledge of the role the gut microbiome in this condition (Tonelli, Lumngwena and Ntusi, 2023). If edentulism (the loss of teeth), has occurred, then the content of the oral microbiome can resemble that of a child prior to dental eruption. In these cases if dentures are used increase in bacteria and fungi such as *Candida* can occur, as they will spend some of their time outside the mouth and maybe subject to poor hygiene, denture surfaces are also different than natural teeth so biofilm formation will be different (Cho *et al.*, 2014; Delaney *et al.*, 2019; Wenger *et al.*, 2021). During old age an increase in chronic long-term conditions combine into comorbidities, some such as Barrett's oesophagus are linked to a change in oral microbiome structure, the oral microbiome is also hypothesised to be a causative factor in Osteonecrosis, finally while a change is known in diabetes there is no clear causative relationship (Ruggiero *et al.*, 2018; Snider *et al.*, 2018; Almeida-Santos *et al.*, 2021). Due to an increase in infections associated with old age antibiotics are more frequently taken, currently this has not been linked directly with an increase in antimicrobial resistant species within the oral microbiome but it does warrant further investigation (De Sousa *et al.*, 2020; Stewart *et al.*, 2021). These examples all highlight

the complex role the oral microbiome has in not just infections but in other chronic diseases which have an impact on healthy ageing.

1.2.3. Gut Ageing

Unlike the skin and the mouth, the gut does not express the external signs of ageing, but it still undergoes changes with age which are summarised here to give context and comparison to the other aspects of this thesis. Briefly intrinsically it experiences a thinning of the epithelial layer, a reduction in blood flow leads to a reduction in nutrient absorption function, and an impaired immune response. Beyond diet and infection extrinsically the gut is not exposed to many stressors, however medical interventions for conditions such as appendicitis or *Clostridioides difficile* infection can be as extreme as the removal of significant portions of the gut tissue or associated systems. As previously mentioned, microbes from the oral microbiome can be found in the gut in older age. Aside from this the microbiome can be implicated in an increase in gut inflammation which will contribute to the reduction in gut function. The gut microbiome can improve insulin resistance a precursor to type 2 diabetes which often arises in older age (Morley, 2007; Piggott and Tuddenham, 2020). An increased number of medications taken in older age such as blood thinners and pain relief, the majority of which are absorbed through the gut, these active pharmaceutical agents can be altered by the gut microbiome, potentially making them inactive or toxic (Flowers, Bhat and Lee, 2020), this is detrimental to the host and can lead to increased risk of long term conditions such as diabetes or outright harming the host gut (Sarkar and Sil, 2021; Dikeocha *et al.*, 2022).

1.3.Sampling Methods for the Skin and Oral Microbiome

When investigating any microbiome, the effectiveness of a sampling method can be crucial in the quality and reproducibility of the results obtained (Bjerre *et al.*, 2019; Kool *et al.*, 2023). Currently no standardised approach for sampling the skin and oral microbiome has been agreed upon withing the scientific community (Lamaudiere and Sergaki, 2024), this is in part due to the different requirements and goals of the studies and research being conducted. Also, many outside factors must be considered, such as if the preservation of anaerobic microorganisms is prioritised, or if the collection of metadata from the participants is required. Some techniques can be invasive and require a medical practitioner to perform safely (Fig. 1.4). Before obtaining microbiome samples from participants within these studies, ethical approval for the research study must be granted by an appropriate body and informed consent obtained from the participants in accordance with the local laws and regulations.

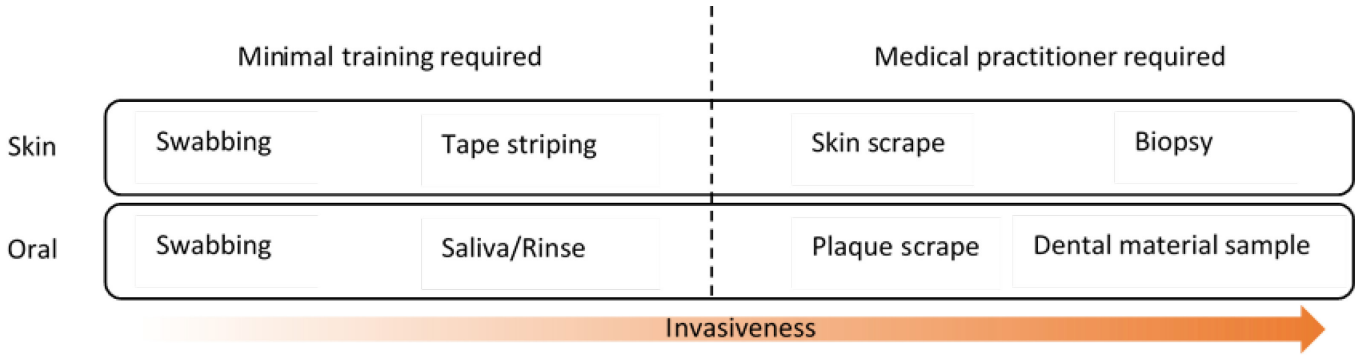


Figure 1.4 A representation of methods for sampling the human skin and oral microbiome. *Detailing the invasiveness and ease of performing.*

1.3.1. Skin Sampling

For obtaining skin microbiome samples from humans there are four methods with their own advantages and disadvantages, however they all share the key need to maintain a repeatable sampling area of the same size and skin region to keep the niches similar (Pistone *et al.*, 2021). The simplest method is swabbing the skin, this is non-invasive and only needs to have the designated area sampled and the sampling technique to be controlled. Briefly, the method comprises a sterile swab being rubbed against the area of the skin being sampled, usually the swab has been moistened in a suitable buffer to assist in the stabilisation of the microorganisms or enable their effective removal, and for standardisation the swabbing will be done in a repeated pattern (Grice *et al.*, 2008, 2009). However, as it is only sampling the surface layer of skin it will not detect any microorganisms living in the deeper layers (Cundell, 2018), but commercial kits are readily available for skin swabbing.

Tape stripping is another method and is relatively non-invasive, a sterile piece of sticky tape is attached to the skin for a set period of time before its removal and subsequent investigation for microorganisms. This has the advantage of removing some of the upper skin layer allowing for some deeper microorganisms to be detected, the tape can also be cut to the same size to enable easy control of sample area. This technique does have more disadvantages than swabbing, as the tape needs to be sterilised prior to use, and no kits are currently commercially available meaning more time would be required to set it up and process the samples afterwards (Berkowitz *et al.*, 1974). Also, some participants may find the removal of the tape uncomfortable. Another drawback is removing the bacteria from the tape, this would either involve washing the tape in a sterile buffered solution to dislodge the microbes for further investigations or adhering and removing the tape on a suitable agar plate. However, it would be difficult to ensure that all the microbes have been removed from the tape, or that none of the chemicals in the adhesive in combination with the tape material has been affecting the viability or enhancing growth of any of the microbes removed during transit or sampling (Marples and Kligman, 1969; Ogai *et al.*, 2021; Pistone *et al.*, 2021).

To go deeper into skin sampling a scraping method may be employed, this would involve the use of a scalpel or similar object to remove or disturb the top layers of the skin which could then be processed, this method has shown effectiveness in detecting the fungal content of the skin. However, this is a potentially invasive procedure and would need to be performed by a trained professional if a scalpel is used, sampling sites would be limited to non-sensitive and non-functional areas. Handling of the sample can also be complex since the top layers of the skin when removed may not be cohesive, meaning that all fragments need to be collected into a sample container. A less invasive adaption to this method is where the sample area is disrupted using a hard blunt object like a plate spreader, as no or little skin would be removed the area would need to be flooded with a sterile solution to contain and extract the lifted material; for this the sample area would also have to be contained to prevent run off (Grice *et al.*, 2008; Boulesnane *et al.*, 2020; Lagler *et al.*, 2021; Pistone *et al.*, 2021; Ruuskanen *et al.*, 2022). Again, no commercial kits are available for this procedure.

Finally, a skin biopsy technique can be used. This would involve taking punch holes of the skin, as such this would allow for the detection of any microorganism living on any level within the skin. However, such a procedure is extremely invasive and harmful, it would need to be performed in a clinical environment by trained professionals and a local anaesthetic maybe required for the comfort of the participant. This decreases the number and frequency of sampling possible along with reducing the likelihood of recruiting participants and increasing the costs of the study. It does also generate a large amount of human biomass which can interfere with the microbe extraction process (Grice *et al.*, 2008; Pistone *et al.*, 2021; Ruuskanen *et al.*, 2022).

1.3.2. Oral Sampling

To sample the oral microbiome there are generally two main methods that can be employed, again both with advantages and disadvantages. The first method is themed around sampling the saliva. One form is taking a pure saliva sample where a few millilitres of saliva is expectorated into a sample tube, chewing on a sterile piece of paraffin wax or similar maybe employed to help generate enough volume. The saliva method helps provide a “fingerprint” of the planktonic microorganisms within the oral cavity and is completely non-invasive; having a saliva sample also allows for its non-microbial chemical content to be analysed, this can be used to inform on the potential formation of biofilms or there molecular output. Another technique of sampling saliva is the rinse method where participants rinse their mouths with either a sterile buffered solution or commercial mouthwash and then expectorating the sample into a sample tube. This method has the potential to dislodge some biofilm content that maybe present within the oral cavity. This method does have some drawbacks as commercial mouthwash may contain substances that will damage cells within the sample and their contents, and any rinsing will have to be standardised by the participants (Brookes *et al.*, 2021). Both of these methods share disadvantages as they can be mechanically difficult to perform by some subject participants, as they may lack the muscle control to swill or chew properly along with other factors such as dry mouth that could limit the

sample volume (Ogawa *et al.*, 2017; Verma, Garg and Dubey, 2018; Nearing *et al.*, 2020; Adami, 2021; Zaura *et al.*, 2021).

The second method is to perform an oral swab, this technique has the ability to pick up microbes bound within biofilms and content can be isolated to the area sampled. Unlike the skin swabbing method, it uses a dry swab, but placing the swab in a buffer solution after sampling is sometimes used to ensure microbial viability during transport or prior to processing. It is however an invasive method, but it does allow for quick replicates and can be performed by a trained person thus standardising the sampling technique and removing any difficulties in producing a set volume (Aas *et al.*, 2005; Iwauchi *et al.*, 2019; Adami, 2021; Zaura *et al.*, 2021). Both these sampling methods can be performed with commercially available kits allowing for ease of use, setup, and costings.

Other oral sampling methods include plaque sampling, dental calculus sampling, and sampling from damaged dental material. These have advantages in their specificity and potential to be used to understand dental caries, however they require a trained dental professional to obtain the sample and can suffer from difficulties in obtaining a suitable sample volume. Overall, a key factor when sampling the oral cavity is to consider the time since the participant last ate or drank, as this may increase the presence of transient microorganisms within the sample (Adami, 2021; Zaura *et al.*, 2021).

1.4. Study Design and Ethics

When designing this type of relatively larger scale study many factors and limitations must be considered, some are universal to all research projects while others are specific (Qian *et al.*, 2020). The first main factor is the number of participants to be recruited, there must be sufficient participants to give the study statistical power for its main research aims but this can be limited by factors including the budget and study timescale along with the capabilities of the researcher team involved. In addition, the ethical appropriateness of the nature of the research study and burden placed on participants must be considered. For ethical considerations the study needs to be justified in its methods and rationale, ensuring that it will not cause harm (unless required) to participants and that the research requires participant input to be carried out. The burdens placed on the participant vary from time required for participation in the study, the information or activities required of the participants, to the taking of clinical samples if required. For some burdens compensation can be considered an appropriate mitigating measure, while for others clearly explaining the reasoning behind its inclusion in the project to the participant is acceptable (Campbell and Walters, 2014; Schulz and Grimes, 2019).

Additionally, when trying to create a cohort panel of participants for research advertising must be considered, unless an already established cohort is being used. Project advertising should be eye catching and simple enough to quickly inform the potential participant but must also be detailed enough to adequately explain the project without misleading them. Following this, the location of advertising will also have an impact upon

recruitment, physically distributing posters or fliers etc. allows for the recruitment of participants within certain geographies; meanwhile using social media and other online platforms can allow for a wider audience, but this can hamper recruitment of older or less technologically literate participants (Holden *et al.*, 1993; Hapca *et al.*, 2014; Burke-Garcia and Mathew, 2017; Akers and Gordon, 2018). It is also possible to send out advertising to already established cohorts who have given permission to be contacted.

Finally of ethical consideration when studying the human microbiomes are the gathering of metadata. As the human microbiome is complex, research studies into this field need to gather other relevant data from participants to enable a clear picture of which factors are contributing to the microbiome composition detected. These data need to be considered for appropriateness to the study, i.e. are they relevant to understanding the microbiome or are they to intrusive or irrelevant. Similarly, how these data are gathered factors into this, some factors can be collected through questionnaires and interviews, while others require some form of physical or clinical assessment which a participant may find intrusive. All these factors will have to be considered and then demonstrated to the appropriate ethics committee prior to approval, after this then the study can begin.

1.5. Quantification of the Microbiome

Once samples and any applicable metadata have been obtained from the informed participants the investigations into the microbiome can begin, these can either be culture-based or genomic. As some organisms cannot be cultured most studies prefer to utilise genomic methods unless they are investigating specific species which can be isolated in culture (Hiergeist *et al.*, 2015; Rimoldi *et al.*, 2023). For a genomic investigations' a key initial decision is down to whether to use 16S sequencing or whole genome sequencing, both have various drawbacks and strengths in their use, such as taxonomy resolution and microbial coverage, due to cost reasons 16S sequencing has remained popular until relatively recently (Jovel *et al.*, 2016; Durazzi *et al.*, 2021).

The 16S gene encodes for the RNA sequence used as part of the small ribosomal subunit in the bacterial ribosome (Schluenzen *et al.*, 2000). This key functionality in bacterial life makes it an ideal target for sequencing, as every bacterium has 16S genes. The 16S sequence is broken into consistent regions and variable regions, these variable regions differ between bacterial species. Therefore, it is possible to classify unknown sequences in a sample with a database of known sequences to compare them to (Patel, 2001). In microbiome projects these classified sequences are then bundled with the abundance data generated from sequencing and processing, these can then be analysed using a variety of metrics to investigate the questions of the study being performed.

The 16S sequences must be processed into a final dataset, some researchers outsource this to the sequencing company or another industry dry lab, whereas others will prefer to perform processing in house. Regardless of the processing body the sequences will be processed in a manner like that outlined in Figure 1.5. Here the first stage is the check of the Phred score quality information, which measures the uncertainty of the correct base being reported by the sequencing tool (Ewing and Green, 1998; O'Rawe, Ferson and Lyon, 2015), researchers will have a threshold for minimal acceptable uncertainty in their project. Based on this information tools will then be used to remove any sequences below the threshold, along with trimming any primers or adaptor sequences left over from the sequencing process which are non-informative sequence sections that may have ambiguous mismatch's that interfere with downstream processes (Sipos *et al.*, 2007). Additionally, researchers may also remove sequences below a certain length at this stage. Having generated a set of higher quality sequences they are then denoised (i.e. ambiguous bases left over from the sequencing process are clarified), and where applicable are merged into a consensus sequence, as a last quality check using only the sequence information chimeric reads are also removed (Ju and Zhang, 2015). During this process sequences are also dereplicated, saving downstream computational memory and processing time by simply recording the number of occurrences of each unique sequence rather than having a copy for each sequence in the final dataset.

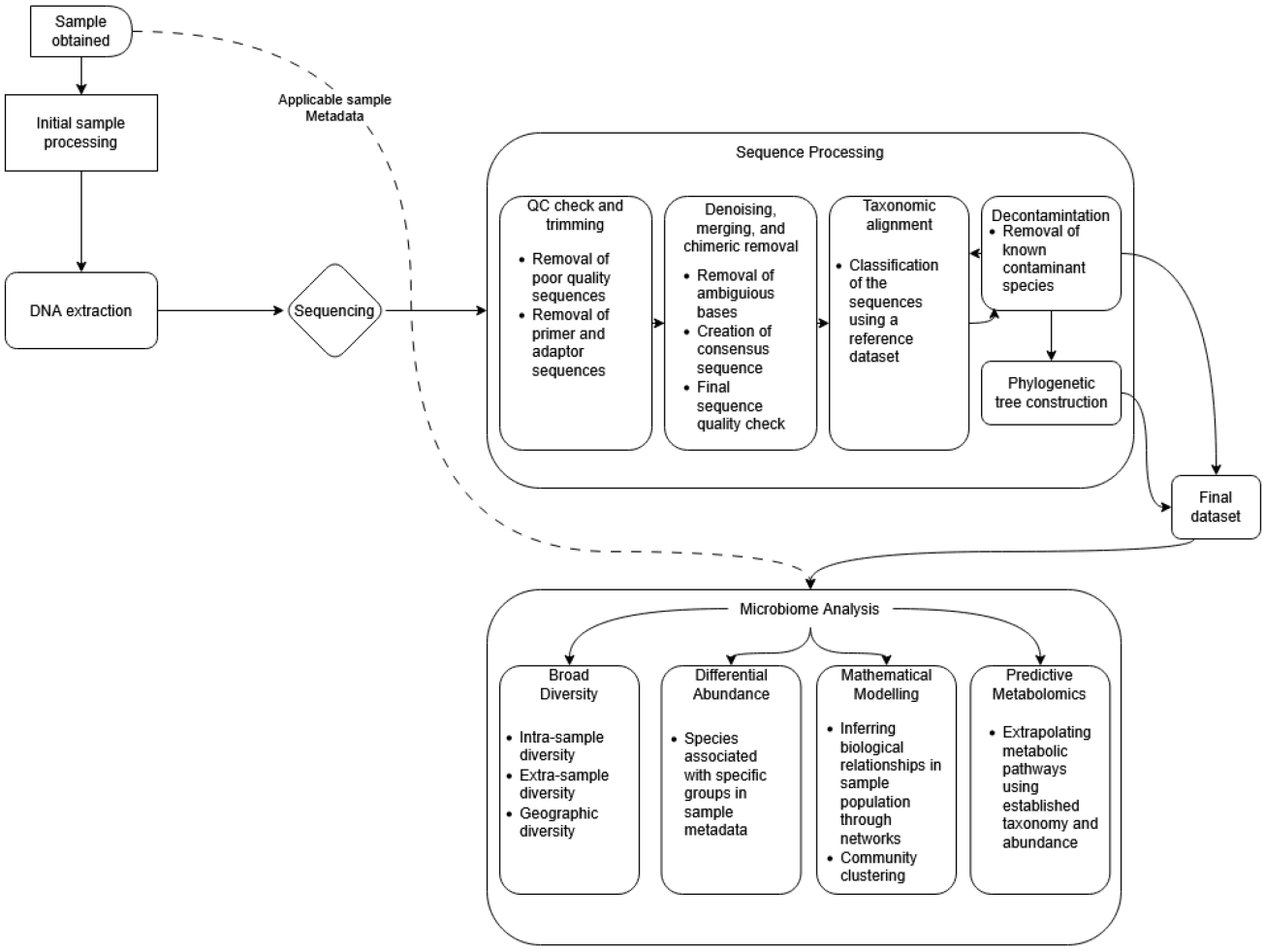


Figure 1.5 Flowchart outlining the typical stages in a 16S bioinformatics pipeline used to study the microbiome.

Having generated a dataset of sequences from the raw sequencing output they can now be classified taxonomically. Here a reference database will be processed to construct a classifier, key to this is trimming the reference reads in the database with the same primers of the region amplified in sequencing (Church *et al.*, 2020). Once the sequences have been classified then the process of decontamination can begin (if required by the study), here known contaminant species such as mitochondria, chloroplasts, or other human derived sequences are removed, along with low occurrence sequences or whatever criteria are considered contaminants to the research (Salter *et al.*, 2014; Walker *et al.*, 2020). This decontamination can take place over several stages, during which the dataset may need re-classification depending upon the limitations of the software used. After decontamination has created a finalised dataset of sequences the further analysis may require the use of a phylogeny tree of the dataset, here the sequences are aligned and a hierarchy developed, this completes the construction of the dataset suitable for analysis.

To analyse the microbiome using a 16S dataset various approaches can be taken, not all will be used depending upon the questions being asked by the study or the computational expertise, again these are outlined and clustered in Figure 1.5. Additionally, most of these methods are the same as those used in environmental ecology when dealing with macroorganisms, as such the term sample is used in place of ecosystem here. An initial broad diversity overview can be generated starting with alpha diversity, which is a measure of intra-sample diversity (Galloway-Peña and Hanson, 2020). Alpha diversity has various metrics available to measure it, some of these are a mixture of: species richness which is the total number of species present, species evenness which is the equalness of the species present, and species dominance which is focused on the species most dominant within a sample. Figure 1.6.A, uses colours to represent species, in doing so it demonstrates how high species richness can result in low species evenness and how low richness can lead to high evenness. This also shows that alpha diversity is limited by only focusing on intra-diversity, as here two samples have the same diversity score but different species. Beyond alpha diversity is beta diversity, which is the extra-sample diversity, simply this compares the presence of the species across samples. Again, there are several metrics that can be used to measure beta diversity, these tend to use both species presence and abundance, although some also utilise the phylogeny tree. Figure 1.6.B, uses colours to represent species, in doing so it demonstrates how more complex beta diversity is than alpha. Here each sample can only be considered when compared to the other samples, i.e. there is no one single diversity number to assign, this means that samples are considered in terms of their similarity, through statistical analysis similar samples can be clustered together and considered distinct from other samples. Finally of the broad diversity metrics there is gamma diversity, which is a measure of diversity

changes over geographical gradients. This measure is only possible with multiple samples from the same participant over different areas of the body, thus it is not always applicable in microbiome studies.

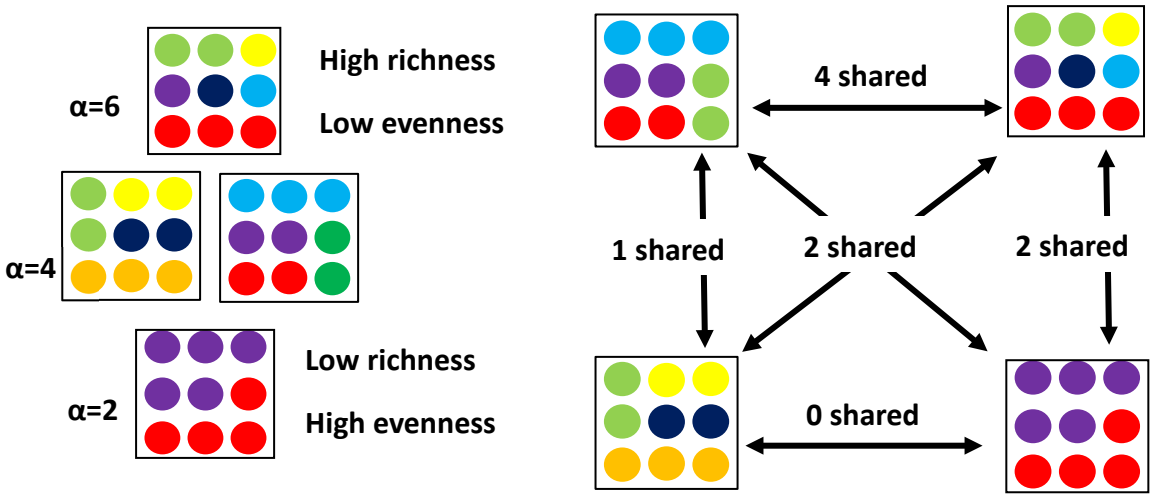


Figure 1.6 A representation of alpha and beta diversity. *Using coloured dots in the place of species and boxes as ecosystems. A. Alpha diversity, as the ecosystems ascend from the bottom alpha diversity increases along with species richness at the cost of evenness. B. Beta diversity demonstrated here is the cross-species comparison present in beta diversity, the bottom two ecosystems are highly dissimilar while the top two are very similar.*

Differential abundance analysis is a form of mathematical modelling which attempts to identify species associated with a certain condition of the sample metadata, these species can then be subject to further investigation or may become candidates for biomarker organisms. There are multiple methods used in differential abundance, however, there is disagreement around the most appropriate method to use and evidence of different results generated from the same dataset (Cappellato, Baruzzo and Camillo, 2022). Therefore, studies will use multiple methods and then filter through the results to find common results.

Beyond identifying specific species within groups, it is possible to infer the community relationships within samples or sample groups, this is achieved by constructing and analysing co-occurrence networks. In these networks each node represents a species in the microbiome and the edges between them are the association of the species occurring in the same sample group. Therefore, those species not connected are not likely to occur in the same sample group, this could be due to the species only being present in a small number of samples. Similarly, not every connected node is connected in a main network, it is possible for several smaller networks to be constructed from one dataset. Through analysing these networks it is possible to establish if a group has a stable microbiome capable of acting as a community to share resources and protect from invasive species (B. Guo *et al.*, 2022; Kajihara and Hynson, 2024). Another method for identifying communities in sample groups is Dirichlet Multinomial Mixtures, this generates community types which samples can be organised into, along with identifying abundant species in each group (Holmes, Harris and Quince, 2012).

Finally, it is possible to move beyond the taxonomic information and infer the metabolic processes present in the microbiome sample. Here the taxonomic information is used in conjunction with metabolic

databases to identify known pathways present for each taxon, these are then combined with the abundance data to generate pathway abundance tables for each sample. These tables can then be analysed in a similar manner to the sequence abundance data, to identify differentially abundant pathways associated with certain sample metadata groups. This allows for an understanding of the microbiome metabolites being released into the host environment and inferences of the effect they may have on the host made.

Comparatively whole genome sequencing as the name suggests will sequence all the DNA within a sample, this is achieved by creating short fragments of the sample DNA which are sequenced and then the sequences are reassembled computationally. Some of the stages to create a final dataset for analysis are similar with those in 16S sequencing, for example to taxonomically classify each assembled genome a comparison to a reference database is still required, for which there are several tools available along with discussion within the community about the length of sequence used for classification (Tran and Phan, 2020; Kutuzova *et al.*, 2024; Defazio *et al.*, 2025). Once the final dataset has been created then similar ecological analyses carried out on a 16S dataset can also be used such as alpha and beta diversity metrics and differential abundance investigations. However, with the increased taxonomic resolution and gene information that whole genome sequencing provides additional investigations can be carried out, including accurate assessment of microbiome functional profile, modelling of shared genes within a personal microbiome, and strain related analysis allowing for microbiome evolution along with bacterial transmission tracking (Pérez-Cobas, Gomez-Valero and Buchrieser, 2020; Kim *et al.*, 2024). These additional benefits come at extra costs, including the increased financial cost of sequencing there is the additional computational requirements for processing the data compared to 16S sequencing (Aminu *et al.*, 2024). This is due to the increased complexity of the data that can lead to contamination issues from false positives (be they single genes or whole species) caused by dataset assembly (Goig *et al.*, 2020; Li, Jiang and Dai, 2022; Sun *et al.*, 2023), or the additional contamination from the human host; as the sequencing process sequences all DNA within a sample then any human derived sequences can risk dominating the sample, this also raises ethical concerns around the consent of the participant (Chrisman *et al.*, 2022), comparatively 16S sequencing will only sequence mitochondrial DNA and any short sections that resemble the primer sequences from the participants genome. Nonetheless, despite these risks whole genome sequencing for the microbiome is still pursued due to the increase in data produced when compared to 16S sequencing (Ranjan *et al.*, 2016; Lewis *et al.*, 2021; Bars-Cortina *et al.*, 2024) along overcoming other 16S weaknesses, but 16S sequencing studies into the microbiome are long established with large resources available and produce data that correlates with other techniques (Gotschlich, Colbert and Gill, 2019; Malla *et al.*, 2019), and therefore can still be considered a valid method, particularly in small scale studies with limited budgets and resources (Muhamad Rizal *et al.*, 2020).

1.6. Aims and Objectives

The main aim of this thesis has been to characterise the diversity, composition, and metabolome of the skin and oral microbiome of humans at different ages using a UK population. This thesis also aimed to identify additional driving factors that may affect these microbiomes during ageing. These were achieved by implementing the following objectives:

- Identifying an optimal sampling method for each microbiome suitable for deployment at Aston University (Chapter 2).
- Analysing the various metadata factors gathered from the participants during the sampling phase (Chapter 3).
- Developing a bioinformatics pipeline to process the raw sequences into a dataset suitable for analysis (Chapter 4).
- Using computational tools to characterise the skin microbiome in terms of diversity, composition, community structurer, and metabolome, through the lens of ageing and the other identified metadata groups (Chapter 5).
- Using computational tools to characterise the oral microbiome in terms of diversity, composition, community structurer, and metabolome, through the lens of ageing and the other identified metadata groups (Chapter 6).

Chapter 2. Preliminary Investigations in Sampling and Culturing Participant Microbiomes

2.1.Introduction

To address the central theme of this thesis samples of the skin and oral cavity will need to be obtained, so they can be analysed for their microbiome content. As outlined in section 1.3 of the introduction there are multiple methods that can be used to sample these areas, each of which has their own advantages and disadvantages. Ideally the sampling method used needs to balance species recovery with practicality, while some methods may be high yield they may also involve a clinician and cause discomfort to the participants. Therefore, for this thesis to be conducted a suitable sampling method for the skin and oral cavity must be established.

This method must be practical to be conducted at Aston University. As this study is not being carried out in conjunction with any health care facility this limits the invasiveness of the sampling methods, similarly invasive methods may reduce participant recruitment figures. Two methods were chosen for comparison for the skin and oral cavity. To assess the yield recovered viable counts were chosen for convenience as viable count data is comparative to molecular analysis such as PCR (Ilha *et al.*, 2016), along with saving money and time. As this stage of the study is preliminary and only quantitatively comparing yield, only a small number of participants is needed, and no additional information is required from them.

2.1.1.. Aims

- Obtain ethical permission to undertake the study.
- Perform a small-scale study to compare and discover suitable sampling methods to use in the main study.
- Use the small-scale study to determine which microorganisms can be recovered from the oral and skin microbiome using standard culture techniques.

2.2. Methods

2.2.1. Ethical Approval

Ethical approval for the project and this pilot study phase was granted by Aston University Health and Life Sciences ethics committee in 2022 (HLS21008).

2.2.2. Small-scale study

2.2.2.1. Participants

For this stage of the study 10 participants were recruited with minimal restrictions beyond being over the age of 18 and being able to give free and informed consent. Participants were asked to refrain from eating for at least 30 minutes prior to giving the oral sample.

2.2.2.2. Sampling Methods

Based on current published research, selected microbiome sampling methods were carried out on each participant and the yield compared. These consisted of an oral rinse compared to an oral swab, and a skin swab compared to a modified skin scrape. A sterile sampling solution was prepared consisting of 0.1% (v/v) Tween® 20 (Sigma-Aldrich® USA) and 0.15M NaCl, this was used to moisten the swabs prior to use with the skin sampling techniques and to store the swab used in the oral swabbing technique. Three participants consented to having multiple samples taken to enable statistical replication to be established, for each replicate sample taken from these participants occurred at least 36 hours after the previous sample was taken.

2.2.2.2.1. Oral Sampling

For the oral rinse, participants were asked to take 10ml of sterile distilled water into their mouths and then rinse for 30 seconds in a standardised rinse pattern. The rinse pattern consisted of 10 seconds in a up and down motion, 10 seconds in a back to front motion, and finally 10 seconds of a side-to-side motion. After rinsing the participants expectorated into a sterile container. Tween® 20 was added afterwards to a final concentration of 0.1% (v/v) lyse any human cells within the sample.

For the oral swab, a dry swab was inserted into participants oral cavities and rubbed against both cheeks and the tongue in an approximately 9cm^2 area for 10 seconds per area. The swab was then stored in 2ml of the sterile sample solution. The order the oral samples were taken in was alternated between participants.

2.2.2.2.2. Skin Sampling

The sample area was maintained using a sterile sampling template of 20cm^2, the area sampled was the interior of the right inner elbow. Each sampling technique was carried out in separate 20cm^2 area of the right arms inner elbow.

For the modified skin scrape a sterile L-shaped spreader was rubbed against the participants skin within the sampling area 10 times from top to bottom and 10 times left to right. Then the area was swabbed with a sterile swab that had been soaked in 2ml of the sterile sampling solution for 2 seconds. A consistent 'zigzag' swabbing pattern (Fig.2.1) was maintained. The swab was then stored in the 2ml of sterile sampling solution.

For skin swabbing a sterile swab was soaked in 2ml of sterile sampling solution for 2 seconds and then used to swab the participants skin within the sample area in a constant 'zigzag' pattern. The swab was then stored in the 2ml sterile sampling solution.

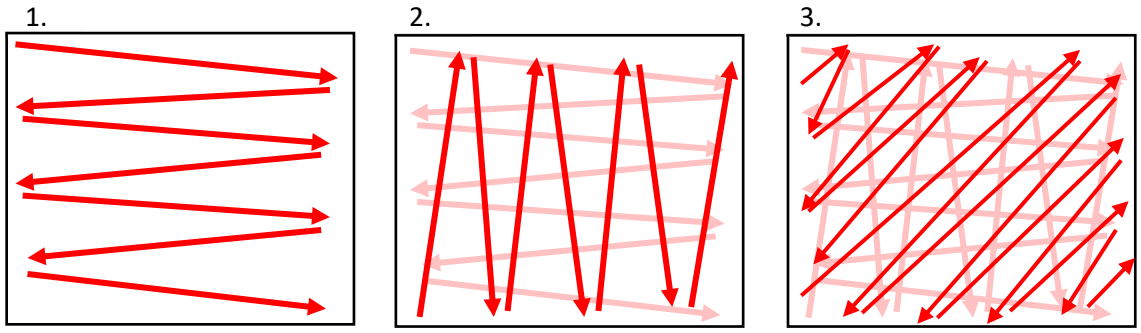


***Figure 2.1 A schematic demonstration of the 'zigzag' skin swabbing pattern.**. The swabbing is indicated in the red arrows and the black box is the sample area, the numbers show the order they are performed in.*

2.2.2.3. Viable Count

The viable counts were prepared immediately after the samples were taken; the samples were stored between 0-4^{O}C whilst work was carried out with them. All samples were vortexed for 5 seconds before being serially diluted in phosphate buffered saline (PBS) (Oxoid^{TM} UK). The oral samples were diluted to 10^{-6} and the skin samples were diluted to 10^{-4}. These dilutions were inoculated in accordance with the Miles and Misra method (Miles, Misra and Irwin, 1938). For the oral samples only dilutions from 10^{-2} – 10^{-6} were used and for the skin the neat sample down to 10^{-4} where used. Briefly Tryptone Soya Agar (TSA) (Oxoid^{TM} UK) was inoculated with 5 20µl drops of each diluent and allowed to dry, this was replicated a total of three times for each sample. The plates were incubated aerobically at 37^{o}C for 18 hours. After incubation the colonies were counted and averaged, and then expressed as CFU/ml for the oral samples or CFU/cm^2 for the skin samples. Prior to further investigative work all samples were made up to 10% (v/v) glycerol prior to storage at -20^{o}C.

2.2.2.4. Presumptive Identifications of Microorganisms

To attempt to presumptively identify the microorganisms recovered using the skin and oral sampling methods; all samples were cultured on a panel of selective and differential agar (Table 2.1) for appropriate incubation times, temperatures, and atmospheres. Agar sourced from Oxoid^{TM} UK were made in accordance with manufacturer's instructions, agar sourced from EO Labs LTD UK were supplied as pre-poured plates and stored in accordance with the manufacturer's instructions prior to use. After the incubation period macroscopic and

microscopic morphologies were noted, Gram staining, catalase, and oxidase tests were performed as further preliminary identification tests on selected colonies.

Table 2.1 Table detailing the microbiological agar. ***Details cover its source, incubation details, and sample used in presumptive identifications of microorganisms recovered from the oral and skin microbiome samples.***

Agar	Manufacturer	Inoculation source	Incubation temperature (°C)	Incubation duration (Hours)	Incubation atmosphere
SDA	Oxoid™ UK	Skin and Oral	30	48-72	Aerobic
MAC3			37	24-48	
MAN			30	48-72	
CBA	EO Labs LTD, UK		37	24-48	
FAA		Oral	37	24-48	Anaerobic
CXCA		Oral	30	48-72	Aerobic
CXSA		Skin	37	24-48	
SCBA		Oral	37	24	

2.2.2.5. Statistical Analysis

A two-way ANOVA was carried out on the viable count data using GraphPad Prism ('GraphPad Prism', 2023). This was carried out on oral rinse vs oral swab, and skin swab vs modified skin scrape, finally a comparison was made between rinse first and swab first for the oral samples.

2.3. Results

2.3.1.Viable Count

2.3.1.1.Skin Sampling

To determine which method of skin sampling between a skin swab and a skin scrape produced the most yield, and thus be suitable to use for microbiome investigations, a viable count was conducted on samples taken from 10 participants using both methods on the right inner forearm. From the skin count data (Fig. 4), there are no more than approximately 400 CFU/cm^2 on the inner elbow. No microorganisms were recovered from sample 4 using either method nor from the skin scrape from sample 10. There is a wide range of counts for each sample and method, with many overlaps in standard deviation and no clear pattern emerging beyond an indication that the skin scrape is producing a slightly higher yield in 7 of the samples. However, when subjected to statistical analysis using a two-way ANOVA there is no significant difference (P=0.3223) between the sampling methods. This suggests that the optimal method for skin sampling will be based on practicality.

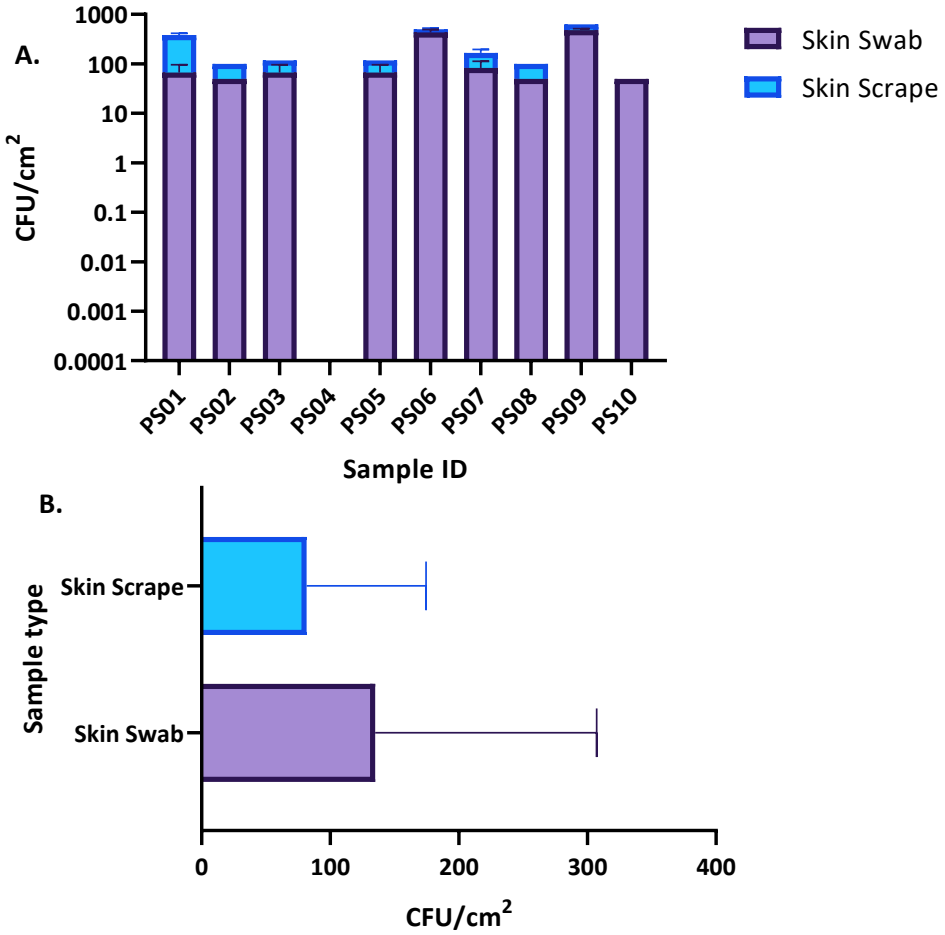


Figure 2.2 Viable count data for the skin swabbing and skin scraping methods. *No significant difference was found when compared using a two-way ANOVA. A. Logarithmic graph showing the CFU/cm2 from skin swabs and skin scrapes from participants including standard deviation. B. Graph comparing the mean of the sample types with standard deviation (P=0.3223).*

2.3.1.2. Oral Sampling

To determine which oral sampling method between an oral rinse and an oral swab produced the most yield, and thus be suitable for use in microbiome investigations, a viable count was conducted on samples taken from 10 participants using both methods. From the oral count data (Fig. 6) there are somewhere in the region of 1-10 million CFU/ml in the oral cavity. Unlike the skin data the standard deviations do not overlap as often. While again there is no obvious trend in the data beyond the general increase of rinsing over swabbing, there is a significant difference when subjected to a two-way ANOVA between the oral rinse and oral swab sampling methods (P=0.0440) When comparing the difference between rinsing first and swabbing first no significant difference (P=0.3309), indicating that neither method was sufficiently depleting the oral microbiome and effecting the subsequent sample.

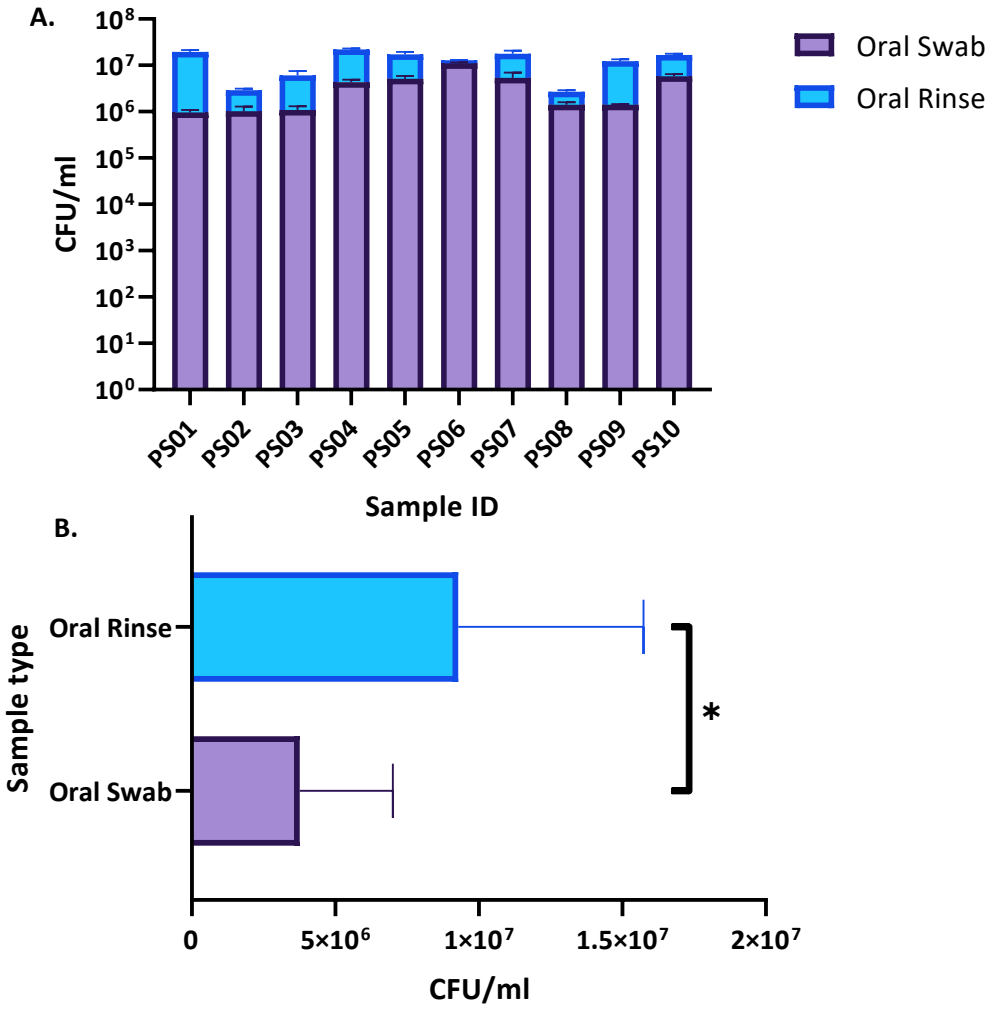


Figure 2.3 Viable count data for the oral swabbing and oral rinsing methods. *A significant difference was found when compared using a two-way ANOVA. A. Logarithmic graph showing the CFU/ml from oral swabs and oral rinses from 10 participants including standard deviation. B. Graph comparing the mean of the sample types with standard deviation *P=0.0440. Figure 2.3 Viable count data for the oral swabbing and oral rinsing methods a significant difference was found when compared using a two-way ANOVA. A. Logarithmic graph showing the CFU/ml from oral swabs and oral rinses from 10 participants including standard deviation. B. Graph comparing the mean of the sample types with standard deviation *P=0.0440*

2.3.1.3. Multi-sample comparison

Three participants consented to have three sets of samples taken, these were processed and subjected to viable counts like all the others. When subjected to a two-way ANOVA with Šídák's multiple comparisons post-test two of the three participants consistently showed a significant difference when comparing the oral methods (Table 2.2), while all consistently show no significant difference when comparing the skin methods.

Table 2.2 Comparing the p-values of three replicates of the skin and oral methods in three participants.

Comparison	Sample ID	p-value
Oral Rinse vs. Oral Swab	PS02	0.6399
		0.9782
		<0.0001
	PS06	<0.0001
		<0.0001
		<0.0001
	PS09	<0.0001
		<0.0001
		<0.0001
Skin scrape vs. Skin Swab	PS02	>0.9999
		>0.9999
		>0.9999
	PS06	>0.9999
		>0.9999
		>0.9999
	PS09	>0.9999
		>0.9999
		>0.9999

2.3.2. Presumptive Identification

2.3.2.1. Presumptive Skin Organisms

Presumptive identifications of skin microorganisms grown from the skin samples on selective and differential agars were carried out to determine which microorganisms can be recovered from these samples the larger scale microbiome study that is the focus of this thesis. Microorganisms were recovered from nearly all samples except PS07 and PS06. *Bacillus spp.* was recovered from 5 participants whereas *Staphylococcus spp.* was only recovered from 4 participants and fungi from 3 (Fig.2.4). PS05 consistently provided *Bacillus*, *Staphylococcus*, and fungi from the samples, other samples regularly provided two of the three microorganisms. Only 1 participant provided an unidentified Gram-negative microorganism, further details of the recovery on each media can be found in supplemental table 9.1.

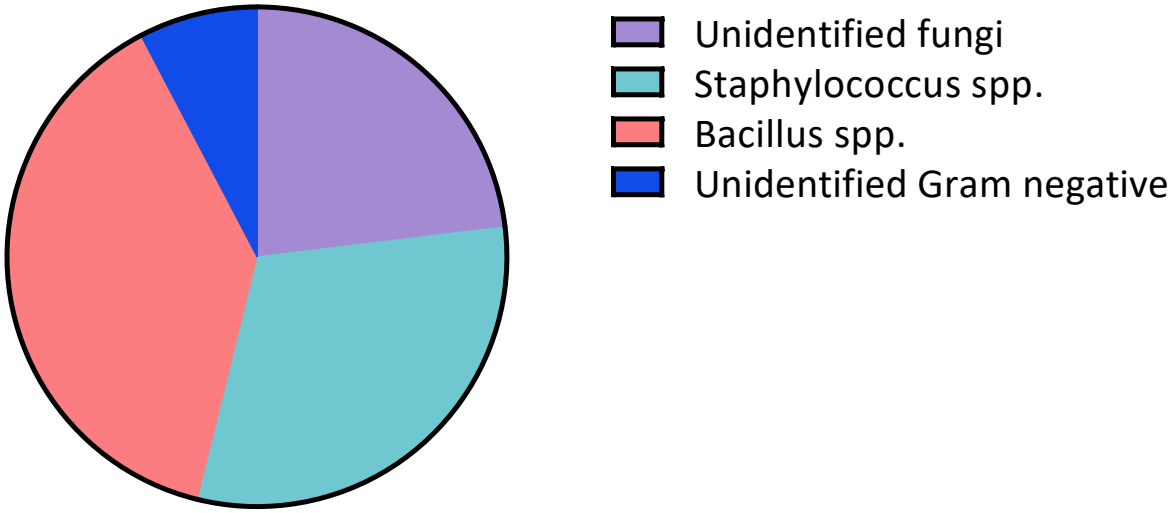


Figure 2.4 Pie chart showing the proportion of skin microorganisms detected using culture-based methods. ***Results from samples taken using both skin swabbing and skin scraping.***

2.3.2.2. Presumptive Oral Organisms

Presumptive identifications of oral microorganisms grown from the oral samples on selective and differential agars were carried out to determine which microorganisms can be recovered from these samples in a larger scale microbiome study. All samples contained *Streptococcus spp.* and all except 2 yielded *Staphylococcus spp.* (Fig. 2.5). *Candida albicans* was detected in 3 participants with one having both *Candida albicans* and *Candida glabrata* present in their sample. Unidentified Gram-negative microorganisms were detected in 4 participant samples. Unidentified Gram-positive, fungi, and *Bacillus* were only detected in 2 participants samples for each organism. Less *Bacillus spp.* and unidentified fungi were detected in these oral samples than the skin samples, conversely more unidentified Gram-negative and *Staphylococcus* were present in the oral samples than the skin. No participants had all these microorganisms present in their oral samples, further details of the recovery on each media can be found in supplemental table 9.2.

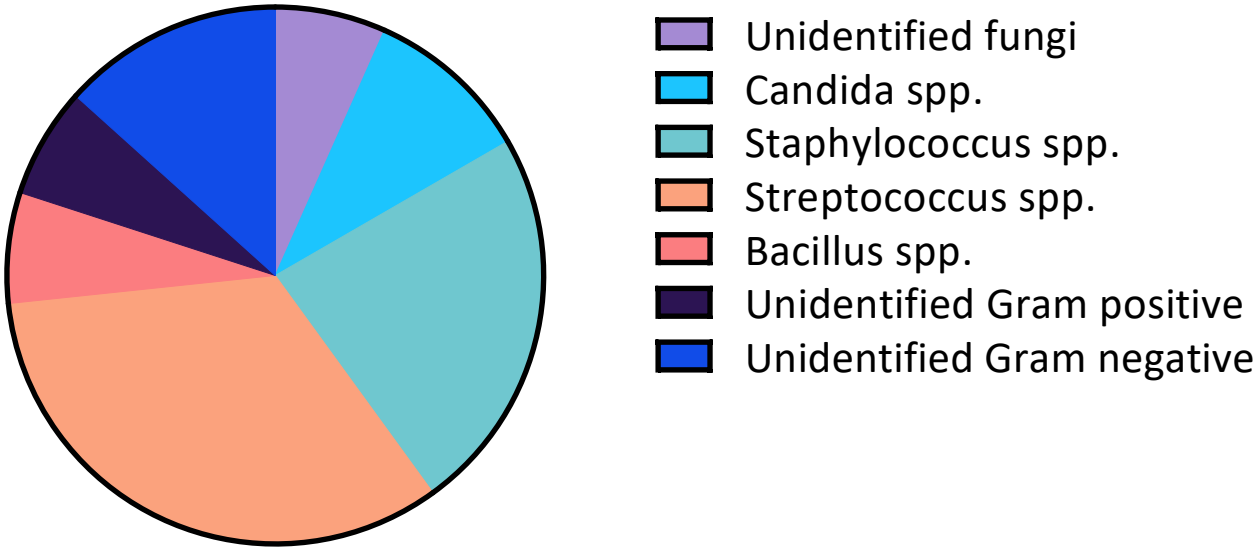


Figure 2.5 Pie chart showing the proportion of oral microorganisms detected using culture-based methods.*Results from samples taken using both oral swabbing and oral rinsing.*

2.4. Discussion

This chapter covered the work to discover which is a suitable sampling method for use in the larger scale study into the content of the skin and oral microbiomes which is the main goal of this thesis. Along with determining which microorganisms can be recovered from these microbiomes using standard culture techniques.

Using the viable count method the 100-400 CFU/cm^2 recovered from the skin samples using both the skin swab and skin scrape methods is within the expected ranges of skin microorganisms which is between 10^3 and 10^4 for an area such as the forearm (Skowron *et al.*, 2021). Similarly, the oral samples provided yields in the region of 1-10 million CFU/ml which can be confirmed by multiple sampling methods (Funahara *et al.*, 2018; Wenger *et al.*, 2021). The lack of growth from PS04 and PS10 skin samples is not due to these participants skin being clean as microorganisms were recovered in the presumptive identifications but is likely to be due to the difficulty in recovering and extracting microorganisms from skin samples. The methods employed here do have limitations, ideally there should have been three samples taken from each participant over a period of time i.e. about a week, however this would have put an extra burden upon the participants as it would have required either tracking or limiting their hygiene routine and ensuring samples were taken at similar times of the day to eliminate external factors from the results. Considering this and that this small-scale study phase is just to establish a suitable method the lack of replicate samples for all participants is tolerable, this is reenforced by the internal controls generated from the three participants who were able to give three samples, as most of the statistical results are consistent across all samples.

The Miles and Misra method has a drawback that the multiple droplets of each serial dilution can merge thus preventing accurate counts being performed, while every practical attempt was made to prevent this, merging did happen on some occasions. More generally any sampling method that used a swab will have problems yielding the microbes from the swab into solution as they may have bound to or become trapped within the swab material making recovery difficult (Van Horn *et al.*, 2008; Jansson *et al.*, 2020), if the study was just using culture based methods then it would be possible to perform an initial inoculation on a general agar using the swab thus increasing the chances of yielding microorganisms. However, as the main study intends to use the same sample for both culturomics and bioinformatics it is necessary to ensure that the microorganisms are evenly distributed in a sampling buffer prior to processing. The skin scraping method undertaken in this current study was drastically different to those carried out by other research where a sharp or rough implement like a scalpel or ground glass is used. These methods were not used for this study as they pose a risk of actual harm to the participants and can have logistical or practical difficulties with sample taking, transport, and processing.

Overall, the visible and statistical difference between the oral swab and oral rinse sampling methods show that as rinsing provides the highest yield and thus it would be the best sampling method to be adopted in the larger scale study of this research project. Conversely as there is no statistical difference between the skin sampling methods, this leaves the option for which is most suitable down to the practicality thus making the skin swabbing method the best choice.

The microorganisms recovered from the skin samples are typical of those normally found as part of the human skin microbiome. While species specific identification is not possible using the techniques used here, it is possible to infer that the *Staphylococcus* colonies are likely to be *S. epidermidis*, *S. hominis*, or *S. aureus*; similarly, the fungi are possibly *Malassezia spp.,* and the *Bacillus spp.* could be a *B. subtilis*. These speculations however are risky as with skin microorganisms they could just as easily be transient microorganisms from the environment on the participants skin, species specific identification would enable confirmation of the colonies and therefore they could part of the normal skin microbiome. As previous commented upon there was a level of difficulty in recovering the microorganisms from the skin sample and them being recovered consistently, for example a *Staphylococcus spp.* would grow on CBA but then there would be no growth on mannitol salt agar (MAN) which is differential for *Staphylococcus spp.* or sabouraud dextrose agar (SDA) which *Staphylococcus spp.* may grow on despite it being designed to be selective to fungi. Therefore, it might be necessary to concentrate the sample being used in culture methods prior to use, possibly via centrifugation and resuspension in a smaller volume of buffer.

Similarly, the oral samples yielded microorganisms typical of the oral cavities normal flora, even the variations in species presence for example *Streptococcus spp.* are present in all samples whereas *Candida spp.* was only present in some participants; given the wide range of microorganisms known to be present in the oral cavity speculation over the exact species present like that done with the skin samples would be unwise (Kleinegger *et al.*, 1996; Caselli *et al.*, 2020). However, further work could be done to identify the unknown Gram-negative and positive strains. For example, on MacConkey agar No.3 (MAC3) *Escherichia coli* and *Klebsiella spp.* will look similar due to their ability to ferment lactose and therefore further work would be required using the Indole test or other biochemical analysis to differentiate them. While participants were asked to refrain from eating for 30 minutes prior to giving the sample this may not have been enough time for their oral microbiome to recover and therefore more transient microorganisms have been detected; therefore, in future studies it may be necessary to consider lengthening this time to 1 hour, however care must be taken to not increase the burden upon participants particularly those who have specific dietary needs. The increased quantity of microorganisms within the oral cavity also means that some of the agar plates were coated with an undefined bacterial lawn, this means that single colonies could not easily be targeted for Gram staining or catalase and oxidase testing thus possibly causing errors in their results. In future work with oral sample recovery, it will be necessary to dilute the samples prior to inoculating the media, using more selective media may also be required.

There is a general drawback to the identification of both the skin and oral samples using these culture techniques, due to the wild nature of such samples the microorganisms present maybe expressing different genes than those expected for the species and thus behave differently on the agars than the known isolates, for example on MAN *S. aureus* may not produce acid and thus not express the typical yellow halo around the colonies on this agar (Ishii *et al.*, 2006). Further and more detailed identification maybe possible with API® testing strips (bioMérieux, France), which run several biochemical tests simultaneously that can differentiate isolates on both a species and strain level, however these are expensive and would require an extra processing step as each suspect colony would need to be grown in liquid broth prior to use; also, atypical isolates may still yield erroneous results and therefore are not an appropriate choice for this work. Exacting identification could have been carried out using sequencing however this was beyond the scope of this stage of the research. While freezing and thawing the samples prior to performing the presumptive identifications work may have killed some cells, however as the samples were treated with glycerol and only underwent one cycle of freeze thawing prior to being used it is unlikely this has had a significant impact upon the number of cells recovered.

2.5. Conclusion

Overall, this pilot study has demonstrated that for the main study an oral rinse and skin swab sampling technique will be suitable for use as they give a higher yield than the others tested and are practical to perform; this was done in a robust manner with a suitable number of participants and these decisions were made with the use of statistically significant data. Similarly, the presumptive identifications have shown the diversity of the oral and skin microbiomes and the challenges that these pose in recovery; thus, considerations such as concentrating the skin sample and more specific selective agars has been factored into the main study.

Chapter 3. Participant Metadata Analysis

3.1. Introduction

Having determined the optimal sampling methods for sampling the oral and skin microbiome in chapter 2 of this thesis, they can now be used to recover samples of the skin and oral microbiome of younger and older adults to address the central theme of this thesis, which is the comparison and exploration of the oral and skin microbiomes between older and younger adults. Additionally, as outlined in the introduction (1.4), the human microbiome is complex and can be affected by many different factors therefore additional data about the participants needs to be gathered to aid with the analysis of the microbiome data. Here this metadata will be reviewed to understand additional differences in the participant population.

An important factor for consideration within this study is the ages of the participants. As ageing is a discrete continuous process to enable the study to have statistical relevance there must be distinction between the age groups in the study. The age ranges chosen must also be clinically appropriate, i.e. it should compare younger adults against older adults, however some younger adults still maybe undergoing late stage adolescence as defined by the World Health Organisation and thus may affect the microbiome composition and prevent a fair comparison to older adults (Sacks, Canadian Paediatric Society, and Adolescent Health Committee, 2003). This is why the younger adults (YA) recruited for this study start at 20 years of age till 40 years of age, additionally to maintain statistical parity the older adults (OA), are aged between 60-80 years of age. On the advice of this thesis funders Colgate-Palmolive the minimum required number of participants in each group is 25 to ensure statistical power during microbiome analysis.

3.1.1. Aims

- Describe the recruitment and sampling of the participants in this main stage of the study.
- Analyse the demographics of the participant population to determine the possibility of creating sub-groups suitable for using in analysis in future components of this research study, either through contributions to the ageing process or the composition of the skin and oral microbiomes.

3.2. Materials and methods

3.2.1. Ethical approval

The original ethical approval used in chapter 2 required amendments to enable this remainder of the work in this thesis to be undertaken, therefore, an Ethical Committee amendment application was submitted to Aston University HLS Ethics Committee. In addition, the accompanying consent form and information sheet were expanded and submitted to accommodate the increased scope of the project; mainly by increasing the number of participants (Minimum of 50 to a maximum of 80), changing the methodology to use only the optimal method

as outlined in chapter 2 (section 3.1), and altering the inclusion criteria. The main restriction on inclusion was the participant ages, participants had to be aged between 20-40 (Younger Adults (YA)), and 60- 80(Older Adults (OA)), years old. Additionally submitted with the amendment was an advert and series of questionnaires to gather appropriate participant metadata. Approval to the amendment was granted by the HLS Ethics Committee (HLS21008).

3.2.1.1. Participant questionnaire design

The questionnaires were created in Qualtrics ('Qualtrics', 2022). The questionnaires were separated by their different topics, the first questionnaire covered both demographic information and acted to screen if the potential participants met the eligibility criteria for the study. The further questionnaires used gathered participant health data and hygiene habits data respectively. An overview of the content of these questionnaires can be seen in table 3.1 (Full questionnaires available in 9.2.1), these questionnaires were designed after considering possible relevant information obtained from reviewing the literature, the questions were then phrased to place the least burden on the participant while also having neutral wording.

Table 3.1 Table outlining the content covered in the participant screening, health, and hygiene questionnaires.

Screening questionnaire	**Health questionnaire**	**Hygiene questionnaire**
Date of birth	Long term health conditions	Toothbrushing routine and frequency
Sex	Smoking habits	Other dental care routine and frequency
Postcode	Alcohol habits	Use of dental hygienist
Ethnicity	Receipt of any current prescribed treatments	Skin hygiene routine and frequency
Relationship to Aston University	Diet and activity	Oral and skin product usage
Ability to attend on campus participation session	Dental/oral health and status	
	Skin health and status	

3.2.2. Participant recruitment

The study advertisement was distributed via email to the Aston Research Centre for Health in Ageing (ARCHA), group and internally within Aston University to groups (Final year undergraduates, postgraduates, and staff), who would contain the ages ranges in the study eligibility criteria. The advert was also placed on social media along with physical copies placed around Aston Universities campus.

3.2.3. Participant sampling

3.2.3.1. Materials preparation

For sampling the participant oral microbiome distilled water (DH_2O), was sterilised by autoclaving at 121^{o}C for 15 minutes. Following sterilisation, 10ml of this was aliquoted into sterile 125ml containers (Corning, USA), prior to use.

For sampling the participant skin microbiome, the following materials were also prepared. Prior to autoclaving PBS (Oxoid, UK) solution was modified with Tween® 20 (Sigma P9416) to create a 0.1% (v/v) solution, this was then sterilised by autoclaving at 121°C for 15 minutes. Two millilitres of this solution was aliquoted into sterile bijous prior to use. Sampling templates with an internal area of 25cm^2 were made from food safe pourable silicone (MB Fiberglass, UK), these were sterilised in pairs inside autoclave pouches (SLS, UK) by autoclaving at 121°C for 15 minutes.

The following materials were prepared for processing of the samples. A solution of 10% (v/v) Tween® 20 (Sigma, UK) was made and sterilised by autoclaving at 121°C for 15 minutes. Finally, a solution of 80% (v/v) glycerol (Melford, UK) was made and sterilised by autoclaving at 121°C for 15 minutes.

3.2.3.2. Sampling procedure

Prior to their participation session appointment, participants were asked to refrain from eating for at least 1 hour. For the oral sample participants were asked to take the 10ml of sterile DH_2O into their mouths and then rinse for 30 seconds with the aid of a stopwatch in a standardised rinse pattern. The rinse pattern consisted of 10 seconds in an "up and down" motion, 10 seconds in an "back to front motion", and finally 10 seconds of a "side-to-side" motion. After rinsing the participants expectorated back into the 125ml container.

Two samples of the skin were taken one from the middle of the right cheek and another of the interior of the right inner elbow (Fig 3.1). For each area a sterile swab was soaked in the 2ml of sterile PBS and Tween® 20 solution for 2 seconds and then used to swab the participants skin within the sample area in a constant 'zigzag' pattern (as described in Chapter 2, section 2.2.2.2). The swab was then stored in the 2ml sterile PBS and Tween® 20 solution.

All samples were stored at 2-8°C prior to processing, and were processed within 24 hours of obtaining them, this decision was extrapolated from an understanding of faecal microbiome sample storage (Choo, Leong and Rogers, 2015). After these samples were collected each participant completed both the health and hygiene questionnaires, through the means of an interview.

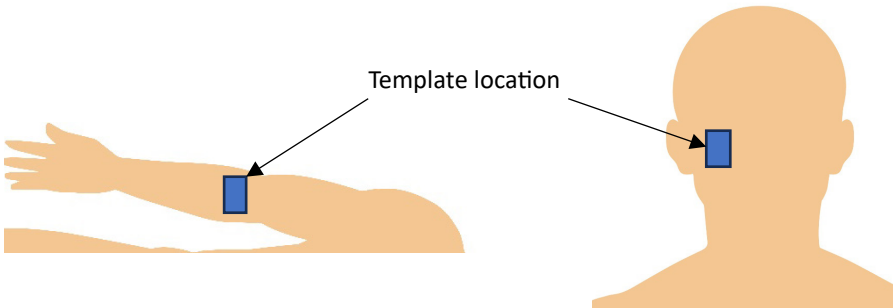


Figure 3.1 Diagram showing the location of the skin sampling templates. *Human shape adapted from Freepik.*

3.2.3.3. Sample processing

Within 24 hours of receiving the sample sterile 10% (v/v) Tween® 20 was added to the oral rinse samples to a final concentration of 0.1% (v/v), and vortexed along with the skin samples for 10 seconds to lyse any human cells within the sample.

A aliquot of each sample was taken and mixed with the 80% (v/v) glycerol solution to create a solution of 20% (v/v) glycerol and sample of 1.5ml in volume. These were then stored at -80^{O}C to act as reserve sample stocks. The remaining sample was then used as described in the following chapters.

3.2.4. Metadata processing and scoring

Participant responses to the questionnaire were extracted from Qualtrics and anonymised. Each participant was given a randomly assigned participant ID code; these were split into the Younger Adults (YA) and Older Adults (OA). Responses to questions on general health and alcohol consumption were scored on a frailty index (FI) and Alcohol Use Disorders Identification Test Consumption (AUDIT-C) scale respectively (Kim and Jazwinski, 2015; Venturelli *et al.*, 2021). Participant ages were calculated from the date of participation. Participant postcodes were scored using English indices of deprivation 2019 and classified using the Postcode to Output Area codes (A glossary of the codes used is available in supplementary table 9.3) for each participant, no identifiable postcode data has been used in the following data sets (Ministry of Housing, Communities & Local Government, 2019; Office for National Statistics, 2022).

3.2.5. Metadata statistical analysis

Metadata was initially analysed using IBM SPSS statistics using Fisher's exact test. All graphs were constructed in R studio using the tidyverse package, GraphPad Prism, or SankeyMATIC (Wickham *et al.*, 2019; 'IBM SPSS Statistics', 2021; 'GraphPad Prism', 2023; *SankeyMATIC: Make Beautiful Flow Diagrams*, 2023).

3.3. Results

3.3.1. Participant demographic data

From a maximum of 80 potential participants 59 were recruited into the study, this comprised of 30 YA and 29 OA. The age ranges (Fig 3.2), of the YA was 20- to 39-year-olds and in the OA 64- to 80-year-olds, the YA median age was 26 years old, and the OA median age was 72 years old; both have a standard deviation of 4 years either side of this median. When split by gender the YA group includes 18 females and 12 males, while the OA group includes 17 females and 12 males. Considering the ethnicity of the groups, the data shows that the OA group members belong entirely to White ethnicities whereas the YA group contained 1 participant from a Mixed/Multiple ethnic background, 2 from Other backgrounds, 8 from Asian backgrounds, and 19 from White backgrounds. In detail the YA group has 8 Asian, 2 Mixed/Multiple, 1 other, and 9 White female members, but only 2 Asian, and 10 White male members. Overall, this demonstrates the initial diversities present within the participant groups, but more detailed investigation into the participants may yield other differences.

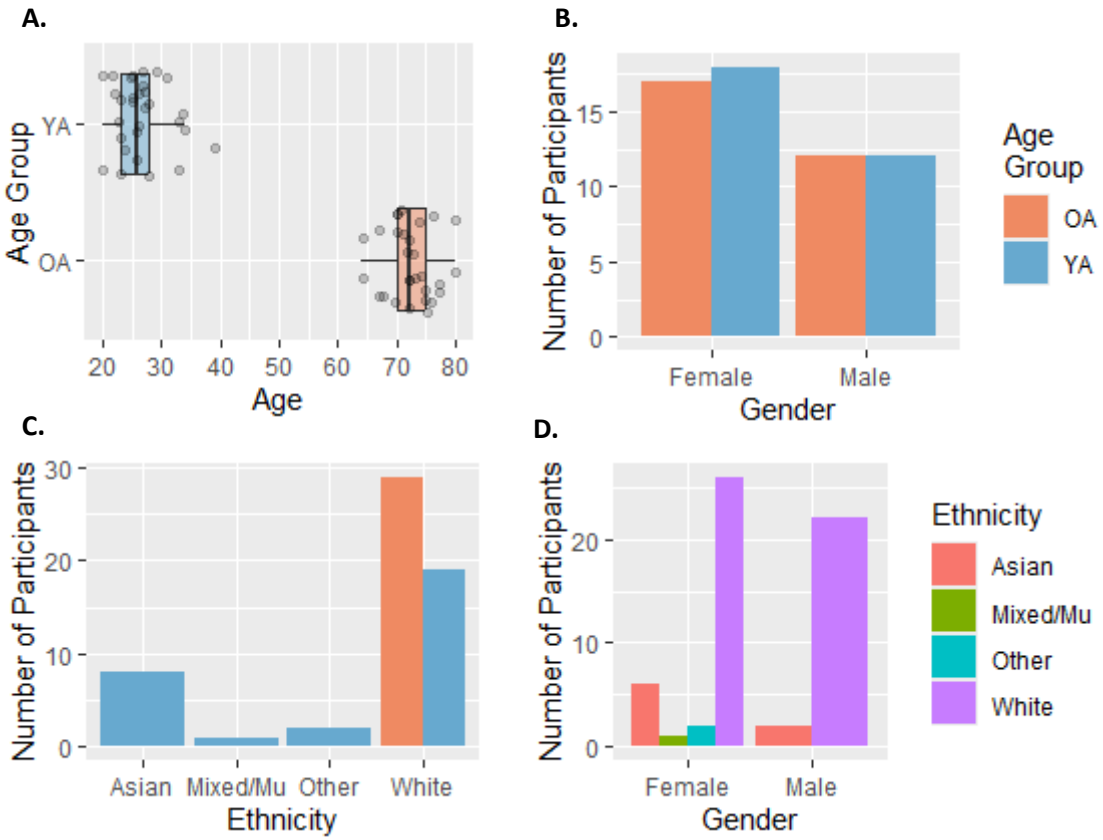


Figure 3.2 Broad demographic data of the participant populations. *A. A box and whisker plot of the YA and OA age groups age ranges. B. The number of male and female participants in the YA and OA groups (p=1.00). C. The number of participants in different ethnic groups in the YA and OA groups (p=<0.001). D. The number of participants in different ethnic groups in the YA group separated by gender (p=0.292).*

3.3.2. Participant sociodemographic data

When participant postcodes were used to generate sociodemographic data using the Output Area Hierarchy and Classifications dataset (code definitions in supplementary table 9.3), this generated 4 subset codes to group participants postcodes into (Fig 3.3); in both the YA and OA groups a single participant was not scorable on this dataset due to providing invalid postcodes. The majority of the YA group are in the 2A1, 2B2, and 4A2 groups of the Output Area code, while the OA group are more representative in the 5 and 6 groupings, a significant difference between the YA and OA groups (p=0.026), was observed when tested using a Fisher's Exact test. Using the Super Output Area code 16 YA are in the 1A or 7A groups, while 11 of the OA are in the 3B and 8A groups; there was a significant difference between the YA and OA groups (p=0.001), when tested using a Fisher's exact test. Using the Workplace Zones codes, the OA group have 8 members in the D2 group and the next most populace groups being C1, C4, and D1, with 3 participants in each. In the YA group the B3 and then the C1 groups are the largest, when tested using a Fisher's Exact test there was a significant difference (p=0.044). Finally, the Local Authority codes showed both the YA and OA group having the most participants in the 4A1 code, with 15 and 7 respectively; there was no significant difference between the OA and YA group using a Fisher's Exact test (p=0.092).

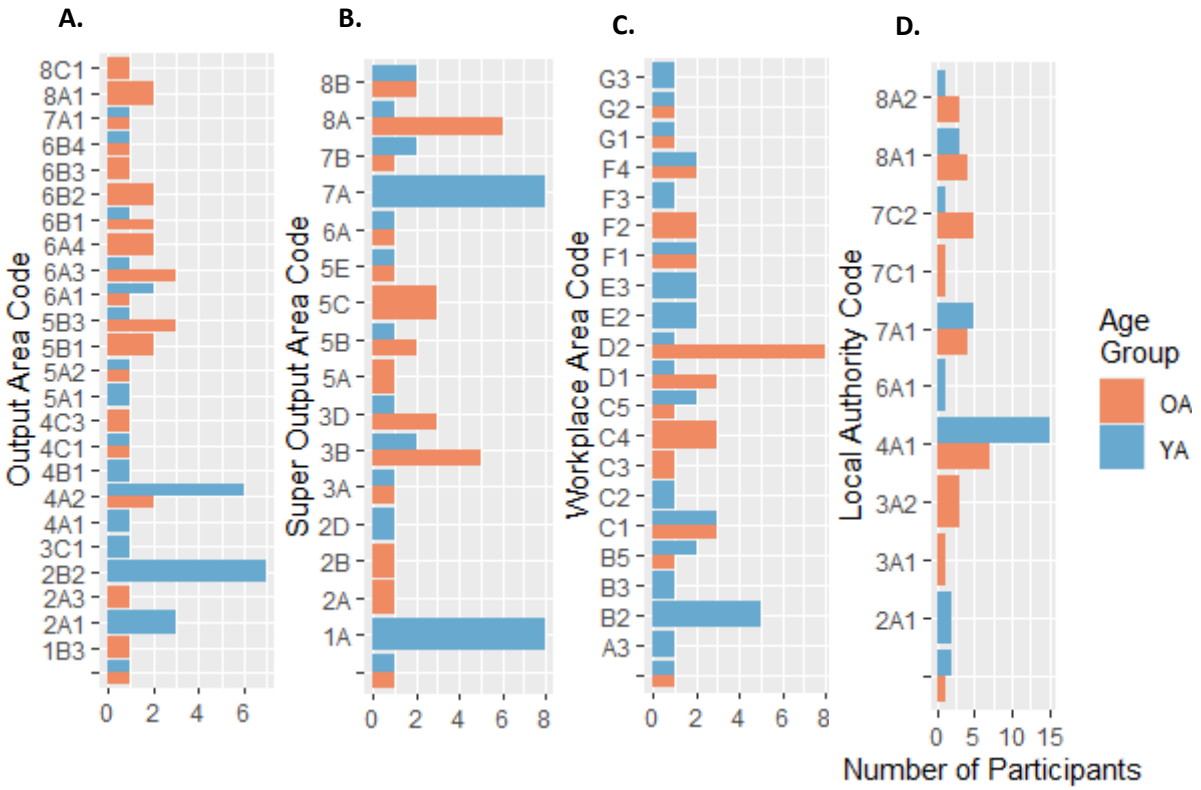


Figure 3.3 Participant sociodemographic data.*Results generated from the participant postcodes using the Postcode to Output Area Hierarchy with Classifications (2022) dataset, split by the YA and OA groups showing significant difference in some metrics. A. The number of participants in each output area code (p=0.026). B. The number of participants in each super output area code (p=<0.001). C. The number of participants in each workplace zone code (p=0.044). D. The number of participants in each local authority code (p=0.092).*

Using the Index of Multiple Deprivation Deciles (IMDD) to group both the YA and OA groups (Fig 3.4), 6 of the YA and OA are in the 3rd and 10th decile respectively, there was no significant difference between these groups when tested with a Fisher's Exact test (p=0.198). Using the Health and Disabilities deciles 9 of the YA group score at the 3rd decile compared to 3 of the OA and 6 OA score at the 7th decile compared to 1 YA, there was no significant difference between these groups when tested with a Fisher's Exact test (p=0.185). The living environment decile scores show 9 of the YA score at the 1st decile compared to 2 of the OA, while the highest decile scored by the OA is the 10th with 5 OA and 3 YA; when subjected to a Fisher's Exact test there was a significant difference between the YA and OA groups in this decile (p=0.048). The Income Deprivation Affecting Older People Index Decile sees 7 and 6 of the YA score at the 1st and 2nd decile compared to 1 and 3 of the OA group, while the 10th decile has the highest number of OA with 6 participants in this category compared to 1 YA; there was no significant difference between these groups when tested with a Fisher's Exact test (p=0.164). These conclude the various aspects of the participant sociodemographic data obtainable from their postcodes, this has shown additional differences between the participant age groups.

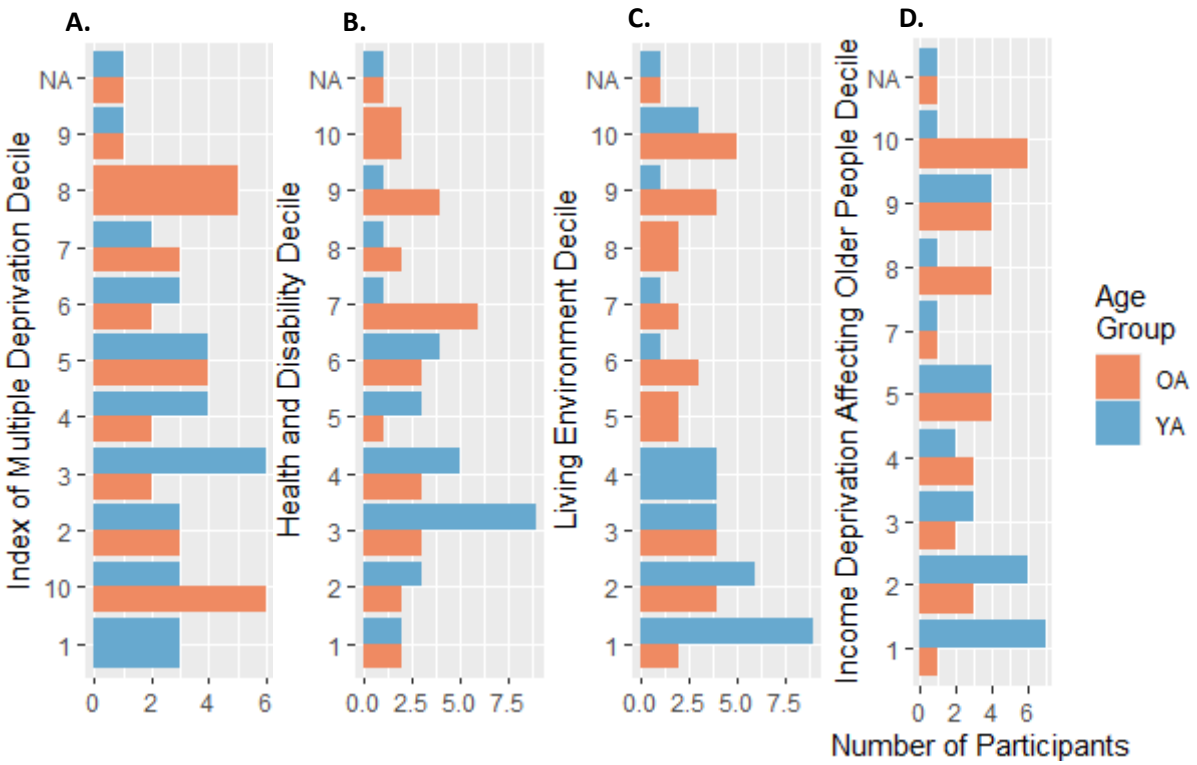


Figure 3.4 Participant deprivation indices data. *Results generated from the participant postcodes using the English Indices of Deprivation (2019) dataset, split by the YA and OA groups showing no significant differences. A. The rank of numbers of participants in the Index of Multiple Deprivation Decile (p=0.198). B. The rank of numbers of participants in the Health and Disability Decile (p=0.185). C. The rank of numbers of participants the Living Environment Decile (p=0.048). D. The rank of numbers of participants in the Income Deprivation Affecting Older People Index Decile (p=0.164).*

3.3.3. Participant health status data

Beyond the sociodemographic groupings it is also possible to view the participants through their health data. After scoring participants on a frailty index (Fig 3.5), there was a significant difference between the YA and OA groups (p=<0.001), with the OA group scoring consistently higher than the YA. When broken down by the genders of the age groups, the OA group has between 2 and 4 participants scoring 0, 1, and 3 on the FI regardless of gender, however only 1 male but 8 females scored 2 on the frailty index, with finally 1 male but no females scoring 4. With the YA group 5 male and female participants each scored 1 on the FI, only 1 female scores a 3 on the FI with no males, nearly twice the number of females score 0 on the FI compared to males, none of either gender score 2, or 4 on the FI.

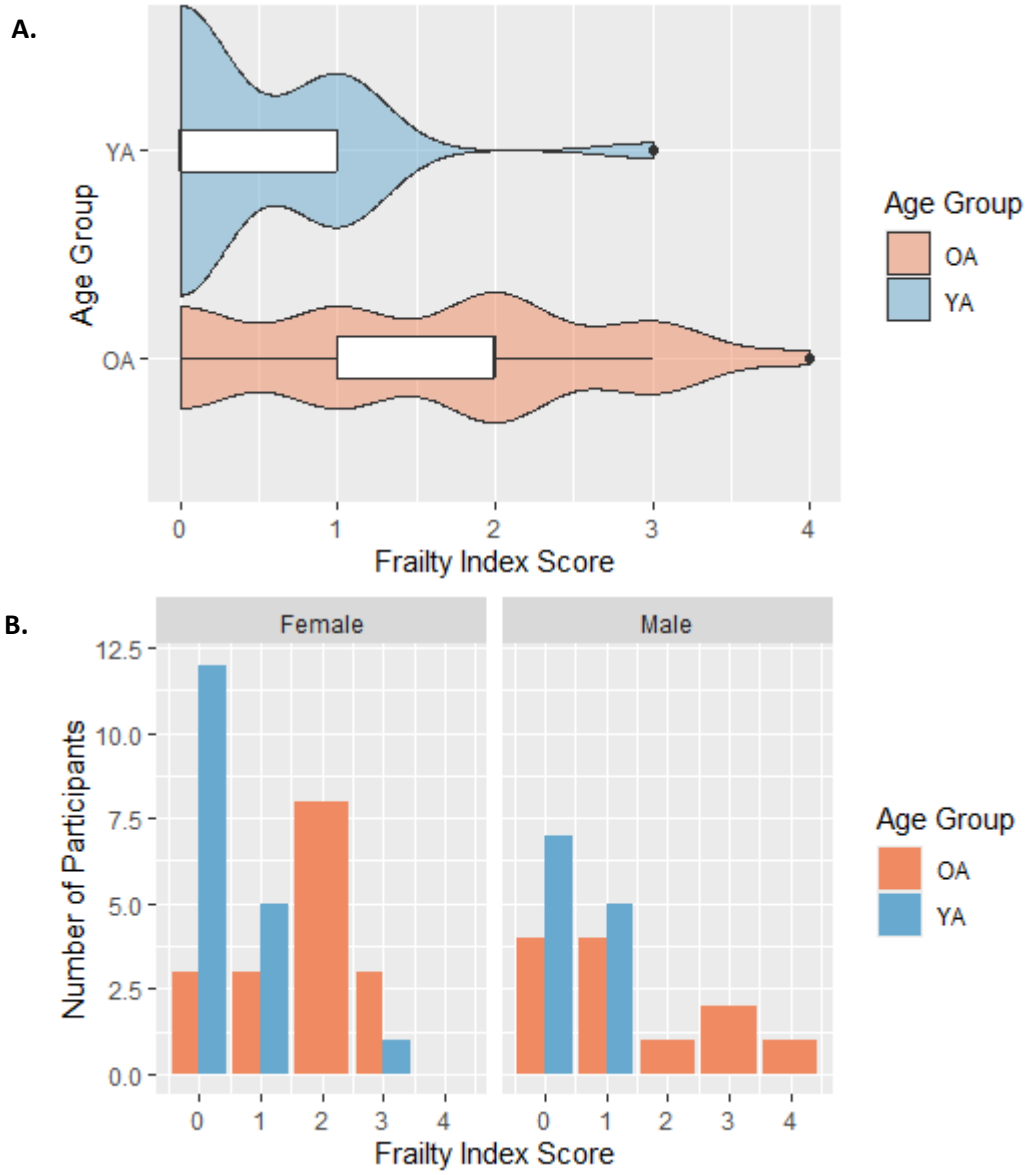


Figure 3.5 Participant frailty index score data. *Showing a significant difference between the age groups. A. Bow and violin plot showing a significant difference (p=<0.001), between the two groups. B. The frailty index scores of the age groups divided by the participant genders.*

An additional overall health factor is alcohol consumption, when scored using an AUDIT-C scoring system (Fig 3.6), both the YA and OA scores ranged from 0 to 6, there was a significant difference between these groups (p=0.017). When each group was analysed in detail in the OA group shows a general trend of alcohol consumption regardless of gender. However there generally there were more female participants than male in all scores apart from 6 where only 1 male participant scores; a total of 9 participants scored 0 on the scale. Amongst the YA group, 11 participants score 0, more female participants score 1 and 3 than the male, but more males scored a 2 and 6 than female; overall there is a significant difference between the genders in the YA group (p=0.002).

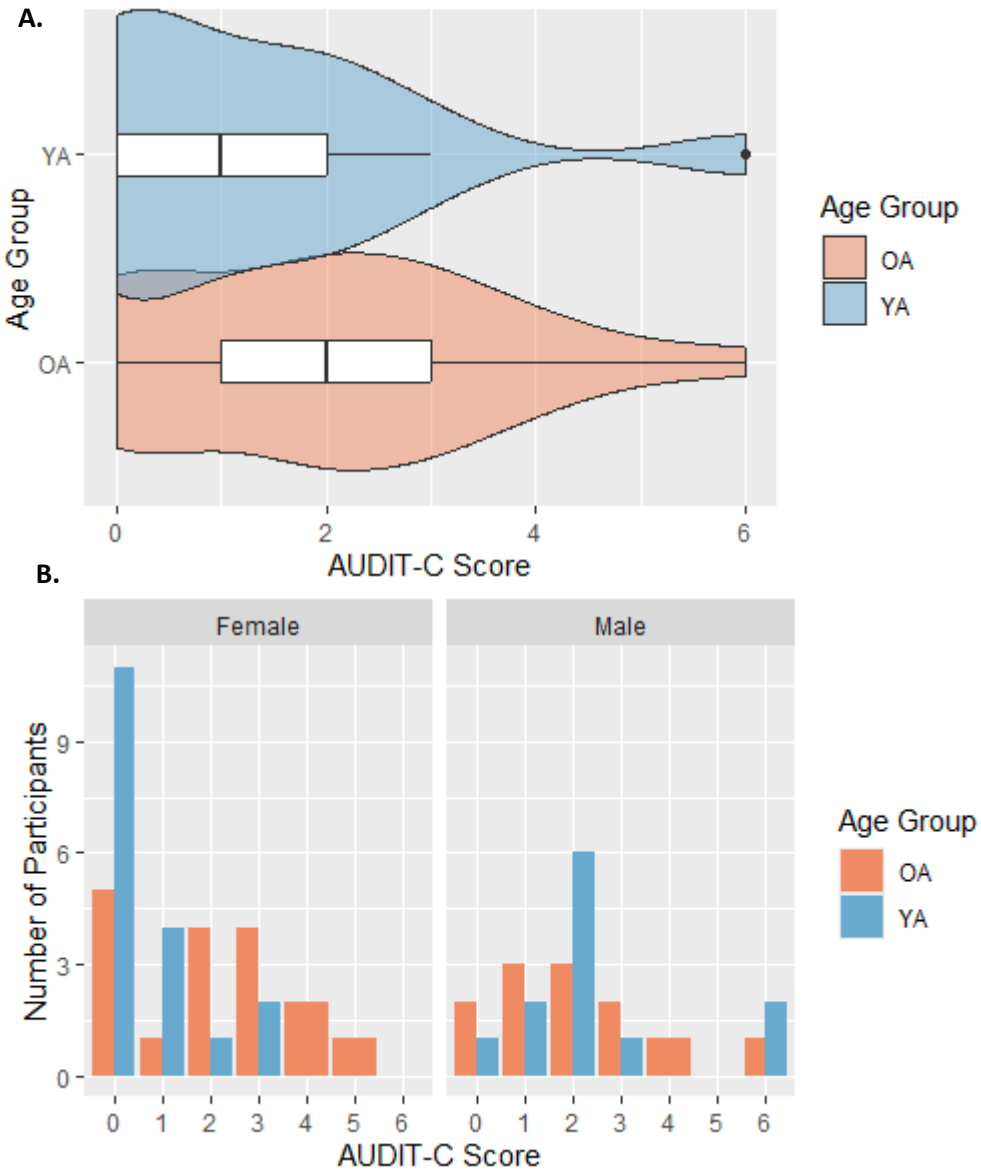


Figure 3.6 Participant AUDIT-C score data. *Showing a significant difference between the age groups. A. Bow and violin plot of participants showing a significant difference (p=0.010). B. The AUDIT-C scores of the age groups divided by the participant genders.*

As an aspect of this thesis is the oral microbiome it is worthy to consider the participants oral health. Comparing the dental health of the participants (Fig 3.7), shows that 23 OA and 11 of the YA groups have had at least one tooth other than their wisdom teeth removed. A total of 18 YA and 28 OA reported having fillings with 2 in each group having had them within the last 6 months, there was a significant difference between these groups using a Fisher's exact test (p=0.001). Of the OA group 4 reported root canals and 8 reported crowns, but only 2 YA reported root canals and 3 crowns. Only in the OA group were dentures, implants, and bridge work reported, while the YA group reported braces and retainers. From the OA group 15 reported having their tonsils removed compared to 3 from the YA group, when investigated using a Fisher's exact test a significant difference was observed between these groups (p=<0.001).

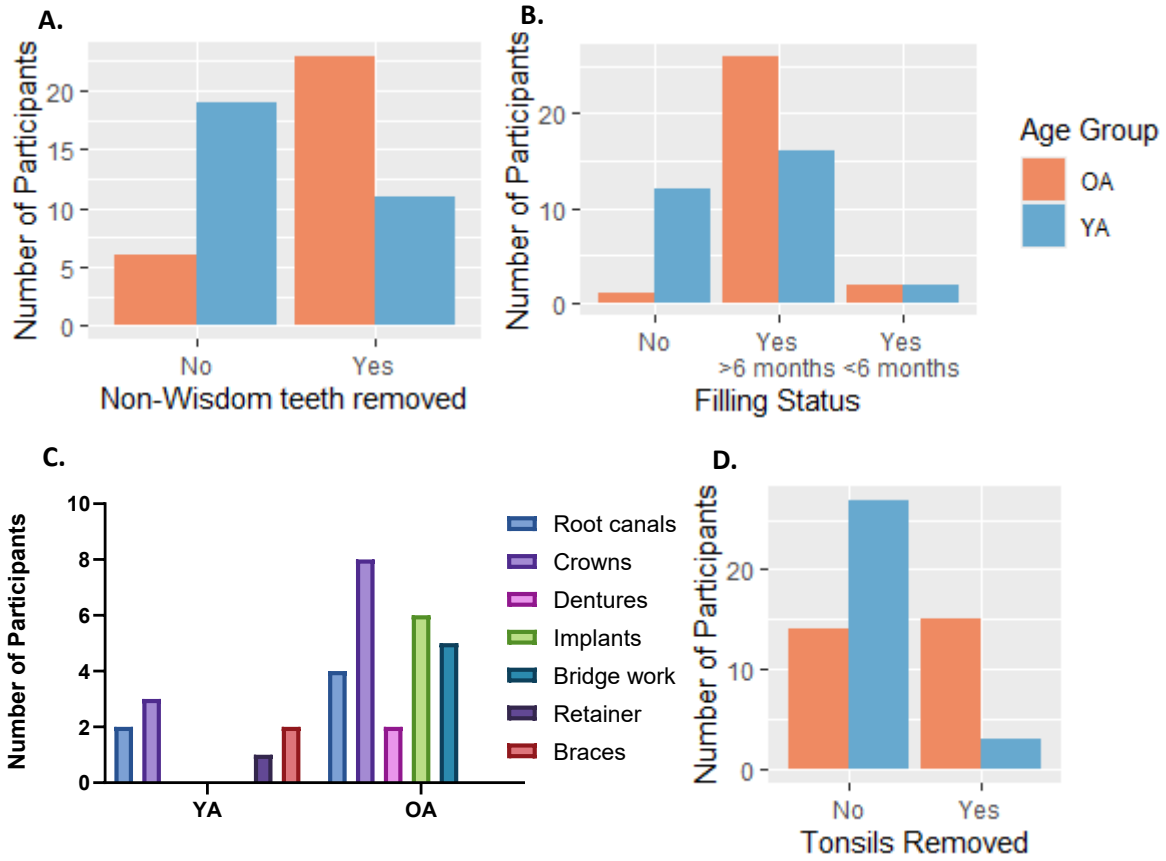


Figure 3.7 Participant dental health. *Data divided by the OA and YA groups shows a significant difference in some metrics. A. The number of participants who have had adult teeth other than wisdom teeth extracted. B. The number of participants who have had fillings within timeframes (p=0.001). C. The number of participants with various other dental work (p=<0.001). D. The number of participants who have and haven't had their tonsils removed (p=<0.001).*

In addition to the oral health of participants, it is possible that their skin health will factor into their skin microbiomes, therefore some assessment of it is justified. When measuring skin health 8 YA and 2 OA reported suffering from acne in the past, a further 5 YA and 10 OA reported suffering from other skin conditions in the past, but this was not a significant difference (p=0.073). Currently only 7 YA and 6 OA reported from a skin condition at the time of sampling, again this was not a significant difference (p=1.000). When asked if they were

currently using a skin treatment that is prescription only medication or pharmacy medicine 3 of the OA and 2 of the YA said yes, when compared using a Fisher's exact test this was not a significant difference (p=0.671). When plotted over time, of those who reported suffering acne or another skin condition in the past 6 reported currently suffering from a skin condition, with an additional 7 who had not suffered from a skin condition in the past also currently suffering from a skin condition; from those using skin treatments 2 reported currently suffering from a skin condition while 3 reported suffering from one in the past.

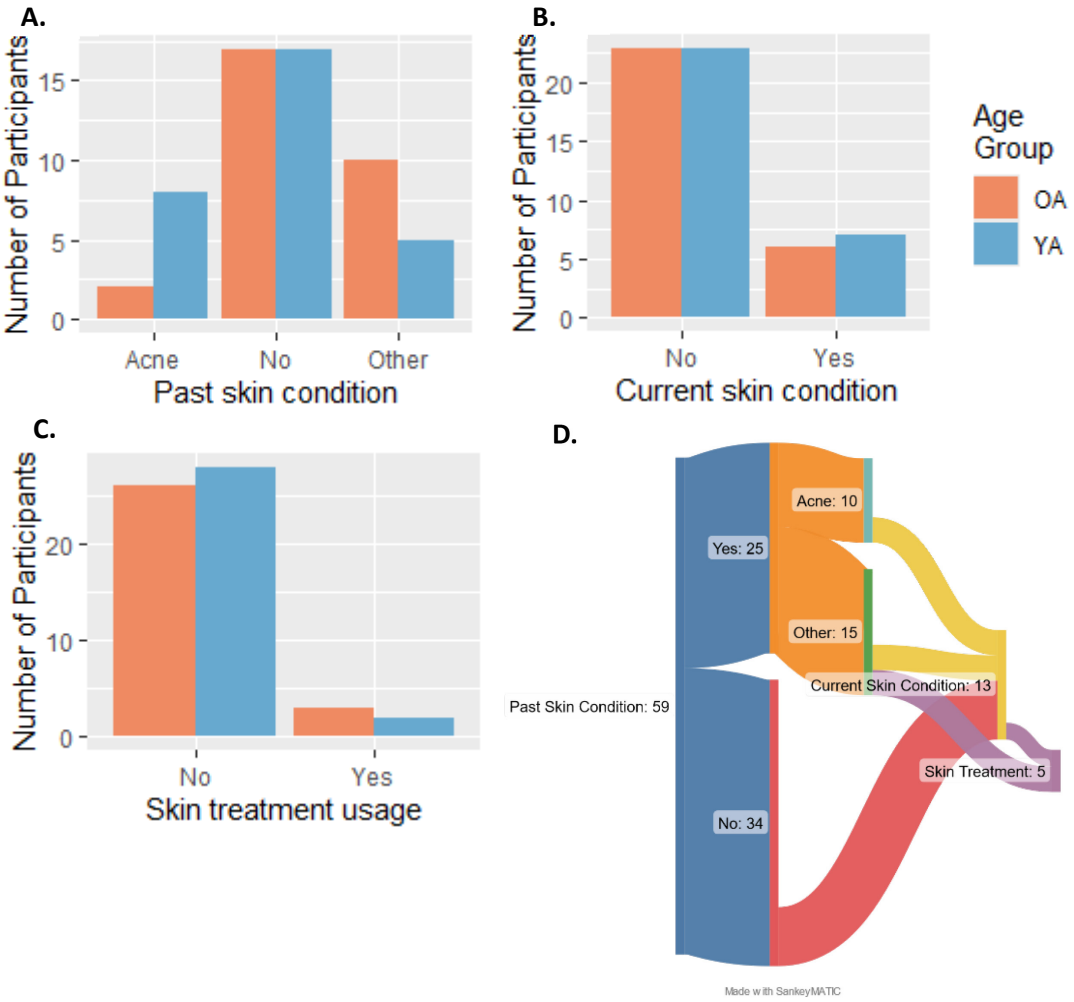


Figure 3.8 Participant skin health data. *Data divided by the OA and YA age groups, showing no significant differences. A. The number of participants who have suffered from skin conditions in the past (p=0.073). B. The number of participants who are currently suffering from a skin condition (p=1.000). C. The number of participants who are currently using some form of medicated skin treatment (p=0.671). D. A network diagram showing the relationship between individual participant responses to those covered in the previous 3 graphs.*

When looking into other aspects of participant health that may be relevant to the microbiome or ageing (Fig 3.9), 4 of the OA and 1 of the YA group reported currently undergoing hormone treatment while 1 YA and 1 OA reported that they are considered immune supressed. 2 OA and 1 YA reported that they were taking or had

been taking within the past 6 weeks of participating an antibiotic. When looking into participant health choices, 3 OA and 3 YA reported actively smoking, with 2 YA reporting that they both smoke and vape. A total of 13 YA and 10 OA participants reported following a range of specific diets, the highest reported diet for both groups were vegetarian with 7 YA and 3 OA members. When asked about their physical activity a total of 20 YA and 23 OA said they considered themselves an active person in some capacity, of these 4 from each group regularly swim in either a pool or open water. When investigated using a Fisher's exact test none of these aspects showed

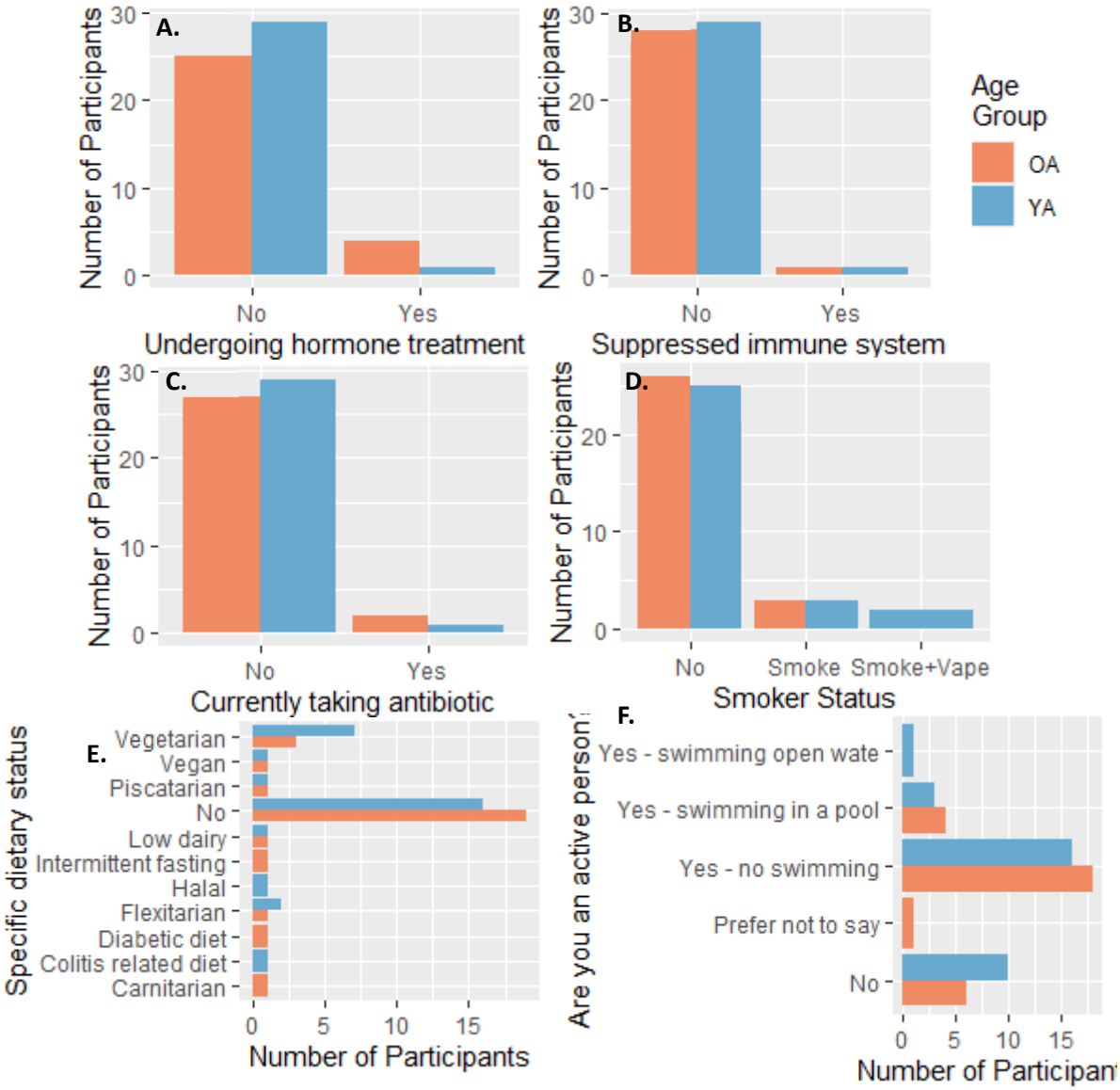


Figure 3.9 Participant health adjacent factors. *Data divided by the OA and YA groups showing no significant differences. A. The number of participants who are undergoing hormone treatments (p=0.195). B. The number of participants who have a suppressed immune system (p=1.000). C. The number of participants taking antibiotics (p=0.612). D. The number of participants who smoke or vape (p=0.706). E. The number of participants who are following specific diets (p=0.596). F. The number of participants who consider themselves active and what types of activity this involves (p=0.719).*

a significant difference.

Finally, it is possible to cross compare some of these health and sociodemographic factors. Comparing the participants with past skin conditions (Fig 3.10), by gender shows a significantly higher number of female participants reporting them than male ($p=0.005$), while there was an equal split between the male and female participants reporting acne. A detailed look of the OAs past skin condition status using the FI showed a significant difference between those who score a zero or one being more likely to not report a past skin condition ($p=0.021$), while those who score a two were more likely to report a past skin condition. Distributing the AUDIT-C scores of the participants by gender showed a significant difference between male and female participants ($p=0.017$), with female participants more likely to score zero. When the OA group AUDIT-C score was distributed via IMDD a significant difference was found ($p=0.036$), with the participants who live in lower deciles more likely to score lower on the AUDIT-C scale. Breaking down the activity habits of participants by Output Area Code showed a significant difference between those who don't consider themselves active living in 4A postcodes ($p=0.013$), while those who consider themselves active live in the 5B, 6A, and 6B postcodes. This concludes the investigations into the participants health using the responses to the questionnaire.

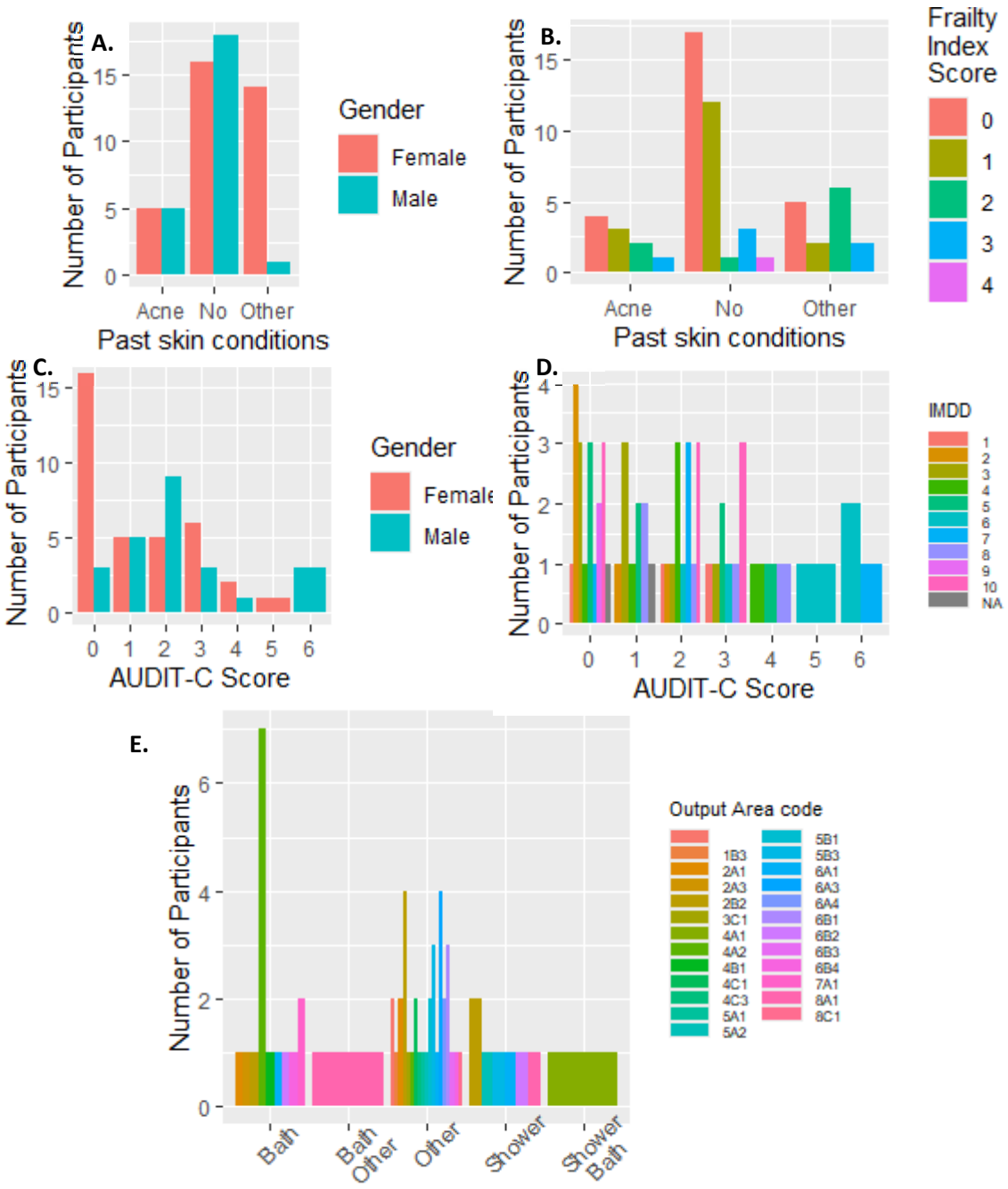


Figure 3.10 Comparisons of various participant health factors. *Results showing significantly different interactions. A. The number of participants who have suffered from skin conditions in the past, divided by gender (p=0.005). B. The number of OA participants who've suffered from skin conditions in the past, divided by frailty index (p=0.021). C. The AUDIT-C scores of the participants divided by gender (p=0.017). D. The AUDIT-C scores of the OA groups divided by Index of Multiple Deprivation Decile (IMDD) (p=0.036). E. The activity habits of participants divided by their Output Area Codes (p=0.013).*

3.3.4. Participant hygiene habits data

Along with various aspects of their health that could contribute to their microbiomes and ageing, participant habits around skin and oral hygiene could also play a role. Regarding participant dental hygiene habits (Fig 3.11), all participants reported that they brushed their teeth. Both groups reported using either a manual or

electric toothbrush, with 12 of the YA and 13 of the OA preferring a manual, verses 18 of the YA and 16 of the OA using an electric. In reporting daily toothbrushing frequency both the YA and OA group had 20 participants preferred twice a day, with only 2 in the YA group preferring to brush their teeth more frequently, 8 YA and 9 OA participants reported preferring to toothbrush once a day. In total 31 participants reported conducting some form of flossing activity, of these 6 OA and 5 YA prefer to floss, with 8 OA and 7 YA preferring to use interdental brushes, only 5 of the YA preferred to do both. Of those that performed a flossing activity 16 of the OA and 9 of the YA performed it once a day on average, while 2 in each group performed it twice a day and only 1 in the YA group performed it 3 times a day on average. There were 6 of the OA and 11 of the YA who reported as using a mouthwash regularly. Finally, 16 of the OA and 8 of the YA reported visiting a dental hygienist within the last 6 months of the participation sessions; only this aspect of dental hygiene showed a significant difference ($p=0.035$), when investigated using a Fisher's exact test.

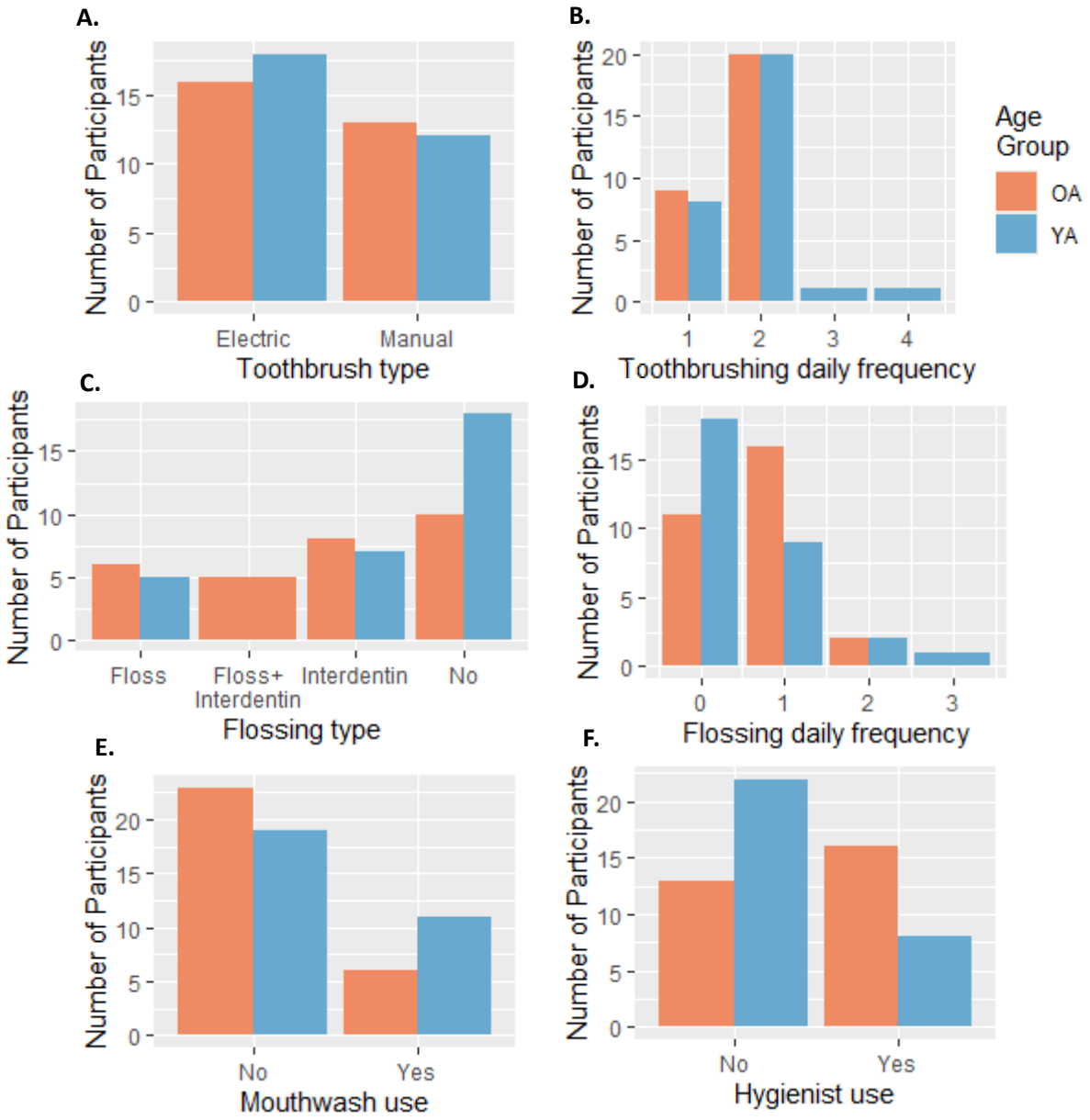


Figure 3.11 Participant dental hygiene habits. *Data divided by the age groups, showing a significantly different interaction in one metric. A. The number of participants who use different types of toothbrush (p=0.795). B. The frequency of daily toothbrushing by each participant (p=1.000). C. The number of participants who perform different types of dental flossing activities (p=0.057). D. The frequency of daily flossing activities caried out by participants (p=0.196). E. The number of participants who use mouthwash as part of their dental routine (p=0.252). F. The number of participants who have visited a dental hygienist within the last 6 months (p=0.035).*

When looking at participants skin hygiene habits (Fig 3.12), all participants reported following a regular bathing routine. Specifically, 20 of the OA and 25 of the YA reported a routine that involved a shower in some form, 3 of the OA preferred to use a bath in some from, 2 OA and 3 of the YA used both a bath and a shower regularly, aside from these 1 OA and 2 YA preferred to use a bathing method other than a bath or shower. The weekly bathing frequency varied from less than once a week on average to 10 times a week or more, the frequency reported by the most YA and OA participants was 7 times a week at 8 and 7 participants respectively, the second most preferred frequency was 4 days a week with 6 OA and 7 YA reporting this frequency. When

cleaning themselves while bathing 20 of the OA and 27 of the YA reported preferring to us a liquid shower gel or body wash, with a further 7 OA and 2 YA preferring a hard soap, only 1 YA and 1 OA prefer another detergent. As a final aspect of their skin routines participants were asked if they applied a moisturiser regularly, 15 of the OA and 20 of the YA reported that they did. When these hygiene habits were compared using a Fishers exact test none showed a significant difference.

Again, these hygiene habits can be investigated further using the other established factors in the participant population. Comparing moisturiser use by gender (Fig 3.13), shows a significant increase in its use by females compared to males ($p=0.007$). Breaking down moisturiser use by past skin condition suffers shows a significant increase ($p=0.015$), in likelihood of use among participants who have suffered from some form of skin condition. Contrasting bathing frequency by activity habit for both age groups did not show significant a difference; however, within the YA group a significant difference was observed ($p=0.007$), with those who consider themselves active being more likely to bathe seven times a week compared to those who do not consider themselves active are more likely to bathe 4 times a week. Dividing dental hygienist use by Output Area Codes showed a significant difference ($p=0.005$), between those who do use a hygienist are more likely to live in areas classified as 5B, 6A, and 6B, compared to those who live in 2B and 4A who are less likely to visit a dental hygienist. A significant difference in bathing method was shown when compared to IMDD ($p=0.039$), with showering being the most common out of all the deciles. Comparing manual and electric toothbrush preference by Index of Multiple Deprivation Decile (IMDD), has a significant difference ($p=0.007$), between those living in lower deciles being more likely to use a manual toothbrush compared to those who live in higher deciles being more likely to use an electric toothbrush. This concludes the review of the participants health hygiene habits data.

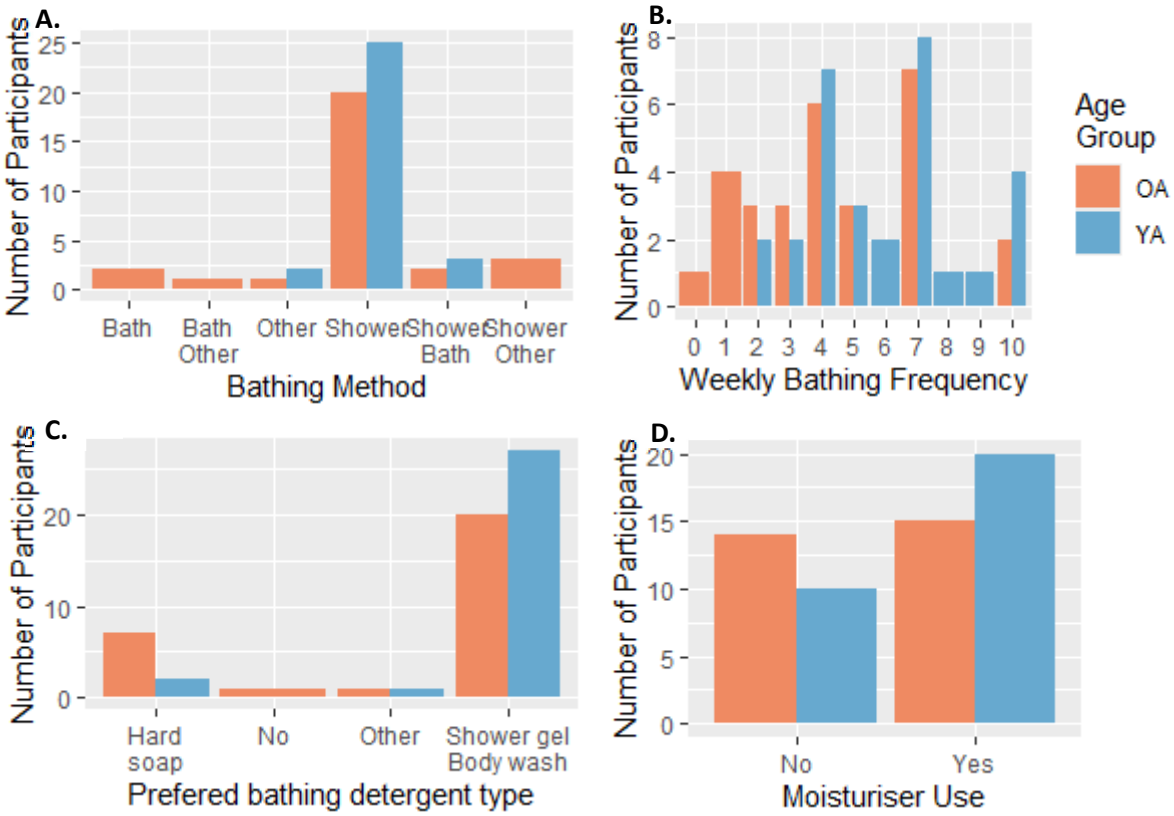


Figure 3.12 Participant hygiene habits. *Data divided by the age groups, showing no significantly different interactions. A. The bathing method carried out by the participants (p=0.321). B. The weekly bathing frequency of participants (p=0.475). C. The preferred detergent type per number of participants (p=0.097). D. The number of participants who use moisturiser on their skin (p=0.295).*

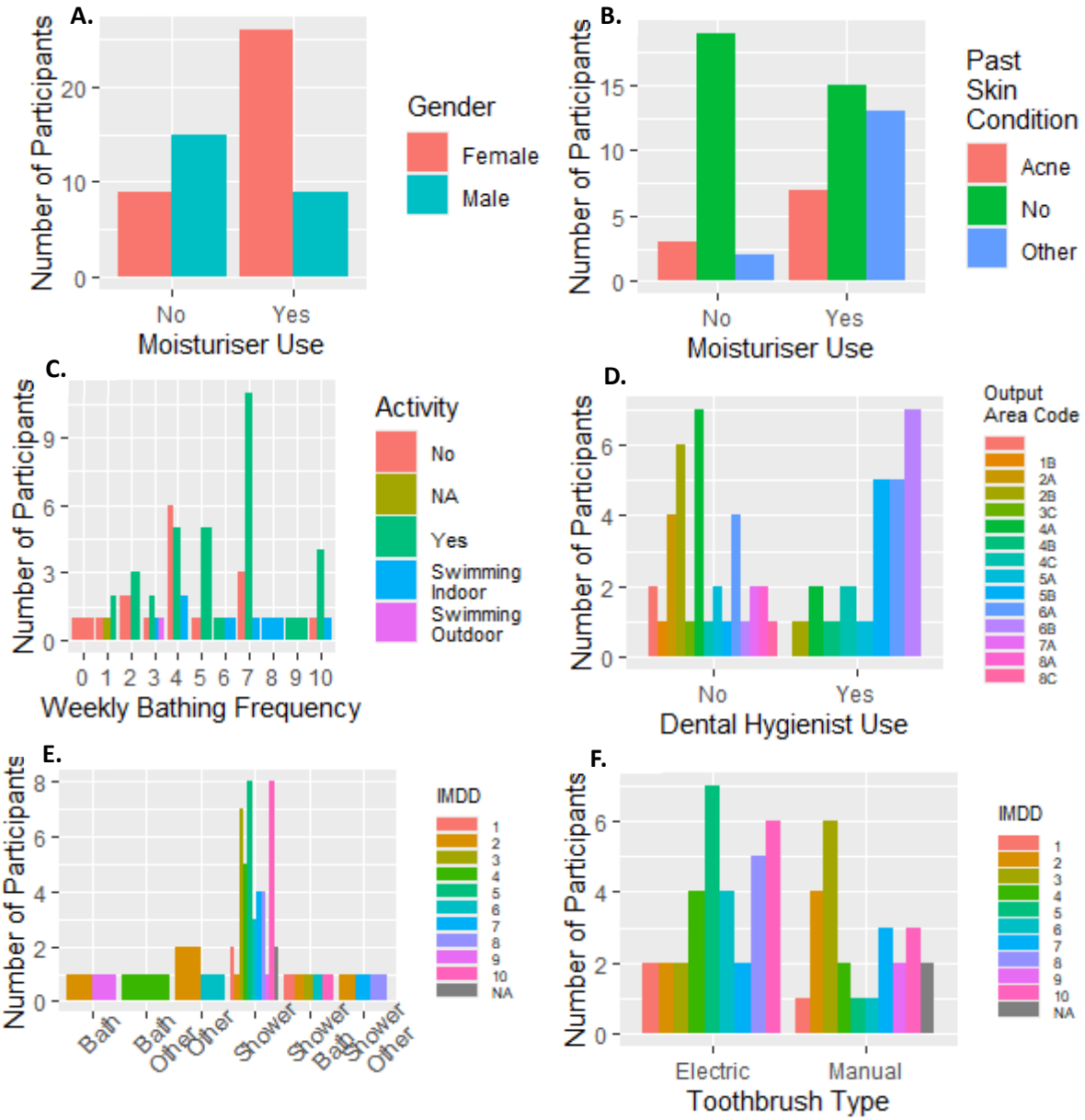


Figure 3.13 Comparisons of participant hygiene habits with other metadata factors. *Results showing some significantly different interactions. A. The number of participants who use moisturiser on their skin divided by gender (p=0.007). B. The number of participants who use moisturiser on their skin divided by past skin conditions(p=0.015). C. The weekly bathing frequency of participants divided by activity type (p=0.366). D. The number of participants who have visited a dental hygienist within the last 6 months divided by their Output Area Code (p=0.005). E. The bathing method carried out by the participants divided by Index of Multiple Deprivation Decile (IMDD) (p=0.039). F. The number of participants who use different types of toothbrush divided by Index of Multiple Deprivation Decile (IMDD)(p=0.007).*

3.3.5. Metadata Summary

Having analysed this metadata in detail it is possible to summarise this into table 3.2, which shows only six significant differences between the age groups when grouped by the 26 metadata variables explored here.

Table 3.2 Summary of the participant metadata split into the age groups. ***Including if a significant difference was found between them.***

Metadata	Factor	OA	YA	Significant difference
Participants		29	30	NA
Gender	Female	17	18	No
	Male	12	12	
Ethnicity	Asian	0	8	Yes
	Black	0	0	
	Multiple/Mixed	0	1	
	Other	0	2	
	White	29	19	
Frailty Index (FI)	0	7	19	Yes
	1	7	10	
	2	9	0	
	3	5	1	
	4	1	0	
AUDIT-C score	0	7	12	Yes
	1	4	6	
	2	7	7	
	3	6	3	
	4	3	0	
	5	1	0	
	6	1	2	
Non-wisdom teeth removed	No	6	19	No
	Yes	23	11	
Filling status	No	1	12	Yes
	Yes <6months	2	2	
	Yes >6months	26	16	
Tonsils removed	No	14	27	Yes
	Yes	15	3	
Past skin condition	Acne	2	8	No
	No	17	17	
	Other	10	5	
Current skin condition	No	23	23	No
	Yes	6	7	
Using a skin treatment	No	26	28	No
	Yes	3	2	
Undergoing hormone treatment	No	25	29	No
	Yes	4	1	
Suppressed immune system	No	28	29	No
	Yes	1	1	
Currently taking antibiotic	No	27	29	No
	Yes	2	1	

Smoker status	No	26	25	No
	Smoke	3	3	
	Smoke + Vape	0	2	
Specific dietary status	No	19	17	No
	Yes	10	13	
Active person status	No	7	10	No
	Yes	18	16	
	Yes-including swimming	4	4	
Toothbrush type used	Electric	16	18	No
	Manual	13	13	
Daily toothbrushing frequency	1	9	8	No
	2	20	20	
	3	0	1	
	4	0	1	
Flossing type	Floss	6	5	No
	Floss and interdentin	0	5	
	Interdentin	8	7	
	No	17	13	
Daily flossing frequency	0	11	18	No
	1	16	9	
	2	2	2	
	3	0	1	
Mouthwash usage	No	23	19	No
	Yes	6	11	
Dental hygienist use	No	13	22	Yes
	Yes	16	8	
Bathing Method	Bath	2	0	No
	Bath + other	1	0	
	Other	1	2	
	Shower	20	25	
	Shower + bath	2	3	
	Shower + other	3	0	
Weekly bathing frequency	0	1	0	No
	1	4	0	
	2	3	2	
	3	3	2	
	4	6	7	
	5	3	3	
	6	0	2	
	7	7	8	
	8	0	1	
	9	0	1	
	10+	2	4	
Preferred bathing detergent type	Hard soap	7	2	No
	No	1	0	
	Other	1	1	
	Shower gel/body wash	20	27	
Moisturiser use	No	14	10	No
	Yes	15	20	

3.4. Discussion

The aim of this chapter of this thesis was to contextualise and analyse the metadata gathered from the participants recruited in this study to determine if sub-groups are present within the population, as those that are present could be used to aid in the interpretation of microbiome analysis in chapters 5 and 6 of this thesis. Although the maximum number of participants were not recruited, both groups incorporated the minimum threshold of 25 participants required to give the study power; failure to achieve recruitment targets is common as an analysis of clinical trial recruitment found only 56% of UK studies achieved their target goals (Walters *et al.*, 2017). Satisfactorily, the age ranges of the participants within the groups meet the requirements, the median of these ranges is not the middle of these groups, however this is to be expected with such small group sizes. The gender divide demonstrated approximately 60% of both groups as female, however this does not reflect the division of gender in the population of Birmingham which is 51% female (Office for National Statistics, 2023). Similarly, neither age groups ethnic diversity matches that of Birmingham; the OA group being purely from white ethnicities has no diversity while the YA group has some members from the White, Asian, and other ethnicities, although much of this diversity is amongst the female members. This prevents comparisons between the ageing skin and oral microbiomes of the different ethnic groups as there are differences in the skin products and diets of these groups along with different attitudes to ageing (Dein and Huline-Dickens, 1997; Laidlaw *et al.*, 2010; Bai, 2014; Gupta, Paul and Dutta, 2017; Li *et al.*, 2019).

The significant difference between Output Area Codes of the YA and OA postcodes is somewhat expected due to how the majority of the OA group is retired and either bought their properties during more affluent times or have recently moved to such areas to enable an easier living standard, meanwhile the YA group are more likely to need to live within means and live in commutable areas. This supported by the 2B2 and 4A2 codes being defined as “Multicultural student neighbourhoods” and “Private renting new arrivals” respectively (supplemental table 9.3), these groups are associated with younger adults and especially in the 2B2 group were targets during recruitment; similarly, the OA groups 5B3 and 6A3 are defined as “Self-sufficient retirement” and “Detached retirement living” respectively, again these are expected due to our targets during recruitment.

Using the Super Output Area Codes yields a similar result where 1A and 7A are defined as “Cosmopolitan student neighbourhoods” and “Urban cultural mix” respectively for the YA group, and for the OA 8A is defined as “Suburban living - Affluent communities”. Whereas 3B is defined as “Ethnically diverse professionals – Asian traits”, it is possible that this difference is seen here due to the fewer codes of the Super Output Area’s compared to the Output Area’s causes more participants to be grouped up, or that as more members of these neighbourhoods sell their properties they are bought up by a more diverse group of people who are able to afford them (Clark and Coulter, 2015; Catney, Wright and Ellis, 2021).

A significant number of the OA group live in Workplace Zone code D2 which is defined as “Primarily residential suburbs” this reflects their previous classifications as living in affluent areas and therefore unlikely to be near industrial or business estates etc. Meanwhile the largest classification code for the YA group is B2 which is defined as “Administrative centres” which confirms their likelihood to live in either inner cities or near business parks. While it is unlikely that the participants work in the industries reported using the Workplace Zone codes, they do provide information regarding the potential exposure to pollution which is a factor in ageing and will also have an effect on their microbiomes (Wang *et al.*, 2021).

There were no significant differences when the YA and OA groups were clustered by Local Authority code. Both groups have their highest clustering in the 4A1 group which is defined as “Ethnically diverse metropolitan living”, this is followed by a cluster of both groups in the 7 codes which are defined as “Town and Country living”. These differences in grouping compared to the Output Area and Super Output codes are likely caused by the differing definitions and metrics used to create them. Overall, these 4 different methods of defining and clustering participants based on using their postcodes to create sociodemographic data should enable limited comparisons between the microbiome and ageing data, although the number of participants within these groups is unlikely to be sufficient to reach power for each case, they will be able to shed light on possible new avenues for future research (House, Kessler and Herzog, 1990; Pac *et al.*, 2019).

Using the English Indices of Deprivation deciles and its sub deciles to categorise participants it is possible to infer that the YA group are more likely to live in deprived areas compared to the OA group. When considered along with the Output area classifications this is supported as an inner-city area like Birmingham where the majority of the YA group are living is more likely to be considered deprived than the suburban areas the OA group live in. Care must be taken when using the English Index of Multiple deprivation to infer sociodemographic information as it can only be used to compare areas and not people and it is best used as a relative measure of relative deprivation between iterations. Such a measure of relative change is not possible as participants did not provide information regarding duration of occupation at their current address or previous addresses, therefore it cannot be established if they have experienced the changes in sociodemographic status that these areas will have over time. However this information is still relevant to creating an overall picture of participants socioeconomic status as it reflects on living and health care standards, which are factors in both the ageing process and microbiome (Lang *et al.*, 2008; Berthoud, Blekesaune and Hancock, 2009).

The significant difference between the YA and OA groups when scored on a FI further ensures that comparisons between the microbiomes of these groups is valid. While it is expected that a few of the YA group would score beyond the 0 on the FI this reflects more on the limitations of the indexing system rather than the health of the participants, in future work this frailty indexing system should also be combined with a more accurate molecular testing system based off participant biomarkers (Woo, Leung and Morley, 2012; Dent, Kowal

and Hoogendijk, 2016; Diebel and Rockwood, 2021; Behr *et al.*, 2023). Differences between the ageing of a population based on gender are not seen here as there is no significant difference between the FI score of the OA population, however this may be due to the small sample size or a limitation of the frailty indexing system where individual conditions are overlooked for the purpose of scoring (Barer, 1994; Kane and Howlett, 2021; Yokoro *et al.*, 2023). A further limitation of the FI in this project is that due to the variation in location for participation sessions and reduction in participant burden, some factors such as measuring participant balance, Body Mass Index (BMI), and , mini-mental state exam was not possible so has been omitted from the indexing system; this may have led to some participants scoring lower on the FI than they would actually score.

When alcohol consumption was compared using the AUDIT-C scoring system a significant difference was observed between the YA and OA group, with the OA group frequently scoring higher while the YA group is more frequently scoring lower. This difference between the age groups is likely a reflection of the general trend of a reduction on alcohol consumption in younger generations compared to older generations (Kraus *et al.*, 2020). Similarly the difference male and female participants may be due to the increased diversity of the female YA members compared to the male, this could cause the cultural differences of the ethnicities to be more influential (Sudhinaraset, Wigglesworth and Takeuchi, 2016). Overall, these differences in alcohol consumption and the range of scores present will enable comparisons between the participants microbiomes to be made along with the influences on the ageing process, some of this is already know in regarding the influence on the gut microbiome and Alzheimer's disease (Moreno-Arribas *et al.*, 2020; McMahan *et al.*, 2023; Ward *et al.*, 2023).

In general, the differences seen between the dental health of the YA and OA groups is expected, older people are more likely to need teeth extractions that younger and similarly older adults will be in more need of the other dental interventions such as crowns and dentures due to having experienced more damage over a longer period of time (Brown, Thomas and Blake, 2018; Hiltunen and Vehkalahti, 2023). The reduction in tonsil removal seen between the OA and YA group is due to the tonsillectomy surgery being seen as unnecessary outside of severe cases within the last 50 years (Dwyer-Hemmings, 2018; Wilson *et al.*, 2023). Although there were no significant differences observed when the oral health data is analysed using the sociodemographic data, however, this does not mean that it has not played a role in creating the oral health profiles seen here, it is more likely that this study has too few participants or did not cover enough data about participant sociodemographic factors (Delgado-Angulo, Mangal and Bernabé, 2019). This information regarding participant oral health will be useful when investigating the differences in the oral microbiome during ageing as oral health is a factor in both general health and healthy ageing, the oral microbiome is also shaped by the conditions inside the oral cavity (Duran-Pinedo and Frias-Lopez, 2015; Brown, Thomas and Blake, 2018; Chowdhry *et al.*, 2023).

In both groups only a small number of participants reported either in the past or currently suffering from some form of skin condition. As the study did not recruit directly for skin conditions and given the relatively small

number of participants the lack of significant difference between these groups is understandable but the proportion who currently have a skin disease does match expected statistics at approximately 20% of the population (Rea, Newhouse and Halil, 1976; Hay *et al.*, 2014). However, a there is a general trend for the OA group to have had some condition other than acne in the past, when analysed by gender there is a significant difference between the OA females reporting more skin conditions than the males, this is likely due to societal factors regarding attention to skin condition between males and females (Kandrack, Grant and Segall, 1991). Similarly, more of the YA participants reported having acne in the past compared to the OA group, this may be due to the YA group having a more recent memory of their adolescence or again a societal shift in the acknowledgement and understanding of the condition. This data regarding skin conditions is important for the interpretation of the overall skin microbiome composition as some skin conditions are either the result of or contributed to by microbiome dysbiosis, similarly the ingredients in these skin treatments could be influencing its composition (Rodrigues Hoffmann, 2017). However, these data do little to contribute towards the understanding of skin ageing and in future studies would be complemented with other measures such as skin thickness and firmness tests, a scale of photoaging can also be created and used to categorise the participants (Farage, Miller and Maibach, 2016).

When looking into the participants other health associated factors only a small number of participants in the OA and YA group reported undergoing hormone treatments, suffering from an suppressed immune system, or taking antibiotics; there was no significant difference between these groups in these measures, however this information is important to know as these participants will be on medications which will alter both the oral and skin microbiome compositions and thusly will need to be considered when interpreting this data (Vich Vila *et al.*, 2020; Gong *et al.*, 2021). Again, while no significant differences were seen in the smoker status, dietary status, or activity status. However, there is a slight increase in the YA group preferring to follow a specific dietary restriction which may be due to different societal values of food consumption between the groups (Parmenter, Waller and Wardle, 2000; Dean *et al.*, 2009). Inversely more of the OA group consider themselves active this may be due to these participants either being more health conscious or having more free time to partake in physical activities (Allender, Hutchinson and Foster, 2008; Bélanger, Townsend and Foster, 2011). This information is important as these choices affect participant health and contribute to microbiome status, interpretation of these data in relation to the participant microbiomes may indicate a requirement of future studies into these factors.

By comparing between the health factors and sociodemographic status of participants the nuances of these relationships can be demonstrated. The significant difference between past skin condition suffers by gender could be due to cultural attitudes towards skin disease awareness or how females tend to be more likely to suffer from long term conditions while males are more likely to have short term infections (Chen *et al.*, 2010). The significant difference observed in the OA groups past skin conditions is unexpected, however the small number of participants in this study prevent any strong conclusions about this link from being drawn as it may

just be a simple correlation without any cause. The gender differences in alcohol consumption seen here are expected as males often drink more than females, however in this study this may be swayed due to the different ethnicities of the age groups and attitudes towards alcohol consumption between them as no significant difference was seen when compared with these groupings (Wilsnack *et al.*, 2009). Similarly the difference in alcohol consumption and IMDD is expected in the OA group, this is not universal to both groups as younger adults have shown recent reductions in drinking activities when compared to older adults, which is supported here by the YA not having a significant difference when compared in this way (Collins, 2016; Visontay *et al.*, 2020; Steffen *et al.*, 2021). The differences between activity and output area code demonstrate the link between income level and physical activity level which is enhanced by the extra leisure time possible on higher incomes, as physical fitness is a factor in healthy ageing this difference between these populations is important for this study (Puciato *et al.*, 2018; Stalsberg and Pedersen, 2018).

The dental hygiene habits used by the participants are diverse with all participants choosing to brush their teeth daily. While there are no significant differences between the toothbrush types, daily brushing frequency, flossing type and frequency, and mouthwash use between the groups there are still some trends that may become apparent with a larger sample size, for instance an electric toothbrush is more popular than a manual for both groups or there appears to be a trend for the OA group to prefer to perform some type of flossing activity. The significant difference between the OA group using a dental hygienist more frequently than the YA group is likely down to proportional disposable income levels between the groups as a hygienist appointment is a private procedure which costs more than a simple check-up. Overall, it is relevant to know what oral health habits participants are carrying out as these will contribute towards both oral health in ageing and the composition of the oral microbiome, for example those that brush their teeth and floss frequently are disrupting oral biofilms more regularly than those who do not (Selwitz, Ismail and Pitts, 2007). Along with this participants also provided information regarding their preferred brand of toothpaste and mouthwash if applicable, for simplicity this data has not been shown here, however the ingredients used in these products will also have an impact on the oral microbiome, particularly if any are antimicrobial in nature (Haraszthy, Raylae and Sreenivasan, 2019; Hong *et al.*, 2020).

Participants provided a wide range of responses for their regular bathing method, with a shower being the most popular, there is also a wide range in weekly frequency of bathing with the popular answers relating to once a day or once every other day. However, there is no significant difference between any of these groups. Differences may still be seen at a microbiome level as the different types of bathing result in changes in how the skin is cleaned and similarly to the oral microbiome the more frequently bathing is performed the more the microbes on the skin will be moved or removed. Similarly, there is no significant difference between the types of bathing detergent used or moisturiser use at this level, however these will affect the skin microbiome similarly to how toothpaste etc. shapes the oral microbiome, as exposure to the ingredients will inhibit or encourage the

growth of species/strains (Wallen-Russell, 2018). There is a significant difference in moisturiser use when divided upon gender, this remains the case in the OA group specifically but not in the YA group, this probably reflects the different societal attitudes to skin product use between males and females of the age ranges here.

The interplay of these habits with other participant factors can help enhance our understanding of the reasoning behind the choices made by the participants and thus be used to understand how they may be influencing the participants ageing process or composition of the microbiome. The increased use in moisturiser by females over males (Fig 3.13), is a known and documented difference which results from cultural views on skin habits (Ertel *et al.*, 2009; Biesterbos *et al.*, 2013). Similarly the increased moisturiser use by participants who have suffered from skin conditions in the past is understandable due to the proven effectiveness of moisturiser at aiding in repair and prevention of some skin diseases (Carville *et al.*, 2014; Hon *et al.*, 2015). The weekly bathing frequency of participants is compared to activity has a positive correlation; this will be due to the increased sweating caused by intense physical activity as such is may also be possible to use this data to infer how frequently participants who consider themselves active perform physically intensive activities. The increased use of dental hygienists among those who live in wealthier output area codes is again reflective of the costs of the procedure, meaning further evidence that participants who are more likely to have increased disposable income to spend on the procedure. The significant clustering seen when participant bathing methods by IMDD is likely due to the popularity of showering when compared to the other methods overall, this may be due to the limited size of the study population and a was not something participants were recruited based on. However, comparing the toothbrush type used by participants by their IMDD does show that those who live in lower income areas are more likely to use a manual toothbrush which is cheaper compared to an electric toothbrush which is more preferred in the higher income areas, this factor is important when considering its effects on participant dental health, thusly effecting their ageing health and may be responsible for the composition of the oral microbiome (Nikias, Fink and Shapiro, 1975; Foley and Akers, 2019).

3.5. Conclusion

In summary this chapter has contextualised the metadata gathered from the participants during the recruitment and participation process. By putting it under deep scrutiny it identified not just differences in age between the participant groups. These included differences in socioeconomic status, overall frailty, and alcohol consumption which are increased in the older adults. In terms of specific oral health, the older adults were shown to have poorer health but be more invested in it than the younger adults, whereas in skin health no differences were discovered. Finally, comparisons of skin and oral hygiene habits yielded no significant differences based on age but did reveal skin moisturiser use was driven by gender. Therefore, this establishes that not only have sufficient participants were recruited to give the study statistical power to compare the effects of ageing on the

skin and oral microbiomes, but that there are other significant differences between the age groups which can be utilised in the analysis of the participant microbiome samples in chapters 5 and 6 of this thesis.

Chapter 4. Sample Sequencing Pipeline Optimisation

4.1. Introduction

Sequencing of the variable regions of the 16S RNA gene allows for taxonomic classification of the source bacterium whether from a pure colony or a mixed microbiome sample. Sequencing and computational processing methods will affect the accuracy of the classification and for microbiome work the population proportions reported (Golob *et al.*, 2017; Siegwald *et al.*, 2019; Jeske and Gallert, 2022). In research groups that regularly carry out standardised microbiome analysis or those that follow already established and published protocols these methods will be consistently applied with minor variations taken due to individual study requirements (Davis-Turak *et al.*, 2017). Here as this is the first microbiome study at Aston University using skin and oral samples sequencing the different variable regions a bespoke optimisation of the pipeline is appropriate. To aid this pipeline creation a culture-based screen of the samples was carried out in tandem, to provide an idea of the species to expect within the sequence data and as a measure of microbiome health, as the ability to classify down to the genus level was established in chapter 2.

Polymerase-based sequencing methods lack the additional error checking enzyme architecture that exists within organisms resulting in a loss in read quality towards the end of each strand. In second generation sequencing methods this issue is countered by sequencing both the forward and the reverse strand of DNA with a sufficient overlap to enable the creation of a consensus sequence in silica (Pfeiffer *et al.*, 2018; Ma *et al.*, 2019). Additionally, shotgun sequencing a whole genome allows for varying fragments to be created from the genome which enables multiple overlaps to be created further reducing ambiguity within the consensus sequence, in short read 16S RNA gene sequencing however this is not possible as only a single part of a single gene is targeted for sequencing. This places more burden on obtaining high sequence quality and the sequencing processing tools.

Briefly as outlined in figure 4.1, after sequencing the primer sequences and other known contaminants such as adapter sequences or short reads of unknown bases should be removed to clean the data and enable the correct calculations for final sequence length prior to the next stage of merging and denoising. The merging and denoising stage creates a consensus sequence for classification ideally this sequence is the same as that present within the samples. Classification and construction of a phylogenetic tree from the consensus sequences enables the dataset to be interpreted in terms of its microbiological ecology. For optimising these stages of this study, the participant age groups, and other factors covered previously will not play a role, instead the location

of the sample and variable region sequenced are important as they will contribute, contaminants and quality issues, or affect the location of the sequence overlap.

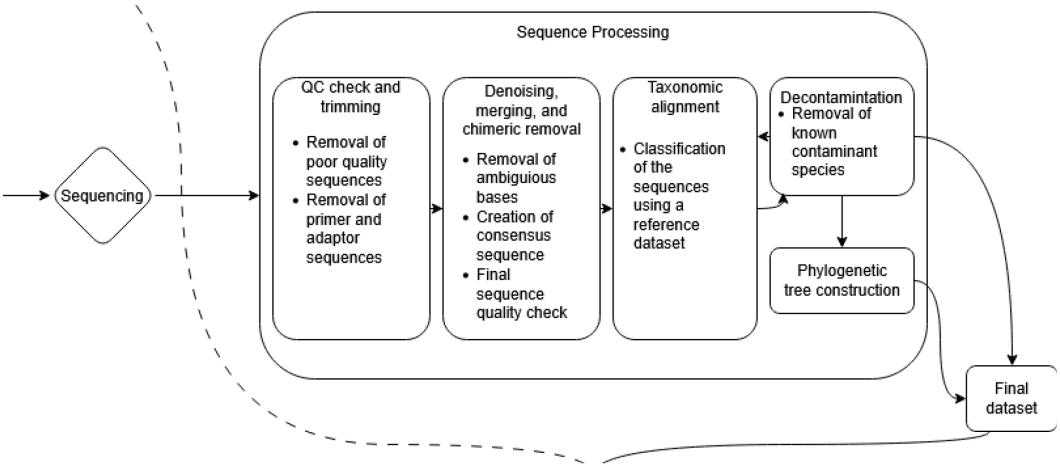


Figure 4.1 Excerpt of Figure 1.5. *Focusing on the sequence processing stage of amplicon based 16S microbiome studies.*

4.1.1. Aims

- Describe the preparation of the samples for sequencing and an overview of the final sequence datasets.
- Create an optimised pipeline that recovers the highest number of sequences from the raw sequence data.
- Develop an understanding of the sequence data that may impact future interpretation of the finalised datasets.

4.2. Methods

4.2.1. Culture based screening

During initial sample processing (outlined in 3.2.3.3), 100µl aliquot of each arm and oral sample was inoculated by lawn spreading onto different agar and incubate for the times listed in table 4.1 bellow. The resulting plates were screened for growth and some initial classification was made, presumptive *Staphylococcus* isolates recovered from MSA were further screened by being inoculated onto DNAse agar and incubated at 37^{O}C for 18 hours.

Table 4.1 Different media and their associated incubation temperatures and times for growth screening from the arm and oral samples.

Agar	Incubation temperature (^{O}C)	Incubation time (hrs)	Inoculum source
Mannitol Salt Agar (MSA)	37	36	Arm and Oral
MacConkey Agar No3 (MAC3)			
Sabouraud dextrose agar (SDA)	30	72	Arm
Sabouraud dextrose agar + Clotrimazole (SDAC)			Oral
Colorex™ Candida (CXCA)	37	36	
Streptococcal selective Columbia blood agar (SCBA)			

4.2.2. DNA Extraction

The oral samples were filtered through sterile disposable filter funnels with a 0.2µm filter membrane (Nalgene, UK). To extract microbial DNA these membranes were then used in the standard protocol of the DNeasy PowerWater Kit (Qiagen, Germany). Both arm and face samples were extracted using DNeasy PowerSoil Pro Kit (Qiagen, Germany), the standard protocol was modified by using the aseptically removed swab head in place of soil in the protocol. Blanks of each sample type were created as described previously (3.2.3.1.), with a swab head soaked in the sterile PBS and Tween® 20 solution used for the skin samples, and an unused sterile DH_2O aliquot used for the oral blank. These blanks were then subject to the previously described DNA extraction. Each DNA extract was then stored at -80^{o}C until required for further processing.

4.2.3. DNA Processing

The absorbance for each DNA extract was measured on a NanoDrop 1000 spectrophotometer (Thermo Fisher, UK), across 230-340nm. DNA extracts were then dried using a Concentrator plus (Eppendorf, Germany), set to V-AQ at 30^{o}C for 45 minutes.

4.2.4. Sequencing

The dried DNA extracts were shipped to Colgate-Palmolive (Piscataway, NJ, USA), for sequencing. Sequencing was carried out on an Illumina MiSeq system (Illumina, USA) at Rutgers University (USA). For all

samples the adaptor sequence CTGTCTCTTATACACATCT was used. For the arm and face samples the primer sequence used was 28F- GAGTTTGATCNTGGCTCAG 519R-GTNTTACNGCGGCKGCTG (Targeting V1-V3), while for the oral sample the primers were 347F-GGAGGCAGCAGTRRGGAAT 803R-CTACCRGGGTATCTAATCC (Targeting V3-V4). These should result in sequences of 490 and 450 base pairs long respectively.

4.2.5. Initial Sequence Processing

The forward (R1), and reverse (R2), sequences for each sample were received in a fastq.gz file format. These were uncompressed using 7-Zip 21.07 software to become a fastq file format.

Each file was assessed using the FastQC tool (Simon Andrews, 2023), these FastQC reports were then compiled by MultiQC (Ewels *et al.*, 2016), into one report for the R1 and R2 reads for each sample type to enable comparison.

4.2.6. Qiime2

The sequences were imported into a Qiime2 (Bolyen *et al.*, 2019), environment (version 2024.5), using the import data function, they were imported as the Casava 1.8 paired end demultiplexed fastq format. An initial overview of the sequences was constructed using the demux plugin with its summarise function.

4.2.6.1. Fastq Trimming and Optimisation

Within the Qiime2 environment the plugin Cutadapt (Martin, 2011), was used to remove the primer sequence from the fastq files along with other known sequence contaminants or low quality sequences with the trim-paired function. The parameters for primer overlap and Phred quality score (A measure of the uncertainty of the correct base being reported), were optimised using the variations listed in table 4.2, no adaptor sequence was included in the fastq files received from sequencing, so this sequence was not included in the Cutadapt parameters; however, the relevant primer sequences were included. The other parameters used were a minimum length of 1, match wildcard reads as true (to accommodate the wildcards present in the primer sequence), and no untrimmed sequences were discarded; these parameters were chosen to maximise sequence retention (key for the low quality of some samples), all other parameters were kept at their defaults.

Table 4.2 Parameters of the different variations of the Cutadapt plugin trialled for optimisation. ***Run ID is the name of the variation trialled.***

Run ID	Primer Overlap	Phred Quality Score
CA1	5	30
CA2	10	30
CA3	15	30
CA4	15	20
CA5	15	10

For computational efficiency parameter CA3 was used for merging optimisation and for maximum sequence recovery parameter CA5 was used for generation of the finalised sequence datasets.

4.2.6.2. Sequence denoising and merging optimisation

The DADA2 plugin (Callahan *et al.*, 2016), was then used to denoise, merge, and dereplicate the sequences within the samples; using the denoise paired function. Here optimisation of the merging of R1 and R2 (The forward and reverse sequences) required variation of each sequence length. Due to the variation in the total sequence length sequenced between the skin and oral samples a different optimum length was required for both. Therefore table 4.3 shows the parameters trialled in the skin sample sets while table 4.4 shows the parameters trialled in the oral sample set. The other parameters in DADA2 adjusted were a minimum overlap of 20, and a pooling method of pseudo, all other parameters were kept at their defaults for the version number used.

Table 4.3 Parameters of the different trimming/truncation variations trialled in DADA2 on the skin samples.*Run ID is the name of the variation trialled, the other columns are the names of the DADA2 parameter functions altered. The parameters used for the final dataset are highlighted in green.*

Run ID	trim-left-f	trim-left-r	trunc-len-f	trunc-len-r
Skin0v0	0	0	0	0
SkinF0v5	0	5	0	0
SkinF0v10	0	10	0	0
SkinF0v15	0	15	0	0
SkinF0v20	0	20	0	0
SkinF0v25	0	25	0	0
SkinF0v30	0	30	0	0
SkinF5v0	5	0	0	0
SkinF10v0	10	0	0	0
SkinF15v0	15	0	0	0
SkinF20v0	20	0	0	0
SkinF25v0	25	0	0	0
SkinF30v0	30	0	0	0
SkinF5v5	5	5	0	0
SkinF10v10	10	10	0	0
SkinF15v15	15	15	0	0
SkinR0v230	0	0	0	230
SkinR260v230	0	0	260	230
SkinR270v230	0	0	270	230
SkinR280v200	0	0	280	200
SkinR280v220	0	0	280	220
SkinR280v240	0	0	280	240
SkinR280v250	0	0	280	250
SkinF10R280v230	0	10	280	230

Table 4.4 Parameters of the different trimming/truncation variations trialled in DADA2 on the oral samples. *Run ID is the name of the variation trialled, the other columns are the names of the DADA2 parameter functions altered. The parameters used for the final dataset are highlighted in green.*

Run ID	trim-left-f	trim-left-r	trunc-len-f	trunc-len-r
Oral0v0	0	0	0	0
OralF0v5	0	5	0	0
OralF0v10	0	10	0	0
OralF0v15	0	15	0	0
OralF0v20	0	20	0	0
OralF0v25	0	25	0	0
OralF5v0	5	0	0	0
OralF10v0	10	0	0	0
OralF15v0	15	0	0	0
OralF20v0	20	0	0	0
OralF25v0	25	0	0	0
OralF5v5	5	5	0	0
OralF10v10	10	10	0	0
OralF15v15	15	15	0	0
OralF20v20	20	20	0	0
OralF25v25	25	25	0	0
OralR230v190	0	0	230	190
OralR240v200	0	0	240	200
OralR250v200	0	0	250	200
OralR250v210	0	0	250	210
OralR250v220	0	0	250	220
OralR250v230	0	0	250	230

4.2.6.3. Construction of taxonomic classifier and taxonomic classification

Separate custom classifiers were made for the skin and oral datasets using the appropriate primers. Due to the heavy computational nature of this work, it was carried out on Aston University's Engineering and Physical Sciences college Machine Learning Server, funded by the EPSRC Core Equipment Fund, Grant EP/V036106/1. The Qiime2 plugins feature-classifier (Bokulich *et al.*, 2018), and RESCRIPt (Ii *et al.*, 2021), were used to construct the classifiers. The skin classifier was constructed using the Ribosomal Database Project (RDP), classifier training dataset No19 (Wang and Cole, 2024). While the oral classifier was constructed using the Human Oral Microbiome Database (HOMD), 16S rRNA RefSeq Version 15.23 dataset (Andrew Voorhis *et al.*, 2023).

Briefly, for the skin classifier the RDP dataset was culled of degenerate or homopolymer sequences, and then dereplicated to only retain unique sequences with their relevant taxonomies. The sequencing primers were used to extract the region from within the dataset within 480 to 520 bp long and truncated to 280 bp, these were further dereplicated before being used to construct a classifier with a 90% alignment confidence. For the oral classifier the same process was carried out using a sequence region between 390 to 490 bp long, truncated to 250 bp from the HOMD dataset. These classifiers were used on the merged sample datasets using the classify-

sklearn function, for the skin samples a confidence threshold of 90% used while in the oral samples a confidence threshold of 80% was used.

4.2.6.4.. Filtering and decontamination of datasets

The now classified sample datasets required filtering to remove known contaminants, prior to undergoing microbiological ecology analysis. Initially the filter-seqs and filter-taxa function were used to remove any sequences and taxa classified as Eukaryota, Chloroplast, Mitochondria, or unassigned. The decontam (Davis *et al.*, 2018), plugin was used in conjunction with the blank samples to identify other contaminants and remove them. Then the filter-features function was used to remove any sequence which occurred less than 10 times in total. Finally, the filter-seqs-length function was used to remove any sequences below 390 bp for all datasets.

4.2.6.5. Phylogeny tree construction

The phylogeny plugin was used to construct a phylogeny tree for each filtered dataset with the align-to-tree-mafft-fasttree function using the default parameters.

4.3. Results

4.3.1. Organism Recovery

After incubation of aliquots from the arm samples successful growth was seen on the majority of MSA and SDA, but only a small fraction of MAC3 plates, where more growth positive plates were seen in the YA group (Fig. 4.2). When colonies isolated from the MSA plates was screened on DNAse agar the majority of samples were positive for DNAse activity, a small number of samples in each group were either negative or did not grow on the DNAse agar. Two participants (YA01 and YA27), samples showed no growth upon any of the agars. This demonstrates that *Staphylococcus* species are expected to be present in the sequence data of the arm samples.

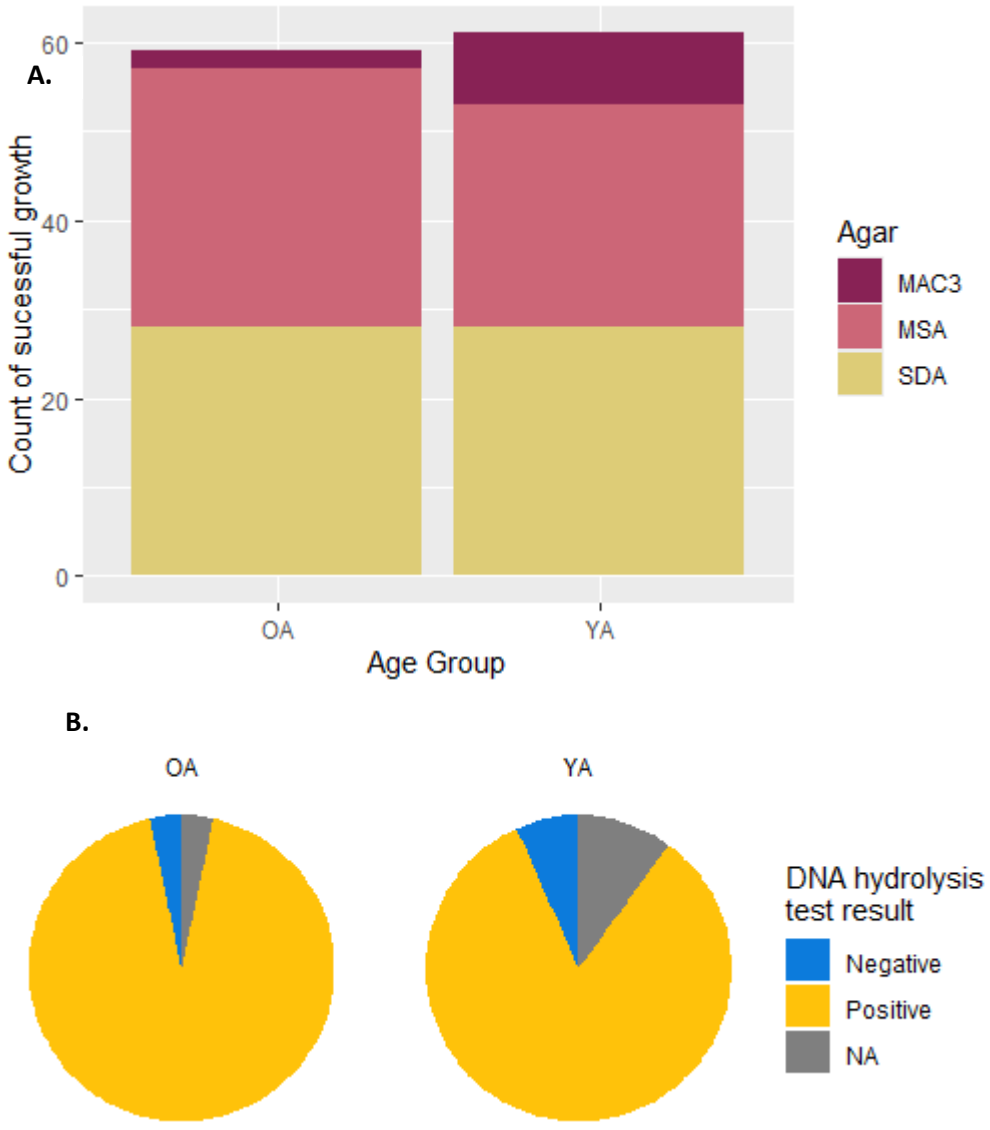


Figure 4.2 The results of culture-based screening of the arm samples split by age group. *Data showing no significant differences. A. The counts of successful growth on MAC3, MSA, and SDA agar. B. The proportion of positive, negative and non-applicable results when a colony from the MSA agar was grown on a DNAse agar plate*

After incubation of the aliquots from the oral samples successful growth occurred across all agars, the lowest growth was in the CXCA in which growth occurred for 12 YA and 14 OA samples, for nearly all samples growth occurred on the MSA and SCBA (Fig. 4.3). The colonies isolated from each MSA plate when incubated on DNAse agar showed a mixture of negative and positive results for DNA hydrolysis, the majority of both age groups were negative. The CXCA provide a colour to each colony to enable a presumptive species identity to be allocated, using this the majority of samples that successfully contained candida were identified as *Candida albicans*, with *Candida auris*, and *Candida glabrata* also seen. When subjected to statistical analysis no significant difference was seen in these results between the age groups. This demonstrates that *Staphylococcus* and *Streptococcus* species are expected to be present in the sequence data of the oral samples.

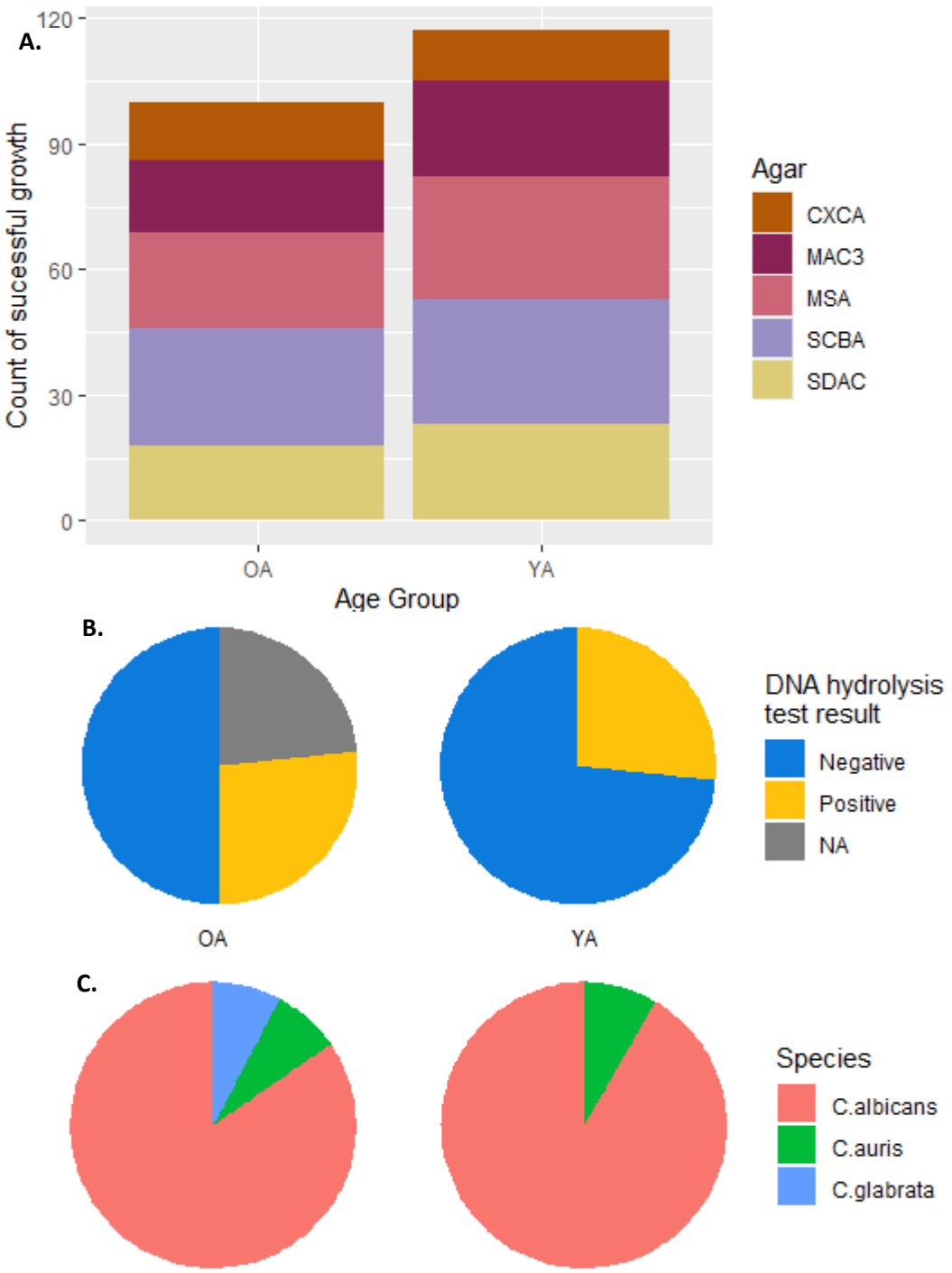


Figure 4.3 Graphs showing the results of culture-based screening of the arm samples split by age group. *Results showing no significant differences. A. The counts of successful growth on CXCA, MAC3, MSA, SCBA, and SDAC agar. B. The proportion of positive, negative and non-applicable results when a colony from the MSA agar was grown on a DNAse agar plate. C. The presumptive species identified on the CXCA after successful growth.*

4.3.2. Nucleotide Quality

The DNA extracts each had their nucleotide concentration measured prior to drying and subsequent sequencing; this information assisted the technicians when preparing the extracts for sequencing. When the DNA extracts for the arm samples were measured using a nanodrop spectrophotometer a range of nucleotide concentrations was found (Fig. 4.4), spanning from 0.67 ng/µl to 57.94 ng/µl, with an average of 5.07 ng/µl. The 260/280 ratio for the arm samples covered -130.92 to 395.97, with an average of 10.73, only 3 of the samples have ratios in the negative. The 260/230 ratio for these samples ranged from 0.01 to 0.5, with an average of 0.1. The face sample DNA extracts (Fig. 4.5) reported a nucleotide concentration of 0.83 ng/µl to 14.35 ng/µl, with a 3.04 ng/µl average. The 260/280 ratio on these samples covers -85.53 to 12.87, with an average of -1.45. The 260/230 ratio for the face samples range -1.12 to 0.66 with an average of 0.1. Finally, the oral samples DNA extracts (Fig. 4.6) have a concentration range of 1.78 ng/µl to 295.35 ng/µl, with an average of 55.6 ng/µl. the 260/280 ratio covers 0.97 to 2.31 with an average of 1.78. While the 260/230 ratio spans 0.14 to 2.25 with an average of 1.4. This gives the oral samples a higher concentration on average than the skin samples, along with a more consistent nucleotide quality, only thorough analysis of the sequence information can reveal the effect this may have on the recovered sequences.

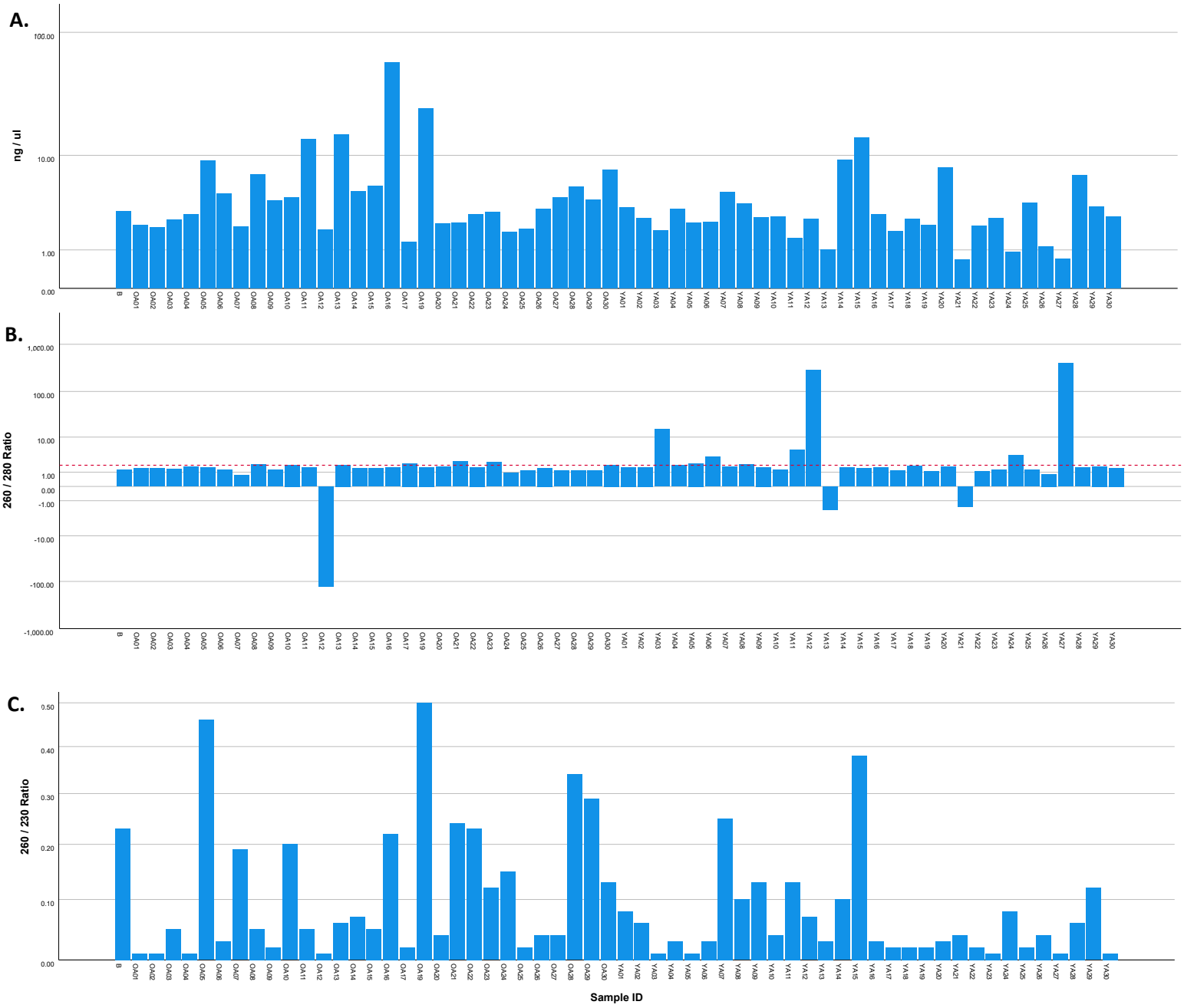


Figure 4.4 Quality information from the DNA extracts of the arm samples for the participants and blank. *A. The concentration of nucleotide in ng/µl. B. The 260/280 absorbance ratio with a marker line for a ratio of 1.8 (red line), of optimal DNA purity. C. The 260 /230 absorbance ratio, as a measure of contaminants.*

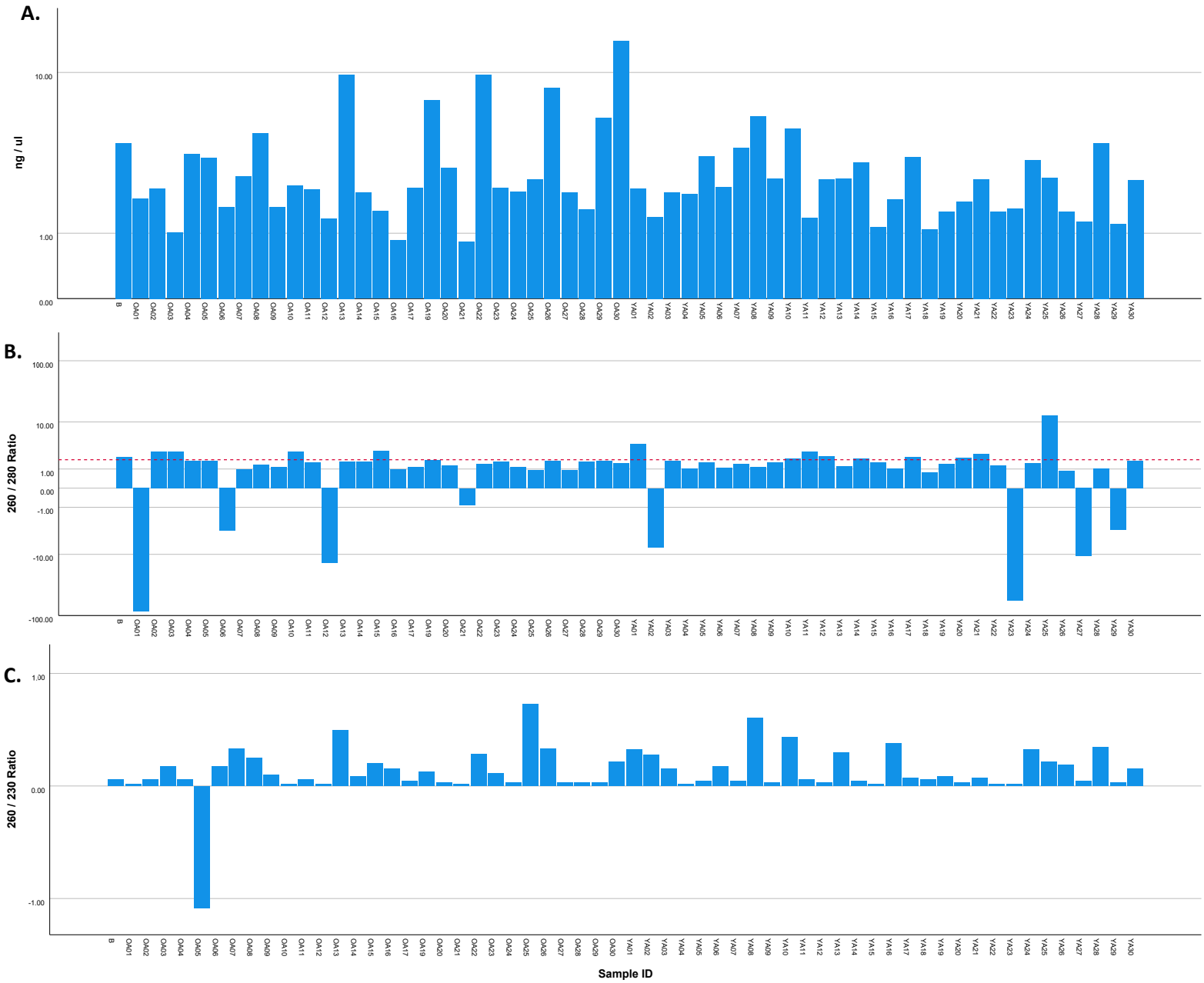


Figure 4.5 Quality information from the DNA extracts of the face samples for the participants and blank. *A. The concentration of nucleotide in ng/µl. B. The 260/280 absorbance ratio with a marker line for the ratio of 1.8 (red line), of optimal DNA purity. C. The 260/230 absorbance ratio, as a measure of contaminants.*

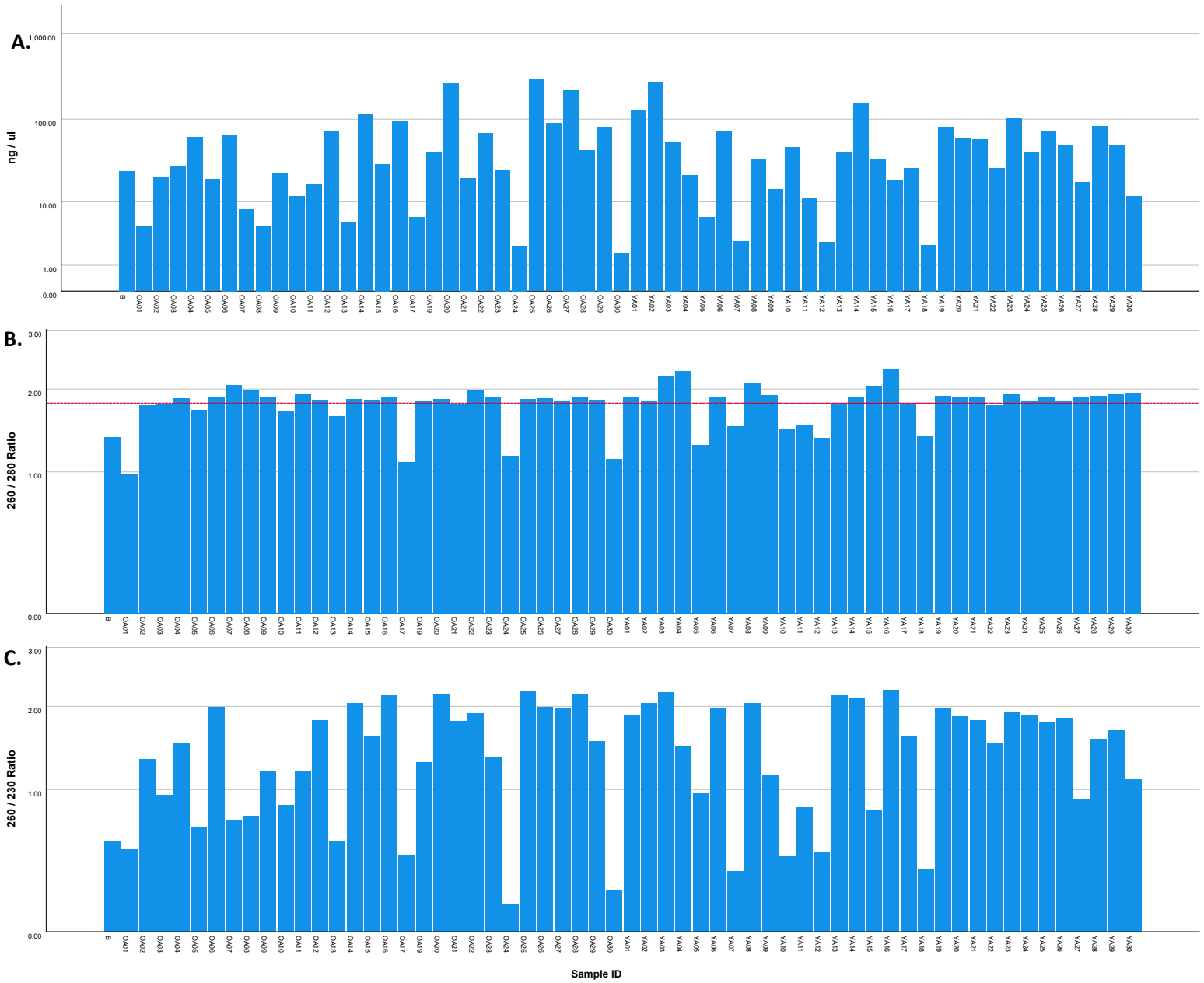


Figure 4.6 Quality information from the DNA extracts of the oral samples for the participants and blank. *A. The concentration of nucleotide in ng/µl. B. The 260/280 absorbance ratio with a marker line for a ratio of 1.8 (red line), of optimal DNA purity. C. The 260/230 absorbance ratio, as a measure of contaminants.*

4.3.3. Sequence Overview

Once the sequences were received and the files were imported into Qiime2 an initial overview of the sequences was generated. The arm and face datasets both had samples that had very low reads 14 and 13 respectively, compared to the oral samples with the lowest read count of over 150000 (Table 4.6); these results however belonged to different participants, OA01, OA20, and YA14 respectively. The maximum number of reads present within a sample was over 350000 for the arm, 630000 for the face, and 29000 for the oral samples; again, these results belonged to different participants, YA02, OA16, and YA13 respectively. Overall, they reported similar medians and means with the arm and oral reads being over 200000 and the face over 230000. When viewed as histograms (Fig.4.7), the arm and face sample minimum sequence counts can be seen as outliers from the main cluster of sequence counts. Similarly, the maximum read counts for all the samples were also distinct from the majority of sequence counts. This overview while informative regarding the successful recovery of sequences from all samples it is limited, a more in-depth picture can be developed by analysing the sequence quality information using Phred scores.

Table 4.5 Summary of raw sequence counts information of all sample types.

	Arm	Face	Oral
Minimum	14	13	153905
Median	203172.5	230364.0	205884.0
Mean	206957	2367723	209296
Maximum	358774	637489	297471
Total	12417434	14206376	12557740

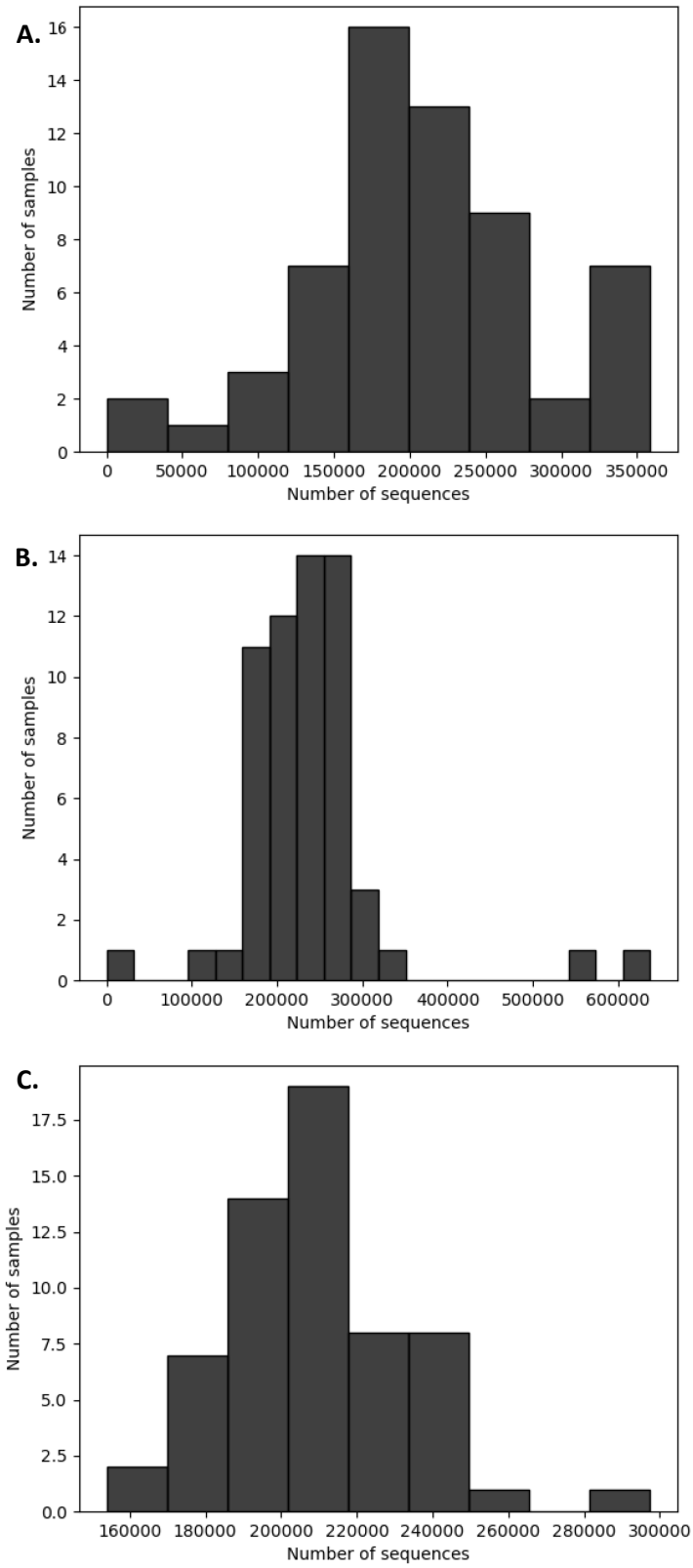


Figure 4.7 Histograms of raw sequence counts from each participant sample and blanks. *A. The arm samples. B. The face samples. C. The oral samples.*

When the Phred score information was compiled through FastQC and MultiQC the majority of the arm samples and all within the face and oral samples forward reads were deemed to have been passed (Fig.4.8),

meaning that the most frequently observed mean quality score within the sample had a Phred score of 27 or above. Meanwhile the majority of all samples reverse reads were deemed to have failed, meaning that the most frequently observed mean quality score within the sample had a Phred score of 20 or below. However, generally the mean of the reverse reads did not drop below the Phred threshold of 27 till after approximately 220bp. By comparison aside from the failed reads, the forward reads did not begin to drop below this threshold till approximately 290bp if at all. Overall, the arm samples appeared to have the most diverse mean quality scores across the read length compared to the face and oral which were consistent. Combining this quality information with the other overview of the raw sequences, meant that trimming optimisation could take place.

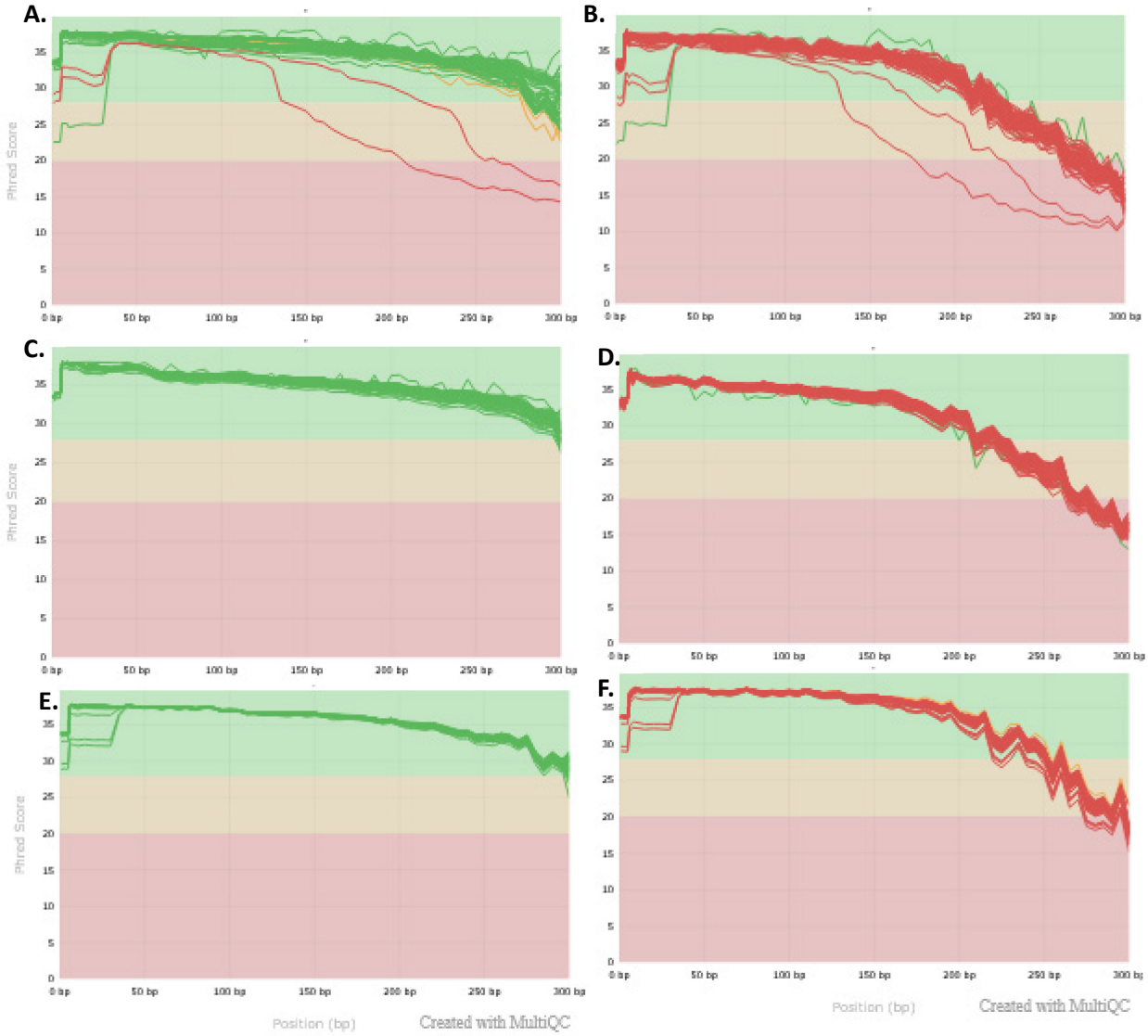


Figure 4.8 Mean Phred quality score for each sample over sequence length.*Constructed using FastQC and MultiQC, green represents a mean score above 27, yellow between 20-27, and red below 20. A. The arm samples forward reads. B. The arm samples reverse reads. C. The face samples forward reads. D. The face sample reverse reads. E. The oral samples forward reads. F. The oral sample reverse reads.*

4.3.4. Sequence Trimming Optimisation

The Cutadapt plugin was used to remove the sequencing primers, known sequence contaminants, and low-quality sequences. Due to the differences in pre-sequencing nucleotide quality across the sample types the different variations in the primer overlap and Phred quality score within Cutadapt were trialled on all the sample types individually.

4.3.4.1. Arm

The variations in Cutadapt parameters trialled (Outlined in table 4.2) on the arm fastq sequences all reduced the sequence counts by all metrics compared to the input values (Table 4.6). However, no variation reduced the minimum sequence count by more than 5 reads. There is little difference between the median, mean, maximum, or total generated from the run variations CA1, CA2, and CA3; all are lower than CA4, or CA5.

Table 4.6 Summary of sequence count information for the arm samples processed through variations in Cutadapt. *A colour scale is applied to each row going from green as the largest number and red the lowest.*

	Input	CA1	CA2	CA3	CA4	CA5
Minimum	14	9	9	9	9	9
Median	203173	199975	199975	199975	202479	202862
Mean	206957	202839	202839	202839	205622	205877
Maximum	358774	352057	352057	352057	357063	357426
Total	12417434	12170346	12170345	12170350	12337293	12352634

When the variations trialled on the arm samples were examined using the quality graphs, the highest quality seen was in the CA3 variation (Fig 4.9), compared to the CA4 and CA5 in both forward and reverse reads. Specifically comparing CA3 to the raw sequence quality all of the forward reads had a Phred score above the 27 success threshold and only a fraction of the reverse reads were now below the 20 failure threshold. As the quality score decreased in the variations trialled so did the resulting quality in the graphs, with the CA5 quality resembling the initial raw sequence quality, however the end of the sequence quality had still improved in both forward and reverse reads.

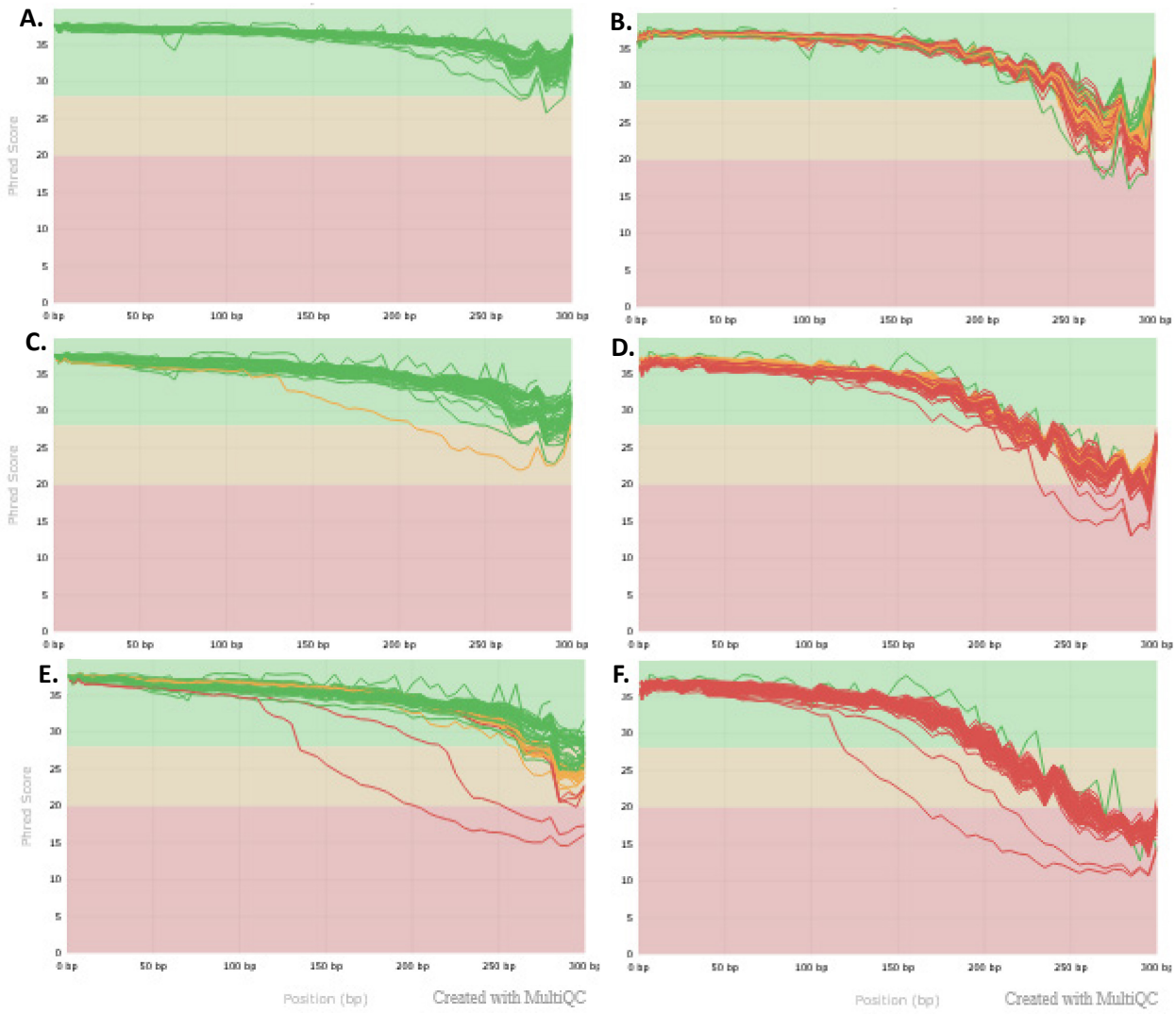


Figure 4.9 Mean Phred quality score for the arm samples over sequence length after Cutadapt variations length. *Constructed using FastQC and MultiQC, green represents a mean score above 27, yellow between 20-27, and red below 20. A. The CA3 variation forward reads. B. The CA3 variation reverse reads. C. The CA4 variation forward reads. D. The CA4 variation reverse reads. E. The CA5 variation forward reads. F. The CA5 variation reverse reads.*

4.3.4.2. Face

The variations in Cutadapt parameters trailed (Outlined in table 4.2) on the face fastq sequences all reduced the sequence counts by all metrics except for the minimum sequence count when compared to the input values (Table 4.7). There was little difference between the median, mean, maximum, or total generated from variations CA1, CA2, and CA3; all are lower than CA4, or CA5.

Table 4.7 Summary of sequence count information for the face samples processed through variations in Cutadapt. *A colour scale is applied to each row going from green as the largest number and red the lowest.*

	Input	CA1	CA2	CA3	CA4	CA5
Minimum	13	13	13	13	13	13
Median	230364	228567	228567	228567	232807	233123
Mean	236773	233313	233313	233313	237262	237551
Maximum	637489	626751	626751	626752	636898	637369
Total	14206376	13765480	13765481	13765483	13998434	14015489

When the variations trialled on the face samples were examined using the quality graphs, the highest quality seen was in the CA3 variation (Fig. 4.10), compared to the CA4 and CA5 in both forward and reverse reads. Specifically comparing CA3 to the raw sequence quality all of the forward reads had a Phred score above the 27 success threshold and only a small fraction of the reverse reads were below the 20 failure threshold. As the quality score decreased in the variations trialled so did the resulting quality in the graphs, with the CA5 quality resembling the initial raw sequence quality, however the end of the sequence quality had still improved in both forward and reverse reads.

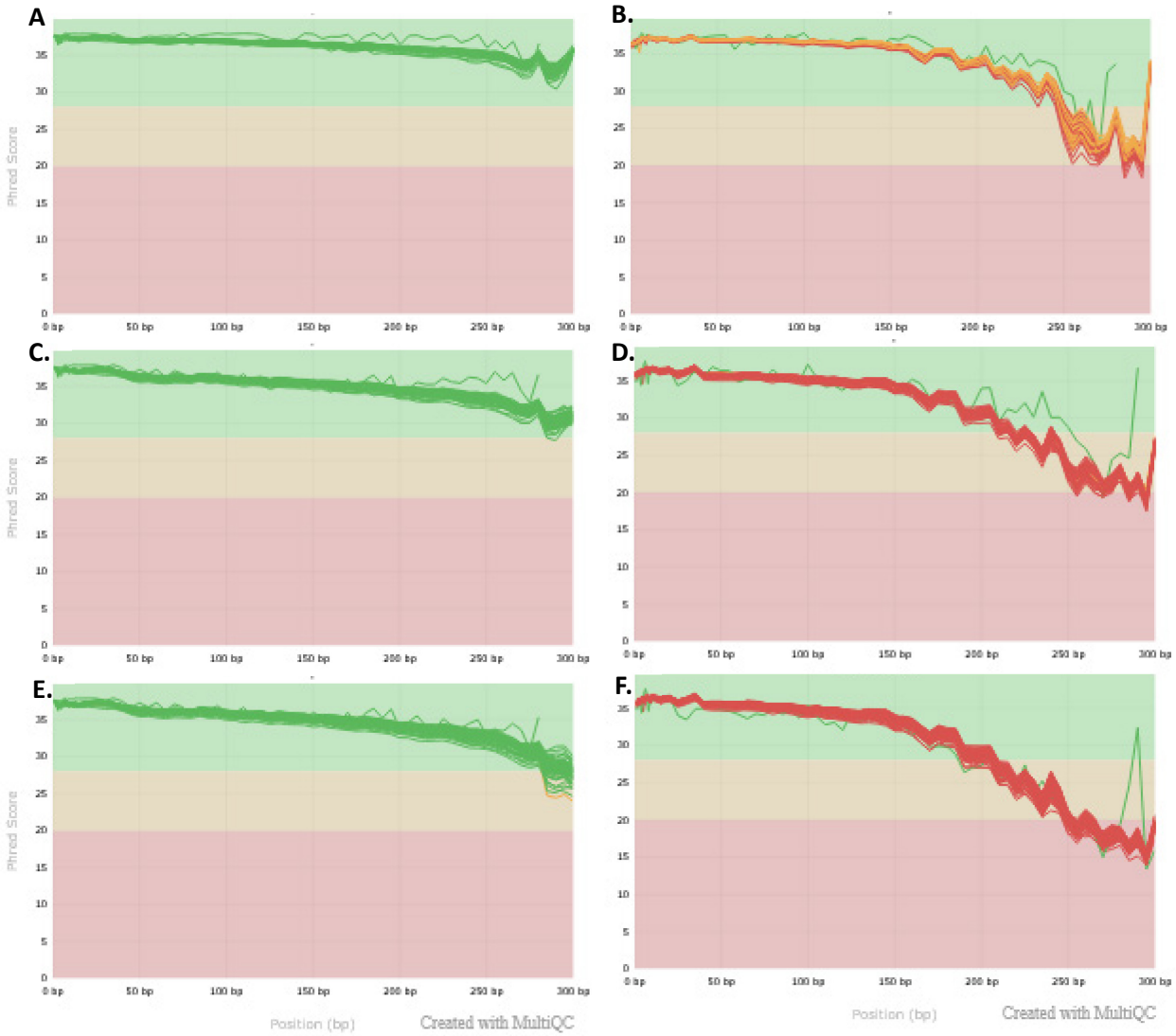


Figure 4.10 Mean Phred quality score for the face samples over sequence length after Cutadapt variations length. *Constructed using FastQC and MultiQC, green represents a mean score above 27, yellow between 20-27, and red below 20. A. The CA3 variation forward reads. B. The CA3 variation reverse reads. C. The CA4 variation forward reads. D. The CA4 variation reverse reads. E. The CA5 variation forward reads. F. The CA5 variation reverse reads*

4.3.4.3. Oral

The variations in Cutadapt parameters trialled (Outlined in table 4.2) on the oral fastq sequences all reduced the sequence counts by all metrics compared to the input values (Table 4.8). There was no difference between all metrics generated from variations CA1, CA2, and CA3; all were lower than CA4, or CA5.

Table 4.8 Summary of sequence count information for the oral samples processed through variations in Cutadapt. *A colour scale is applied to each row going from green as the largest number and red the lowest.*

	Input	CA1	CA2	CA3	CA4	CA5
Minimum	153905	152353	152353	152353	152684	152732
Median	205884	204377	204377	204377	205113	205147.5
Mean	209296	207005	207005	207005	207589	207622
Maximum	297471	294971	294971	294971	295556	295591
Total	12557740	12420320	12420320	12420320	12455362	12457291

When the variations trialled on the oral samples are compared using the quality graphs, the highest quality seen was in the CA3 variation (Fig 4.11), compared to the CA4 and CA5 in both forward and reverse reads. Specifically comparing CA3 to the raw sequence quality all of the forward reads had a Phred score above the 27 success threshold and only a small fraction of the reverse reads were now between the 20 failure and 27 warning threshold. As the quality score decreased in the variations trialled so did the resulting quality in the graphs, with the CA5 quality resembling the initial raw sequence quality, but the end of the reverse reads was lower than the raw sequence before rising back up slightly, similarly a few samples in the forward reads entered the warning threshold.

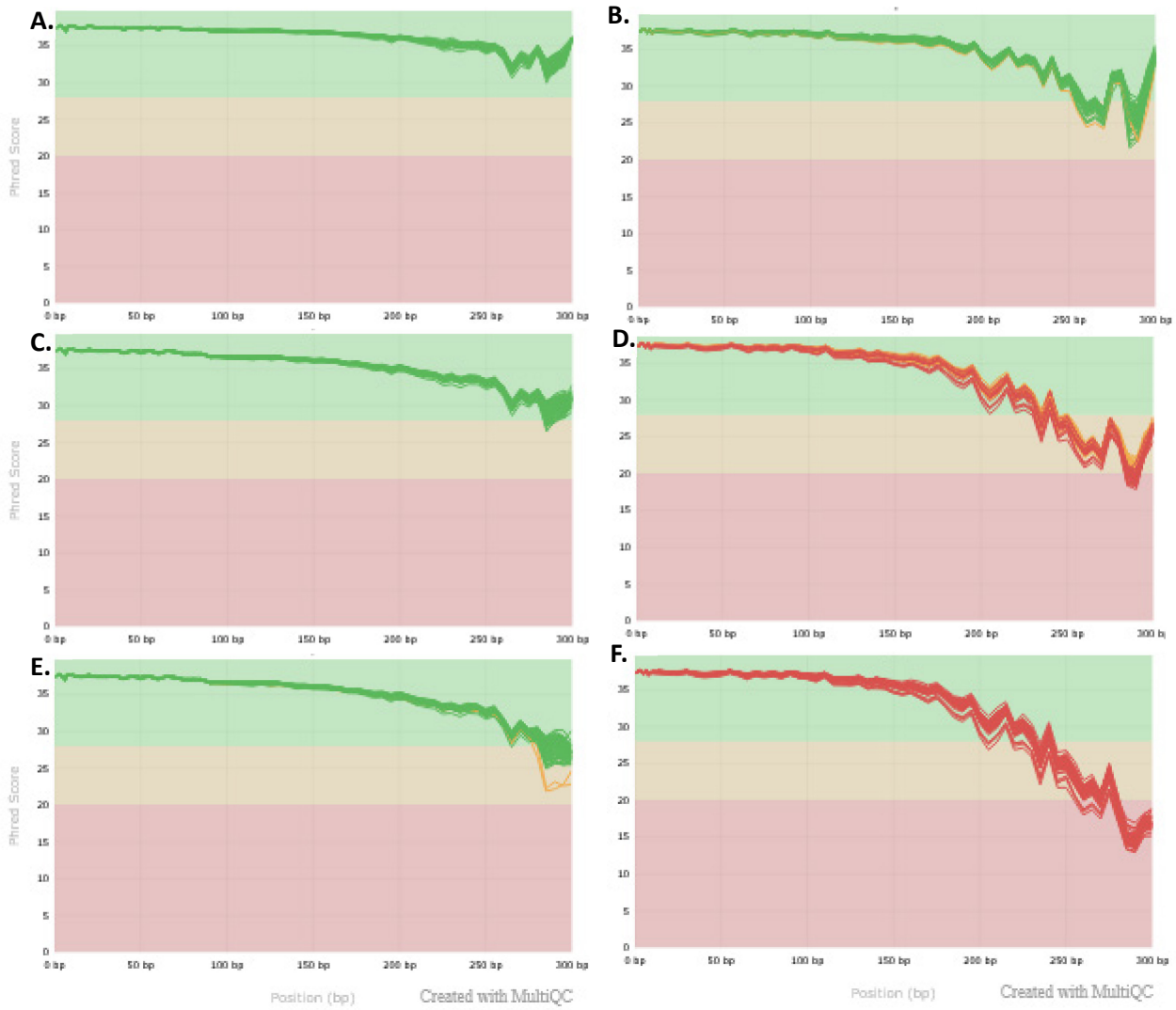


Figure 4.11 Mean Phred quality score for the oral samples over sequence length after Cutadapt variations length. *Constructed using FastQC and MultiQC, green represents a mean score above 27, yellow between 20-27, and red below 20. A. The CA3 variation forward reads. B. The CA3 variation reverse reads. C. The CA4 variation forward reads. D. The CA4 variation reverse reads. E. The CA5 variation forward reads. F. The CA5 variation reverse reads.*

4.3.5. Denoising and Merging Optimisation

The DADA2 plugin was used to denoise, merge, and dereplicate the sequences within the samples; this was optimised by trimming extraneous nucleotides from either end of the sequences prior to merging. Due to the differences in sequence quality across the sample types and the different regions amplified for the skin and oral samples the different variations in nucleotide removal were trialled on all the sample types individually. For computational efficiency the CA3 Cutadapt variation was used during these optimisations, for additional efficiency trims made to the end of sequences were trialled using inputs that had reads below the level of trim trialled removed.

4.3.5.1. Arm

The variations of the arm samples that focus on trimming the front of the sequences yield around 12000 unique sequences with a total frequency of over 3 million (Table 4.9), these showed an average merged percentage between 27-30%. While the trimming of the ends of the sequences yields between 2000-7000 unique sequences with a total frequency between 70 thousand and 1 million, however the percentage merged ranges from 6-85%. Across all there was a variation in the minimum frequency ranging from 0 to 5, this was from the sample with the lowest read count OA01.

Table 4.9 Summary of select merged sequence information for the arm samples processed through variations in DADA2. ***A colour scale is applied to each column going from green as the largest number and red the lowest.***

Run ID	Number of unique sequences	Total sequence frequency	Minimum sequence frequency	Mean sequence frequency	Mean percentage merged
Variations with reads of all length.					
Skin0v0	12006	3348848	5	55814	29.2645
SkinF0v5	12237	3413909	0	56899	28.45833
SkinF0v10	12382	3396699	5	56612	29.37217
SkinF0v15	12199	3415464	0	56924	28.42267
SkinF0v20	11643	3364509	0	56075	27.36883
SkinF0v25	12502	3355476	0	55925	28.0155
SkinF0v30	12391	3342380	0	55706	27.9155
SkinF5v0	11910	3323689	5	55395	28.83033
SkinF10v0	11817	3315144	5	55252	28.69417
SkinF15v0	11907	3362848	4	56047	28.66117
SkinF20v0	11522	3340111	4	55669	27.85033
SkinF25v0	12061	3389560	2	56493	28.63633
SkinF30v0	11877	3372044	2	56201	28.49833
SkinF5v5	12166	3400016	0	56667	28.34017
SkinF10v10	12278	3377480	5	56291	29.168
SkinF15v15	12119	3393224	0	56554	28.02967
Variations where reads below trim level have been removed.					
SkinR0v230	5971	896653	2	14944	59.02983
SkinR260v230	3501	556348	0	9273	32.76083
SkinR270v230	4528	708765	0	11813	43.34267
SkinR280v200	2094	75337	0	1256	6.707
SkinR280v220	4381	674852	0	11248	40.13133
SkinR280v240	6674	1275025	2	21250	84.4345
SkinR280v250	6710	1269492	2	21158	83.492
SkinF10R280v230	4393	680567	0	11343	40.16667

4.3.5.2. Face

The variations of the face samples that focused on trimming the front of the sequences yielded around 20000 unique sequences with a total frequency of over 6 million (Table 4.10), these showed an average merged percentage between 45-46%. Trimming of the ends of the sequences yielded between 3000-9000 unique

sequences, with a total frequency between 80 thousand and 1 million, however the percentage merged ranged from 5-85%. Across all samples there was a variation in the minimum frequency ranging from 0 to 2, this is in the from the sample with the lowest read count OA20.

Table 4.10 Summary of select merged sequence information for the face samples processed through variations in DADA2. *A colour scale is applied to each column going from green as the largest number and red the lowest.*

Run ID	Number of unique sequences	Total sequence frequency	Minimum sequence frequency	Mean sequence frequency	Mean percentage merged
Variations with reads of all length.					
Skin0v0	20637	6139530	0	104060	46.14102
SkinF0v5	21045	6130279	0	103903	46.11034
SkinF0v10	21438	6066780	0	102827	45.68373
SkinF0v15	21414	6111942	0	103592	46.07678
SkinF0v20	21292	6064239	0	102784	45.74373
SkinF0v25	21240	6014088	0	101934	45.33966
SkinF0v30	21015	5968919	0	101168	45.00441
SkinF5v0	20616	6120554	0	103738	46.07136
SkinF10v0	20561	6116439	0	103669	46.03559
SkinF15v0	20650	6113381	0	103617	46.01593
SkinF20v0	20630	6095358	0	103311	45.77102
SkinF25v0	20814	6104906	0	103473	45.9439
SkinF30v0	20434	6041178	0	102393	45.49475
SkinF5v5	21052	6116635	0	103672	46.07068
SkinF10v10	21465	6047687	0	102503	45.59424
SkinF15v15	21444	6092085	0	103256	45.94932
Variations where reads below trim level have been removed.					
SkinR0v230	8848	1619137	2	26986	69.72117
SkinR260v230	5166	1259130	2	20986	51.61
SkinR270v230	6652	1398410	2	23307	58.56633
SkinR280v200	3304	82673	1	1378	5.321
SkinR280v220	6571	1373224	2	20464	50.138
SkinR280v240	9615	1918710	2	31979	85.46583
SkinR280v250	9608	1904411	2	31740	84.6135
SkinF10R280v230	6572	1380390	2	23007	57.48967

4.3.5.3. Oral

The variations of the oral samples that focused on trimming the front of the sequences yielded around 4000 unique sequences with a total frequency of over 9 million (Table 4.11), these showed an average merged percentage between 80-82% and the minimum sequence frequency was around 118 thousand. Trimming of the ends of the sequences yielded between 6000-9000 unique sequences with a total frequency between 2 million and 7 million, however the percentage merged ranged from 20-70% and the minimum sequence frequencies ranged from 12 thousand to 80 thousand.

Table 4.11 Summary of select merged sequence information for the oral samples processed through variations in DADA2. *A colour scale is applied to each column going from green as the largest number and red the lowest.*

Run ID	Number of unique sequences	Total sequence frequency	Minimum sequence frequency	Mean sequence frequency	Mean percentage merged
Variations with reads of all length.					
Oral0v0	4326	9302456	117734	155041	80.7625
OralF0v5	4351	9330966	117445	155516	81.11017
OralF0v10	4355	9339978	117829	155666	81.19083
OralF0v15	4358	9345935	118073	155766	81.23983
OralF0v20	4335	9345295	118114	155755	81.2245
OralF0v25	4312	9353329	118354	155889	81.24883
OralF5v0	4316	9265242	117959	154421	80.73617
OralF10v0	4294	9259070	117921	154318	80.749
OralF15v0	4226	9264525	118092	154409	80.74133
OralF20v0	3937	9360524	118735	156009	80.6285
OralF25v0	3977	9385209	118789	156420	81.03183
OralF5v5	4347	9297263	117672	154954	81.12
OralF10v10	4341	9300272	118022	155005	81.21083
OralF15v15	4274	9315033	118377	155251	81.261
OralF20v20	3907	9409944	119145	156832	81.15667
OralF25v25	3865	9417712	119420	156962	81.16067
Variations where reads below trim level have been removed.					
OralR230v190	6528	2322577	12864	38710	21.03133
OralR240v200	9323	7205565	80377	120093	68.20417
OralR250v200	9679	7105704	79073	118428	67.436
OralR250v210	9575	6554393	64487	109240	61.93167
OralR250v220	9004	6119426	55359	101990	57.655
OralR250v230	7934	5380005	42728	89667	50.42933

4.3.6. Filtering Overview

Once the sequences had been successfully merged and classified it was necessary to filter them to remove known contaminants and transient organisms from the final datasets, the filtering effects the statistics of the datasets in various ways. As seen in table 4.12 the initial taxonomy-based filtering using the taxa filter and decontam filter steps reduce the total number of unique sequences by around 4000 in the arm, 10000 in the face, and 1000 in the oral. The frequency filtering step removed around 2000-3000 unique sequences in the skin samples but 6000 in the oral samples. The change in total sequence frequency however was fairly consistent and less drastic over all sample types, losing a few hundred thousand over the course of the filtering pipeline. When the sequence length was compared across the filtering stages an initial increase in mean length occurred after the taxonomy-based steps, but this did not affect the range of lengths covered within the sequences. Only during the last filtering step did a reduction in the range of lengths occur, this was seen as a reduction in over 100 bp.

During this filtering process sample OA01 was removed from the arm dataset due to its low number of reads, the removal occurred during the frequency filtering stage.

Table 4.12 Summary of select merged sequence information for all samples during the stages of the filtering process.

Arm		Number of unique Sequences	Total Sequence Frequency	Mean Sequence length	Sequence length range
	Un-Filtered	13910	6136616	445.1	230
	Taxa Filter	9476	5961150	479.12	230
	Decontam Filter	9453	5934754	479.11	230
	Frequency Filter	7944	5927303	479.41	230
	Length Filter	7895	5916250	479.65	109
Face	Un-Filtered	22179	6825593	437.85	230
	Taxa Filter	12950	6545988	481.0	230
	Decontam Filter	12929	6530159	480.98	230
	Frequency Filter	9513	6513598	481.41	230
	Length Filter	9464	6498760	481.59	88
Oral	Un-Filtered	11129	8857778	404.84	180
	Taxa Filter	10280	8794310	410.34	180
	Decontam Filter	10232	8663380	410.33	180
	Frequency Filter	4094	8643930	410.43	178
	Length Filter	4089	8643768	410.55	37

4.4. Discussion

This chapter of the thesis has focused on the establishment and optimisation of the pipeline used to generate the final datasets from the raw sequencing output, this is important to understand as decisions made in creating these final datasets will impact the overall results and interpretations of them. The pipeline needed to be established for this work due to the novel nature of it within Aston University, specifically relating to the sample types and variable region of the 16S RNA gene sequenced. Each sample type from the arm, face, and oral cavity was optimised in isolation to ensure any site-specific contaminations that may have affected the sequencing or other common variations did not bias the optimisation of the other samples.

Screening organism growth from the samples yielded no significant differences between the age groups at the minimal level of this study. The original intention of this organism recovery was to generate some isolates for culture-based investigations. However, this was not continued past this preliminary isolation and identification stage due to difficulty in securing access to a MALDI-TOF system or other cost and time effective means of speciation, along with limited time due to prioritisation of this 16s sequence processing and analysis. Of interest within these results the DNA hydrolysis tests suggest that the majority of presumptively identified and recovered *Staphylococcus* species from the arm samples are *Staphylococcus aureus*, whereas there was limited *S. aureus* in the oral samples. However, in the 16S data no *S. aureus* was identified (Chapter 5 and 6). This

may be due to limitations of the taxonomic classifier used, or the 16S rejoin targeted for sequencing not being suitable for *Staphylococcus* species or differences in ease of live recovery from the swab and sampling solution compared to the DNA extraction, either way DNA hydrolysis is not guaranteed to only work with *S. aureus* and may be positive for other *Staphylococcus spp.* (Kateete *et al.*, 2010). Regardless this knowledge of *Staphylococcus* and *Streptococcus* species was invaluable in this pipeline optimisation, as it acted as a positive control for the classification. Additionally of the four samples removed from the sequence arm dataset due to low recovery only the YA01 sample set had no organisms successfully recovered, suggesting that the low yields may have been due to errors in the DNA extraction or sequencing processes. Of similar interest during this stage 3 species of *Candida* were recovered from the participants, while these cannot possibly be identified through or contribute to the 16S analysis, they can be interpreted in the context of current understanding of *Candida* in ageing. The results here are at odds with published literature which shows a significant increase in candida habitation with age (Zaremba *et al.*, 2006; Benito-Cruz *et al.*, 2016), this may be due to the age ranges in the OA group being too young to show a clear difference or differences in the sampling and recovery method used here. Regardless it is possible that the C*andida* recovered here may be contributing to the function of the oral microbiome and impacting ageing, therefore their presence in the oral microbiome in ageing should be further studied. The portions of the samples that were taken during sample processing are still stored along with the *Candida* isolates at -80^{o}C so it would still be possible to isolate organisms from these samples to carry out culture-based analysis or whole genome sequence analysis. While these results do not directly add to the further analyses conducted, they do demonstrate a microbiome with many active component species.

Quantification of the DNA extracts revealed initial disparities between all the sample types. The skin samples had the lower nucleotide concentration and more impurities when compared to the oral samples, however all three contained impurities that absorb at the 230nm wavelength (such as organic compounds). Various studies (Santiago-Rodriguez *et al.*, 2023; Serghiou, Webber and Hall, 2023) report that skin microbiome presents challenges with low biomass yields through different sampling techniques, these then suffer from an increased risk of contamination from both the sampling and extraction processes through both the environment and reagent sources (Salter *et al.*, 2014; Eisenhofer *et al.*, 2019; Weyrich *et al.*, 2019; Manus *et al.*, 2022), in larger biomass samples such as the oral these contaminant sources are mitigated by the higher sample nucleotide yield. Overall while some of these concentration and impurity issues may have affected the quality and quantity of sequence successfully recovered none of these fully prevented sequencing from being completed. The raw sequences recovered were broadly similar in terms of average and total counts in all samples with the face samples being distinct in having the largest total count and maximum count samples (Table 4.4). The issues regarding DNA extract quality were seen to have impacted the sequencing process as the skin samples contain low read samples along with an overall reduction and inconsistency in quality compared to the oral data, however some of this difference may be due to the different 16S regions sequenced for the skin and oral samples

or known variation caused by extraction kit and participant derived variation (Wagner Mackenzie, Waite and Taylor, 2015), additionally 16S amplicon sequencing can be demonstrated to gain low yields compared to other sequencing methods when handling low biomass skin samples (Castelino *et al.*, 2017; Markey *et al.*, 2025). The variation between the sample types seen supports the decision to treat each in isolation for the pipeline optimisation, although the shared similarity seen by most of the samples suggests that for the skin samples the same or similar parameters will be acceptable. Some alterations to the sample gathering or extraction process may help improve the quality issues in future studies.

The removal of the primer sequences along with other contaminants was necessary to aid the following merging and then classification stages, some optimisation was required at this stage due to the quality variations seen in the raw sequence output. The initial 3 variations trialled CA1-CA3 focused on the minimum primer overlap required for removal of the primer length of the sequence, as there was little variation in any of the sequence count metrics resulting from these across all sample types it was therefore sensible to use the 15 bp overlap parameter as this provided the strongest assurance that genuine primers were being removed as a slight increase in returned sequences was seen in the skin samples. Meanwhile the variations CA3-CA5 focused on the Phred quality score, here all sample types showed an increase in overall quality and a specific increase in the reverse read quality, regarding the highest Phred score cutoff at 30. However, this quality score was detrimental to the number of sequences recovered, with the highest recovery seen when the Phred score was set to 10. Therefore, it was deemed appropriate to use the CA3 variation for the optimisation of the merging stage as this reduced the computational time while still being representative of the datasets, and then to use the CA5 variation for the finalised dataset creation which would prioritise maximum sequence recovery; the difference between these is a pipeline that could take 40 minutes or 3 hours to compute.

Optimisation of the merging stage required the trimming of either end of the sequence with the aim of either removing excess material that was still associated with the primer at the front of the read or low-quality sections of the sequence at the rear. Trimming the beginning of the sequences initially showed promise with the arm samples as it enabled a high recovery of the minimum sequence frequency, however percentage of successful mergers is low and there was difficulty in classifying the sequences to those known with-in the samples; specifically *Staphylococcus* species. Within the arm samples no forward trim trialled resulted in successful mergers of the sequences in the sample with lowest of the minimum input meaning a loss of a participant from the dataset, these also suffered from the same problems with classification as the arm samples. The oral samples did not have these issues with trimming the beginning of the sequence. Conversely trimming at the rear of the sequences reduced the minimum sequence frequency for the arm samples while enabling the recovery of them within the face samples, these also saw an increase in successful percentage merged and a reduction in unique sequences, whereas the oral samples saw an inverse. Taxonomic classification within the skin samples however was more successful in these variations. These challenges in trimming for amplicon based

16S studies are known, although trimming does pose an risk in altering the abundances or classification it is key that the decision for the choice made is clear and communicated (Haider *et al.*, 2024). Overall trimming the rear of the sequences was more suitable as this allowed the optimisation of the rejoin overlapping directly while trimming the front did demonstrate some improvement the lack of variation seen suggests that this did not significantly contribute to the merging process. Therefore, the optimal parameters were selected from these rear trims, the suitability of this was enforced by the improvement in taxonomic classification that it was associated with.

Finalisation of the datasets before analysis required the filtering of known contaminants, low read sequences, and sequences of abnormal length; this reduced the number of sequences within the datasets. Generally, this reduction in sequence was seen at the unique sequence level rather than the total sequence frequency, which suggests that the shared core of sequences representing the microbiome of the samples remained untouched which is supported by established literature (Cao *et al.*, 2021; Nikodemova *et al.*, 2023). However, some of these removed sequences may still have been part of the microbiome from the participant sample but due to sequencing error were identified as different from other sequences or generating shorter lengths, similarly the sampling or extraction process may have contributed to this. Likewise, some of the unclassified sequences may have been normal taxa within the microbiome but due to limitations of the classifier miss-labelled. The source of some contaminants will have been the sampling reagents, or extraction kits (Paniagua Voirol *et al.*, 2021), along with environmental sources during sampling and handling. This could complicate their removal as they could appear consistently across all samples, however they should be present in low numbers and identified in the negative control enabling for their removal. The source of the abnormal length reads may be errors in the sequencing process or non 16S sequences that happen to share the primers used, therefore their removal is appropriate to prevent any ambiguity.

Various limitations affect the various optimisations of the dataset assembly pipeline explored here. Additionally, during the culture-based screening aside from the failure to carry the work forward there was also a lack of any possible conclusions about the growth of fungi which is why SDA agar was used, this is due to the lack of antibiotics in the medium that means it is highly likely that any growth on these was bacterial. In the future this should have been done with antibiotics or the study should have used whole genome sequencing as either would have enabled an understanding of the fungi present on in the microbiomes and any role it may play in ageing. An initial limitation already explored is the low quality and content in some of the skin samples, this has required the most focus to ensure maximal recovery. This may possibly be improved by altering the sampling or extraction process, in fact during the process of this study Qiagen have released an extraction kit explicitly for the recovery of microbiome swab samples which may yield better results in future studies along with the publication of studies (Verbanic *et al.*, 2019; Boulesnane *et al.*, 2020; Manus *et al.*, 2022; Shaffer *et al.*, 2022) comparing sample collection, extraction, and sequencing that would enable a more robust sampling and

sequencing pipeline to be developed. All the sample sets lacked a positive sequencing control consisting of known organisms in known proportions to be sequenced alongside, this would have aided in identifying optimal merging parameters, taxonomic classification, and filtering. Having known organisms in known proportions would have provided the merging and classification with a method comparing the different variations in a more easily accessible and less bias way than by comparing numbers (Bokulich *et al.*, 2016). Along with this, contaminants would still occur within a positive control and therefore this would have aided in the filtering process, by enabling the easy identification of some of these contaminants. Currently mock positive controls for microbiomes exist for faecal and oral studies but not the skin, this does not prevent either from being used in a future study but there may be an additional effect on the recovered sequences by not using the appropriate primers for the positive control. Additionally, no kit only negative control was used, preventing the establishment of sequences associated with the kit reagents and knowledge of contaminants from the other sampling materials (Glassing *et al.*, 2016); again, this would have aided in the filtering process. Further to this while two separate skin negative controls were prepared only one was sequenced, this prevented a fair comparison of the materials used in sample acquisition. Another method that could have also aided in the pipeline is proportionally spiking some of the samples with a selection of known extremophilic organisms (such as radiophones, or halophiles), as in a sufficiently sequenced sample they should be identifiable after processing through a robust pipeline, if a suitably low concentration has been spiked then they will be removed by the normal filtering processes (Tourlousse *et al.*, 2017; Tourlousse, Ohashi and Sekiguchi, 2018); again microbiome spiking kits are available for purchase. Finally a limitation here that will also carry over into the analysis of the datasets is that only one sample was acquired per participant ideally at least three should have been taken over a period of time, this prevents the use of a larger sequence pool being used to average out fluctuations in the core microbiome members and would have aided in the differentiation of transient organisms from low abundance microbiome members as they would have likely only been present in one or two time points. Additionally some research suggests that pooling the samples from a group of participants for sequencing is valid and successfully aids in the recovery of sequence information (Ray *et al.*, 2019), although such an action would limit the ability to utilise participant metadata to further any analysis.

4.5. Conclusion

In conclusion this chapter has demonstrated that despite some challenges it was possible to create an optimised and justified pipeline for the processing of the raw sequencing outputs for each sample into the final datasets suitable for analysis. Performing this depth of analysis and optimisation of the sequences enables a deep understanding of any underlying strengths and weaknesses in the dataset, which can be used when interpreting any ecological analysis performed on them. Finally, having presented this level of optimisation enables future 16S microbiome studies using these primer pairs to use the findings for creating their own finalised datasets.

Chapter 5. Skin analysis

5.1. Introduction

With the 16S datasets for the arm and face samples now processed and optimised, they can be characterised to determine any changes or similarities in ageing; which is the main focus of this thesis. As these are both samples from the skin then it is possible to draw direct comparisons between the arm and the face niches as they will be exposed to different microbiomes and environmental stressors.

During ageing the skin undergoes changes in terms of a thinning to the epidermis, along with a reduction in innate immune response and wound healing. Particularly, there are reductions in lipid content along with reduction of blood vessels and deterioration in neuromuscular junctions, resulting in skin sagging and feeling dry. The thinning of the skin is linked to a reduction in collagen and other fibre production, this also increases the skin frailty contributing to an increased likelihood of wound acquisition (Choi, 2019; Fernandez-Flores and Saeb-Lima, 2019; Ho and Dreesen, 2021; Ahlers *et al.*, 2022). These changes in the skin are a mixture of intrinsic and extrinsic factors interplaying, in addition to the mitochondrial dysfunction, cellular senescence, impaired immune function contributing to the previously outlined changes there are several extrinsic factors in play due in part to the exposed nature of the skin to the outside environment. These include prolonged exposure to UV radiation damaging DNA and skin fibres, exposure to other environmental pollutants such as biproducts of fossil fuel combustion, and the impact of long-term socioeconomic status factoring into skin care (Hausauer *et al.*, 2011; Puri *et al.*, 2017). Within the microbiome reductions in *Cutibacterium spp.* and *Staphylococcus aureus* have been reported, along with a reduction in metabolic pathways associated with healthy skin function (Luna, 2020; Kim *et al.*, 2022).

The characterisation of the microbiome can be carried out through various means as outlined in chapter 1 (1.5), including broad diversity analysis, mathematical modelling, and post-taxonomic investigations. Firstly, there is alpha diversity which is a method for measuring the diversity within each ecosystem, these diversity scores can then be compared to each other to discover if the various ecosystems are similar (Fig. 5.1). These metrics operate on similar core principles of species richness, evenness, or dominance/abundance; however, these are unlabelled measures, so they are limited in their capabilities. After this there is beta diversity which involves measuring the diversities across ecosystems (Galloway-Peña and Hanson, 2020), as such it does require comparisons of the taxonomic information. However, beta diversity metrics do not produce single per ecosystem diversity statistics but instead each ecosystems diversity is dependent upon the others in the study, therefore beta diversity can be used to create clusters of similar groups and if they are dissimilar to others. Beyond these broad metrics statistical analysis may be employed to identify specific species within an ecosystem or cluster of ecosystems that has a significantly increased abundance than the others being compared against, once identified

these species may then become the focus of future research (Cappellato, Baruzzo and Camillo, 2022). To analyse any relationships between the species within an ecosystem a co-occurrence network can be constructed using mathematical modelling, in these networks' species are linked when they occur in the same ecosystem and when multiple ecosystems are used commonly occurring species can become linked in chains of nodes. Comparing these networks with further statistical tests then enables an understanding of any similarities or differences between the ecosystems used to construct them, however, such models are purely based on statistical analysis and therefore limited in their inferences relevance without other methods employed to support them (B. Guo *et al.*, 2022; Goberna and Verdú, 2022; Kajihara and Hynson, 2024). Finally post taxonomic toolkits can be used to extrapolate the metabolome of the microbiome, which can be further analysed to investigate if there are any possible metabolomic differences in the microbiomes, this is important as metabolomic studies use a different sampling procedure and involve destructive analysis preventing the microbiome content from being known.

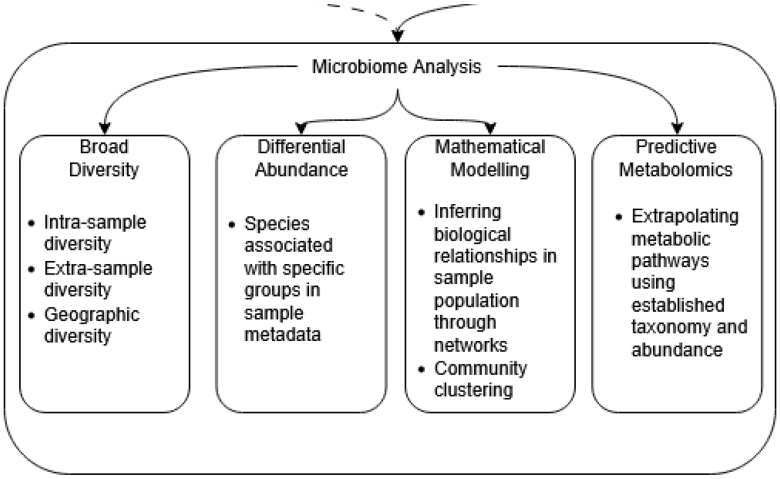


Figure 5.1 Excerpt of Figure 1.5. ***Focusing on the microbiome analysis stage of amplicon based 16S microbiome studies.***

5.1.1. Aims

- Provide an overview of the content of the skin microbiomes using the previously established datasets.
- Describe and compare the arm and face microbiomes diversities using the age groups and other potentially relevant metadata.
- Identify any notable species associated with the YA and OA skin microbiomes.
- Investigate the community structure of the skin microbiomes and any changes during ageing.
- Investigate any changes in the age groups microbiome metabolome in the skin microbiomes.

5.2. Methods

5.2.1. Importing dataset into R

The skin datasets previously created in a Qiime2 environment and the participant metadata were imported into R Studio using qiime2R (Jordan E Bisanz, 2018), and converted into a phyloseq object (McMurdie and Holmes, 2013).

5.2.2. Data normalisation

The rarefy_even_depth function was used to normalise the datasets, the arm was normalised to a depth of 12500 and the face to a depth of 25000. This resulted in the loss of two participants from the arm dataset, YA01 and YA29, along with OA20 from the face dataset. During this stage the blank samples were also removed from the datasets.

5.2.3. Statistical analysis

Broad ecological diversity metrics using alpha and beta diversity were calculated using the normalised datasets, using the metadata the datasets were investigated by the two age groups along with gender, frailty index, and other groups (Outlined in Chapter 3). Alpha diversity metrics were explored using the microbiome package (Leo Lahti and Sudarshan Shetty, 2025). The specific metrics used were, Shannon diversity, Simpson diversity, Observed features, and Pielou evenness. These were statistically compared using the Kruskal-Wallis test implemented using the vegan package (Jari Oksanen *et al.*, 2025). The beta diversity metrics explored included: Bray-Curtis dissimilarity, Jaccard index, weighted and un-weighted Unique Fraction metric (UniFrac); these were computed using Principal Coordinates Analysis (PCoA), ordination method. Statistical comparisons were carried out using a permutational analysis of variance (PERMANOVA), using the pairwiseAdonis package (Pedro Martinez Arbizu, 2017).

Differential abundance analysis was carried out using the non-normalised datasets, due to the different methods used having their own preferred normalisation requirements. Different filtering of the datasets was trialled, removing taxa that occurred less than 10, 50, 100, and 200 times in each sample. For simplicity these filtering thresholds were only screened using the DeSeq differential abundance method. For the full analysis filtering removed taxa that occurred less than 100 times in each sample, this analysis was carried out using the microbiomeMarker package (Cao *et al.*, 2022), which acts as a wrapper script for several differential abundance methods. Taxa identified by at least three methods were considered differentially abundant among the metadata groups analysed.

To construct co-occurrence networks the non-normalised datasets were collapsed to the species level and then filtered to remove taxa that did not occur in more than 30% of the samples. A convenience wrapper script NetCoMi (Stefanie Peschel, 2022) was used to construct the networks using the SpiecEasi (Zachary Kurtz

et al., 2025) protocol. Briefly within the netConstruct function, parameters, method=" mb", lambda.min.ratio=1e-2, nlambda=20, rep.num=50. Then the netAnalyze function was used, parameters clustMethod=cluster_fast_greedy, hubPar=eigenvector, normDeg=FALSE. Finally, the networks were statistically compared using the netCompare function, parameters, permTest=TRUE, nPerm=1000. Due to the limitations of the packages and time-consuming computational complexity only the age groups within each site were compared with this method.

Predictive metabolomics was conducted using PICRUSt2 within the Qiime2 environment (Douglas *et al.*, 2020). The parameters used were the epa-ng placement tool, the mp hsp-method, and a max-nsti of 2. The output generated included was then imported into R as phyloseq objects, they were then analysed in a similar manner to the taxonomy based differential abundance using the microbiomeMarker package. However, the datasets were normalised through converting to relative abundance prior to performing the beta diversity analysis.

5.3. Results

5.3.1. Top species

After finalising the datasets 7985 unique features were in the arm samples and 9464 in the face samples (Table 4.12), this translates into 738 and 780 unique taxonomies at the species level respectfully, though some are unclassified fully to the species level. The top 10 identifiable species within the samples make up most species within the datasets, with over 60% of the arm samples and 50% of the face. Overall, some of the most common in the arm samples are in descending order *Staphylococcus hominis*, *Cutibacterium acnes, Staphylococcus epidermidis, Corynebacterium argentoratense, Moraxella osloensis,* and *Roseateles saccharophilus* (Fig. 5.2). While in the face samples the order has switched making it, *Cutibacterium acnes, Staphylococcus epidermidis, Corynebacterium argentoratense, Staphylococcus hominis*, *Moraxella osloensis,* and *Pseudomonas versuta*, among the most common. There is variation in the distribution of these top organisms within the various samples, some samples only have a small abundance of these organisms whereas in others all or just a few of these organisms make up the entirety of the sample content. For example, in the arm samples YA22, YA17, OA26, and OA21, around 90% of the sample is *Staphylococcus spp.* with small content from the other species, while in YA23, YA21, and OA08 the top 10 species make up approx. 25% of the samples. This summary demonstrated that there are visibly major shifts in the top ten species present in the arm and face samples between the age groups, this required further statistical analysis to investigate the validity of this visible difference.

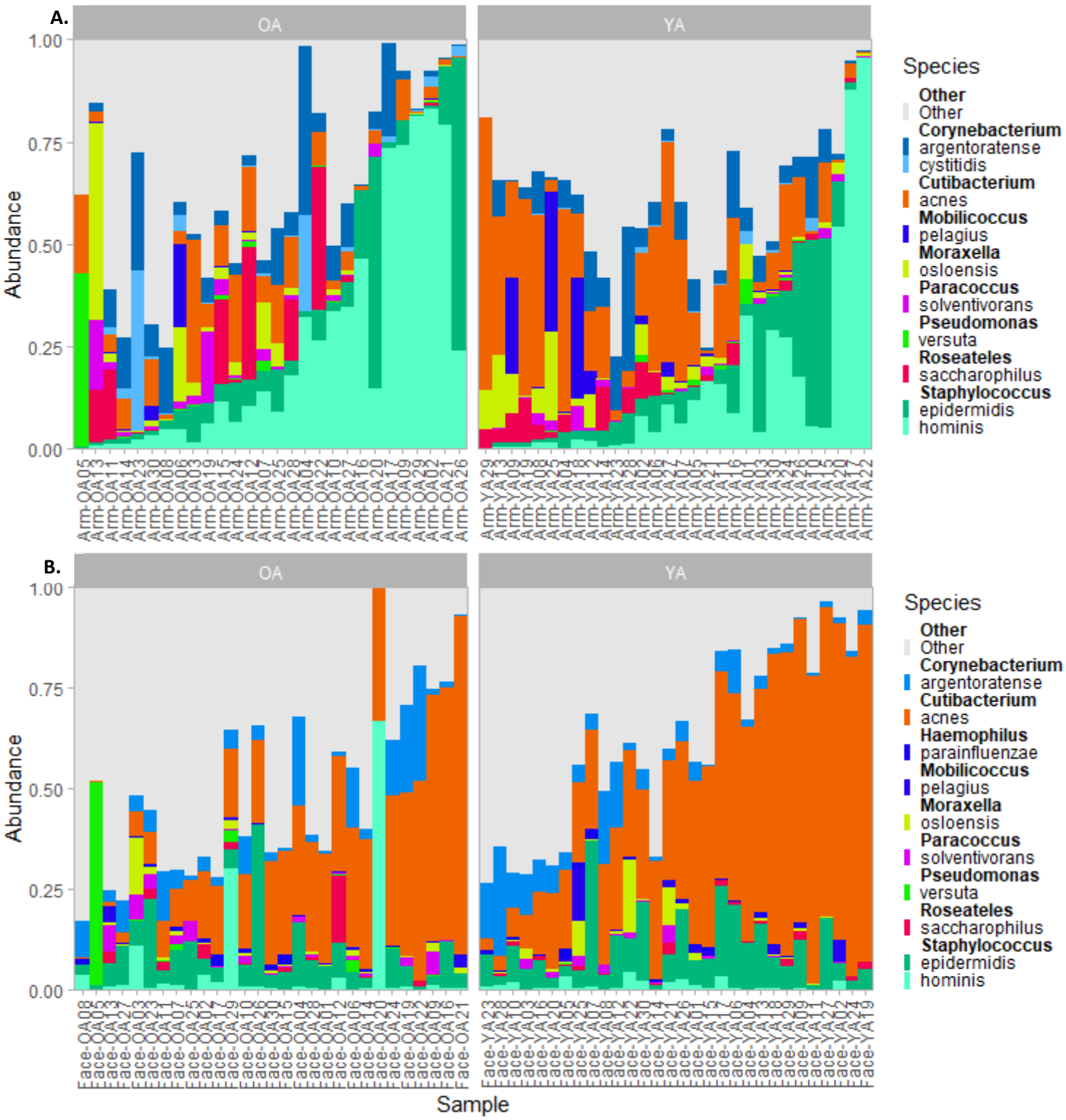


Figure 5.2 Graphs showing the percentage abundance of the species within the skin samples split by age group. *The top 10 species within the sample type highlighted, the samples are arranged in ascending order of the genus with the largest proportion in all samples. A. The arm samples. B. The face samples.*

5.3.2. Microbial diversity

After normalising the datasets through alpha rarefaction, the broad diversity metrics for alpha and beta diversity were then calculated and subjected to statistical analysis to establish any differences or similarities present.

5.3.2.1. Age Groups

When looking at the arm alpha diversity comparing the age groups, no significant differences are seen by Shannon or Simpson diversity, or in evenness, or observed features (Fig. 5.3). By these metrics the OA group cover a larger range of these indices than the YA group who are clustered tighter together. When looking at Shannon diversity and Pielou evenness the OA group mean diversity is higher than the YA group, whereas both share similar means in Simpson diversity, however, the mean of the observed features is higher in the YA group. As no significant differences were seen, but there was a presence of larger overlaps, these were investigated by plotting against individual age.

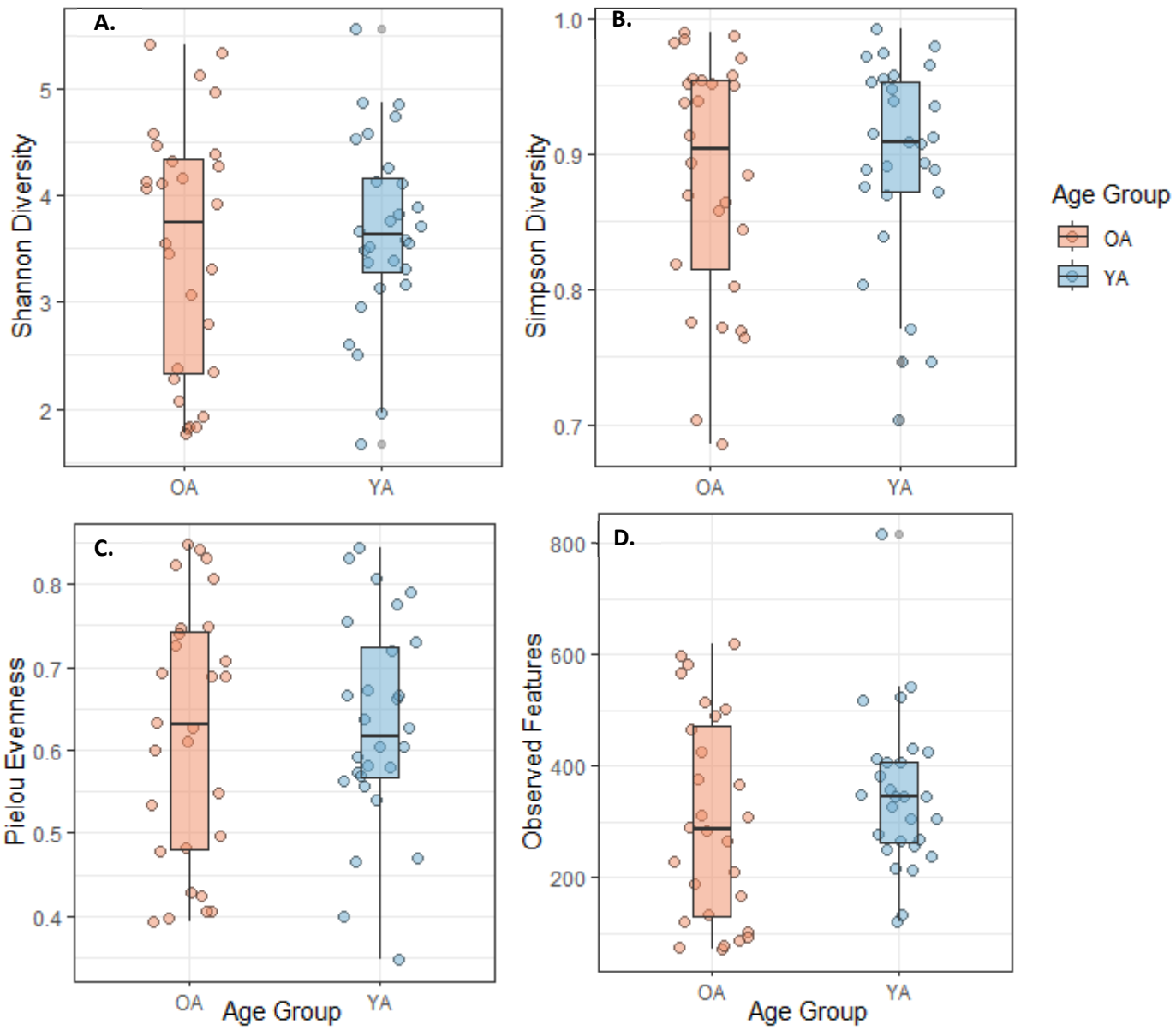


Figure 5.3 Alpha diversity metrics graphs for the arm samples, split by age groups. *No significant difference was observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

When the Shannon diversity metric is plotted against age within both age groups a general downward trend was seen in both groups (Fig. 5.4), with the regression line for the YA having an R value of -0.43 and the OA of -0.31, the YA group regression line has a significant p value of 0.023 while the OA is not significant. This is supported in the graph by the YA diversity scores still being tightly clustered when plotted in this way, comparatively while specific ages in the OA group may be clustered together these clusters are more spread across the graph. This concludes the analysis of the arm alpha diversity data; more ecological analysis was explored with beta diversity.

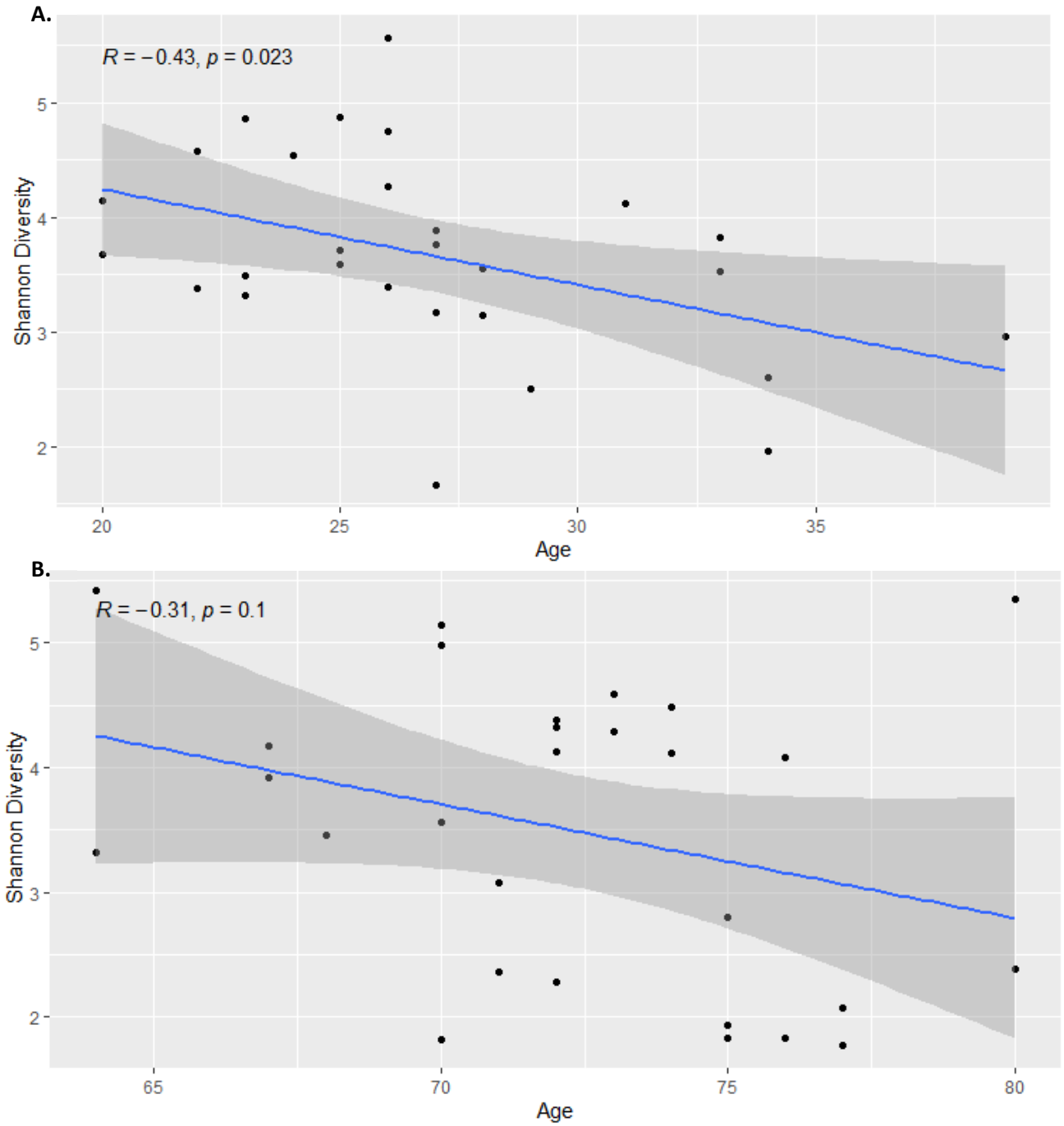


Figure 5.4 Shannon diversity graphs for the arm samples plotted by sample age. *Including regression line indicating a decrease in diversity with increased age. A. YA age group. B. OA age group.*

When beta diversity metrics were explored using Bray-Curtis dissimilarity, Jaccard index, and weighted and unweighted UniFrac methods a significant difference between the age groups was observed; with p-values of 0.002, 0.004, 0.007, and 0.001 respectively (Fig. 5.5. and Fig. 9.1). This is demonstrated as distinct clusters of the age groups being formed in the ordination plots, these clusters follow different directions in the plot's axis. Therefore, significant ecological differences were observed between the age groups arm microbiomes when cross comparing taxa and abundance through beta diversity.

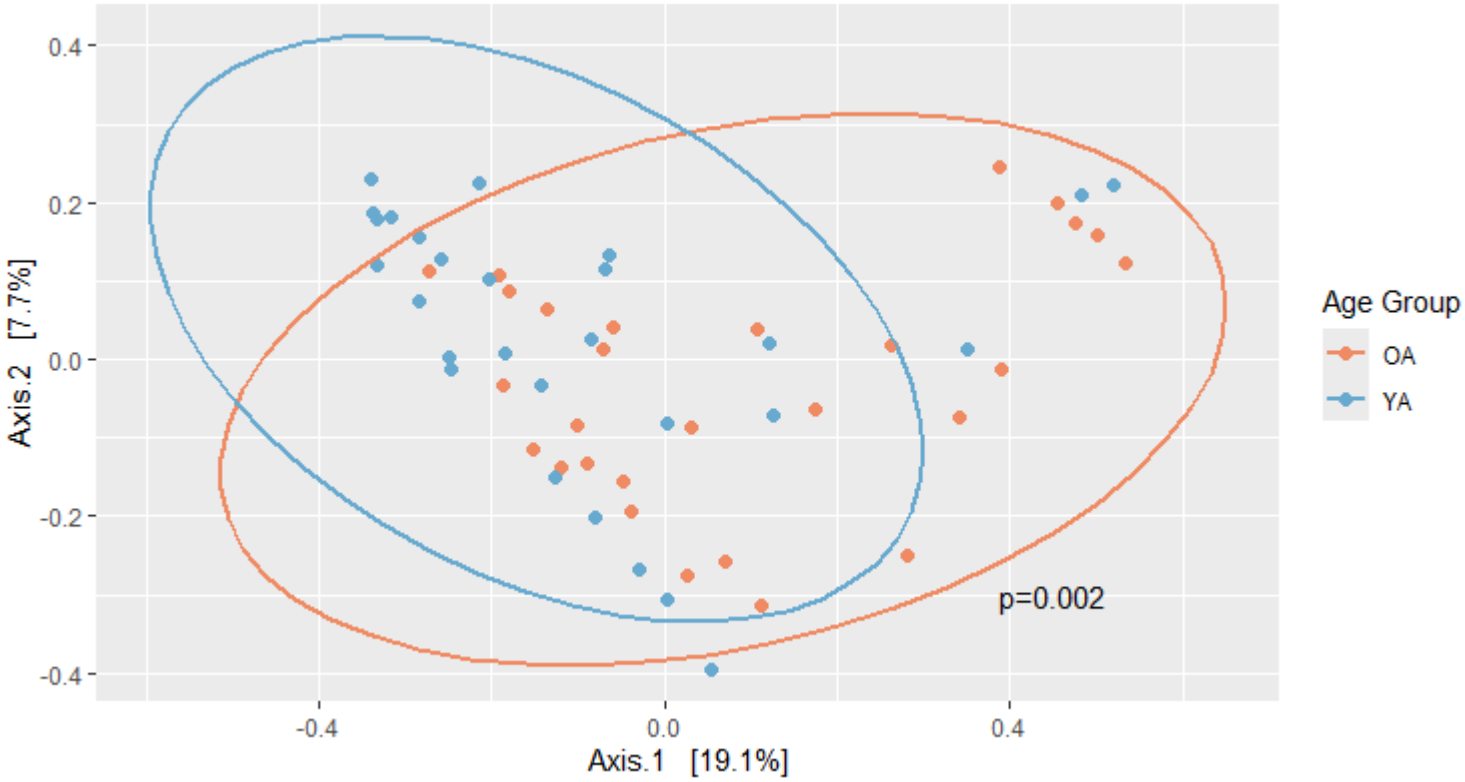


Figure 5.5 Bray Curtis dissimilarity measure of beta diversity as an ordination plot of arm samples divided by age group. ***A significant difference was seen (p=0.002).***

In comparison the face alpha diversities displayed no significant differences by Shannon or Simpson diversity, or observed features (Fig. 5.6), however a significant difference was seen in Pielou evenness with a p-value of 0.024 when subjected to a Kruskal-Walis test. The evenness in the OA group is higher than in the YA group, which is more diverse, this trend is similarly seen in the Shannon and Simpson diversities but is without significance. While the average observed features in the OA group is higher than the YA group, both groups are spread across a wide range of results. Directly comparing the face to the arm reveals a similarity in range covered by the Shannon diversity and Pielou evenness, this is not true for the Simpson diversity where the face samples cover a wider range than the arm, and the face has more observed features than the arm with an average in both groups of more than 400 compared to the arm where they are below 400. Further exploration of the face alpha diversity was possible by plotting against individual participant age.

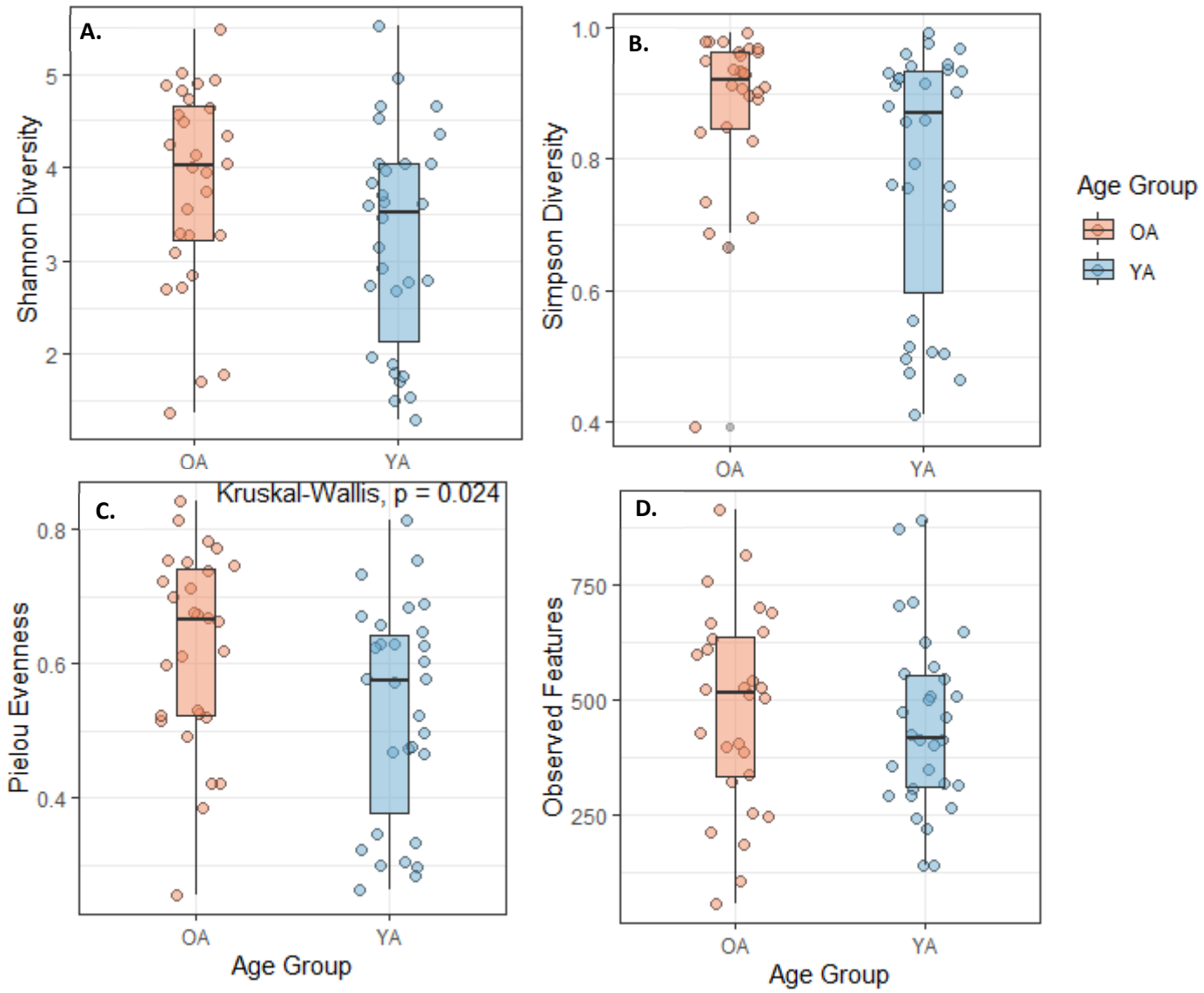


Figure 5.6 Alpha diversity metrics graphs for the face samples, split by age groups. *Where a significant difference was seen in one metric. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness, which shows a significant increase in evenness in the OA group (p=0.024). D. Observed Features.*

Again, when Shannon diversity is plotted by the ages within the age groups a downward trend was seen in the YA group with an R value of -0.25, the OA group however is nearly trendless which was reflected in an R value of -0.036, in both groups regression lines no significant differences are seen (Fig. 5.7). This is supported by both groups having a wide range of diversities across the ages. Again, this concludes the limits of alpha diversity analysis on the age groups, further exploration was completed using beta diversity.

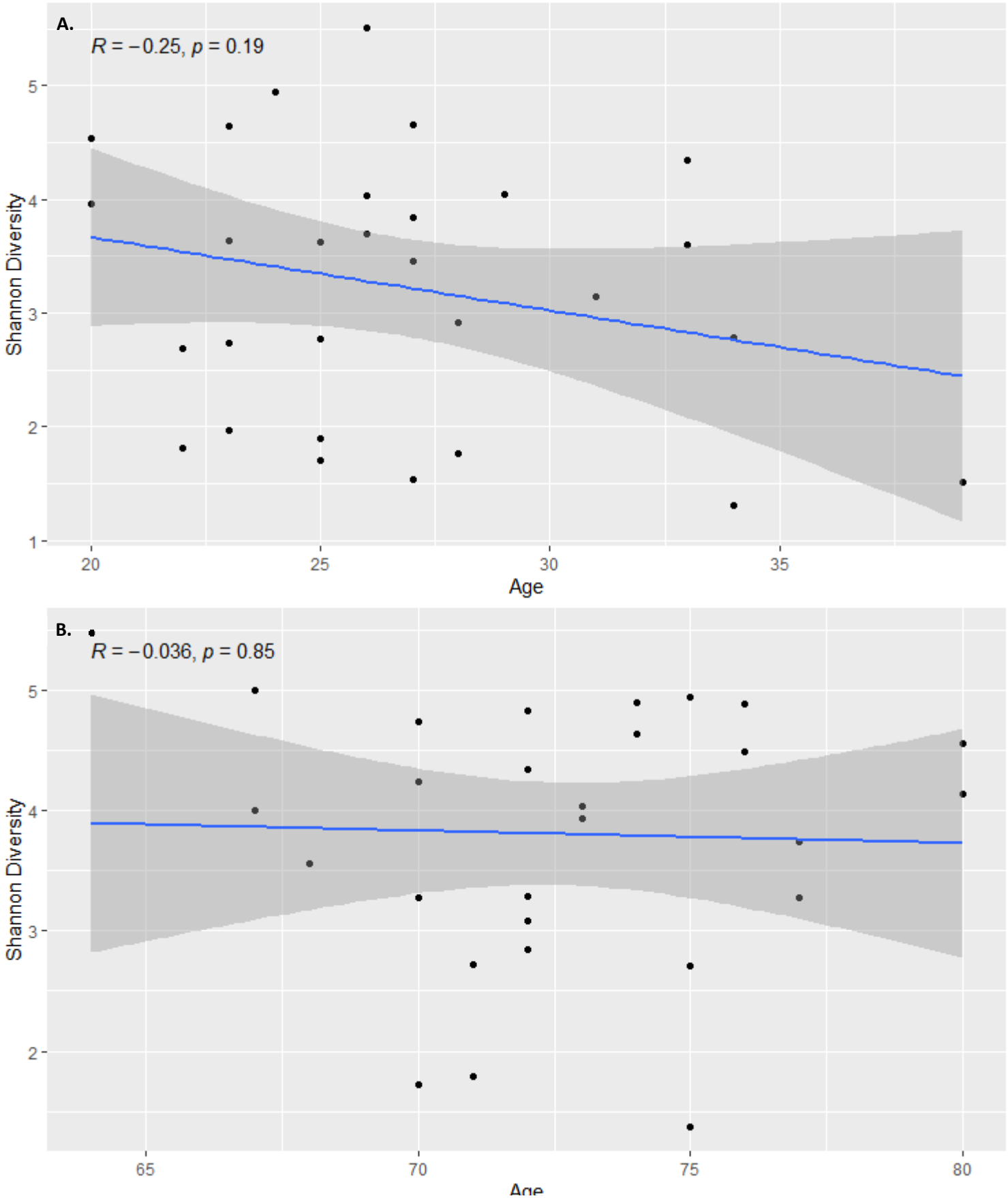


Figure 5.7 Shannon diversity graphs for the face samples plotted by sample age, including regression line. *Indicating a decrease in diversity over YA age but not in OA. A. YA age group. B. OA age group.*

Exploring the face samples beta diversity showed a significant difference in Bray Curtis dissimilarity, Jaccard index, and unweighted UniFrac metrics; with p-values of 0.009, 0.007, and 0.001 respectively (Fig.5.8 and Fig. 9.2). However, weighted UniFrac does not show any significant difference between the age groups, with a p-value of 0.219. Again, these significant differences showed distinct clusters in the ordination plots of the two different age groups, characterised by following different directions in the axis. Therefore, there were significant ecological differences were observed between the age groups face microbiomes when cross comparing taxa and abundance through beta diversity, with the OA having a significant increase in evenness in alpha diversity. This concluded the broad ecological diversity investigations using only the participant age groups.

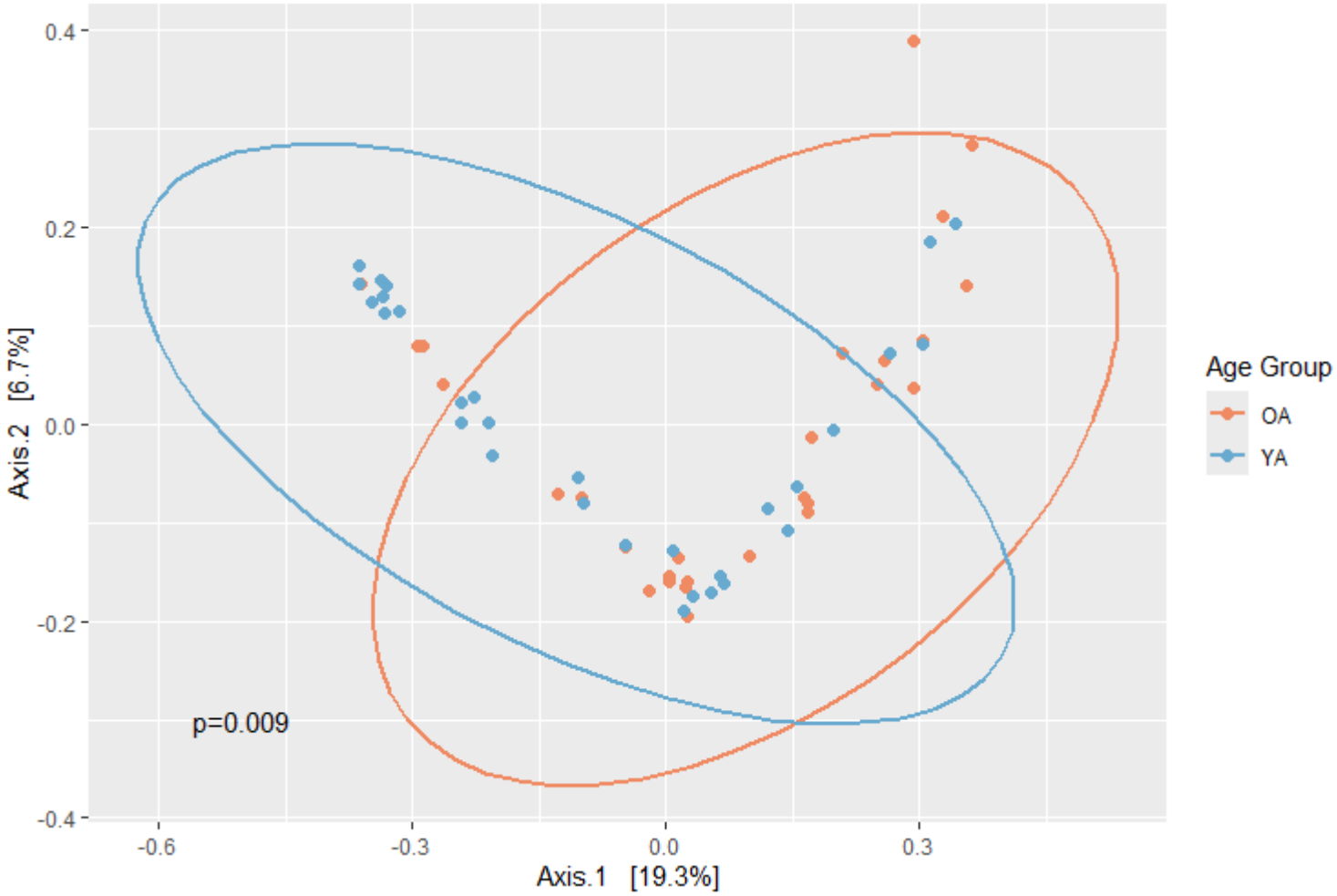


Figure 5.8 Bray Curtis dissimilarity measure of beta diversity as an ordination plot on face samples. *Split by age group, where a significant difference was observed (p=0.009)*

5.3.2.2. Gender

Another metric that can be used to compare the samples by is their gender. While alpha diversity metrics used in the arm data share the lack of a significant difference observed when grouped by age group different clustering can be observed in the genders. Across Shannon and Simpson diversity, and Pielou evenness a tight clustering can be seen in the OA males, and both YA genders when compared to the OA females; additionally, the male groups share a striking similarity (Fig.5.9). Comparatively the observed features are consistent between both age groups genders, while still sharing the overall differences seen previously. Overall, the OA females have the largest range of diversity scores across all the metrics used here, but no significant differences were observed, as these alpha diversities could not be reasonably broken down further through gender the next stage was to explore their beta diversities.

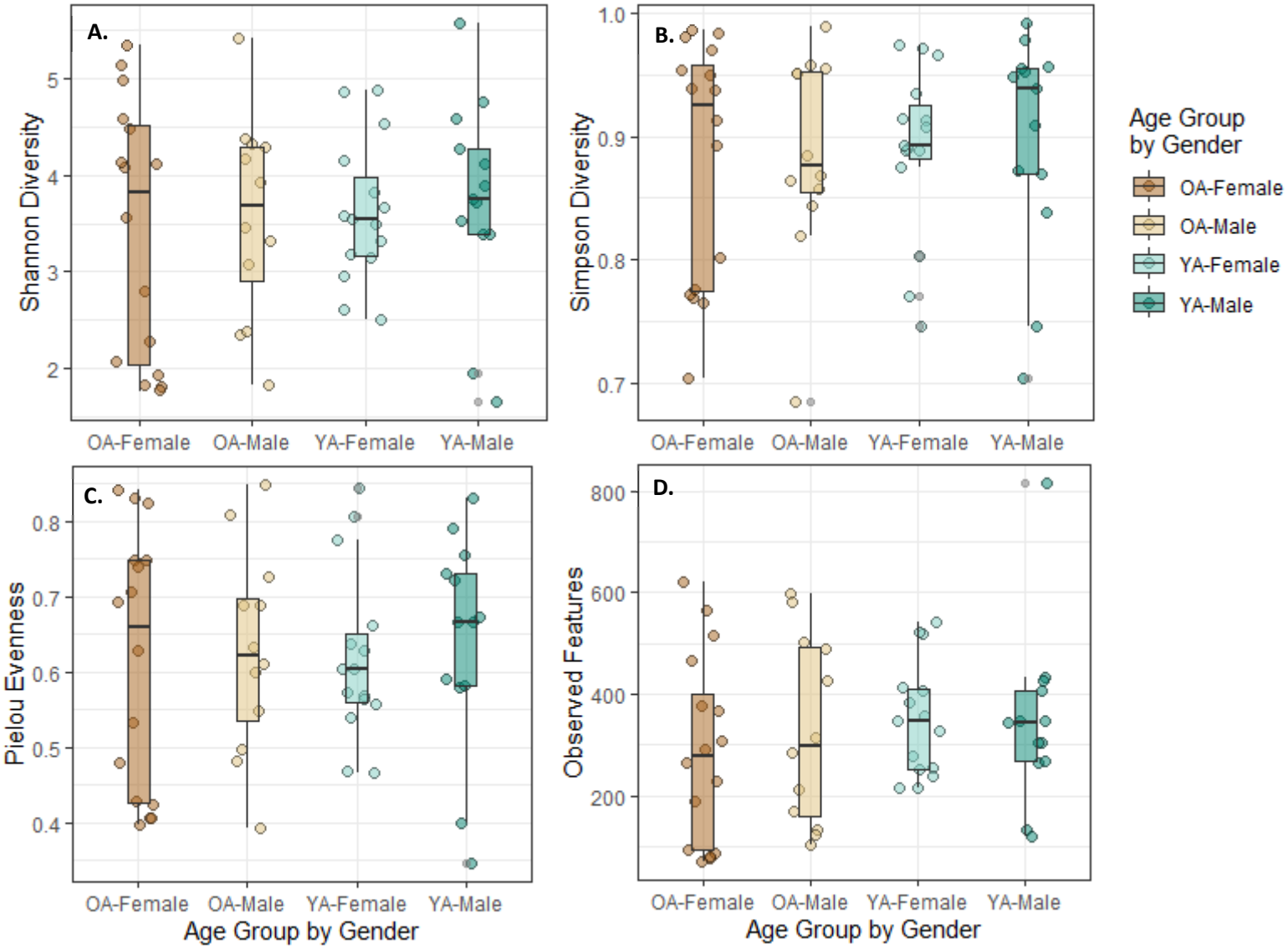


Figure 5.9 Alpha diversity metrics graphs for the arm samples, split by age group gender. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

This gendered grouping effects the beta diversity metrics, compared to the age grouping. There is now only significance between the OA and YA females across all measures, and the OA male and YA females when the Bray Curtis dissimilarity and Unweighted UniFrac measures (Fig. 5.10, Table 5.2, and Fig. 9.3). This results in distinct clusters of the significantly different interactions on the ordination plot but in clusters that have increased overlaps in the non-significantly different interactions. This change in significant differences in arm beta diversity in the gendered age groups compared to just the age groups warranted further consideration throughout the study.

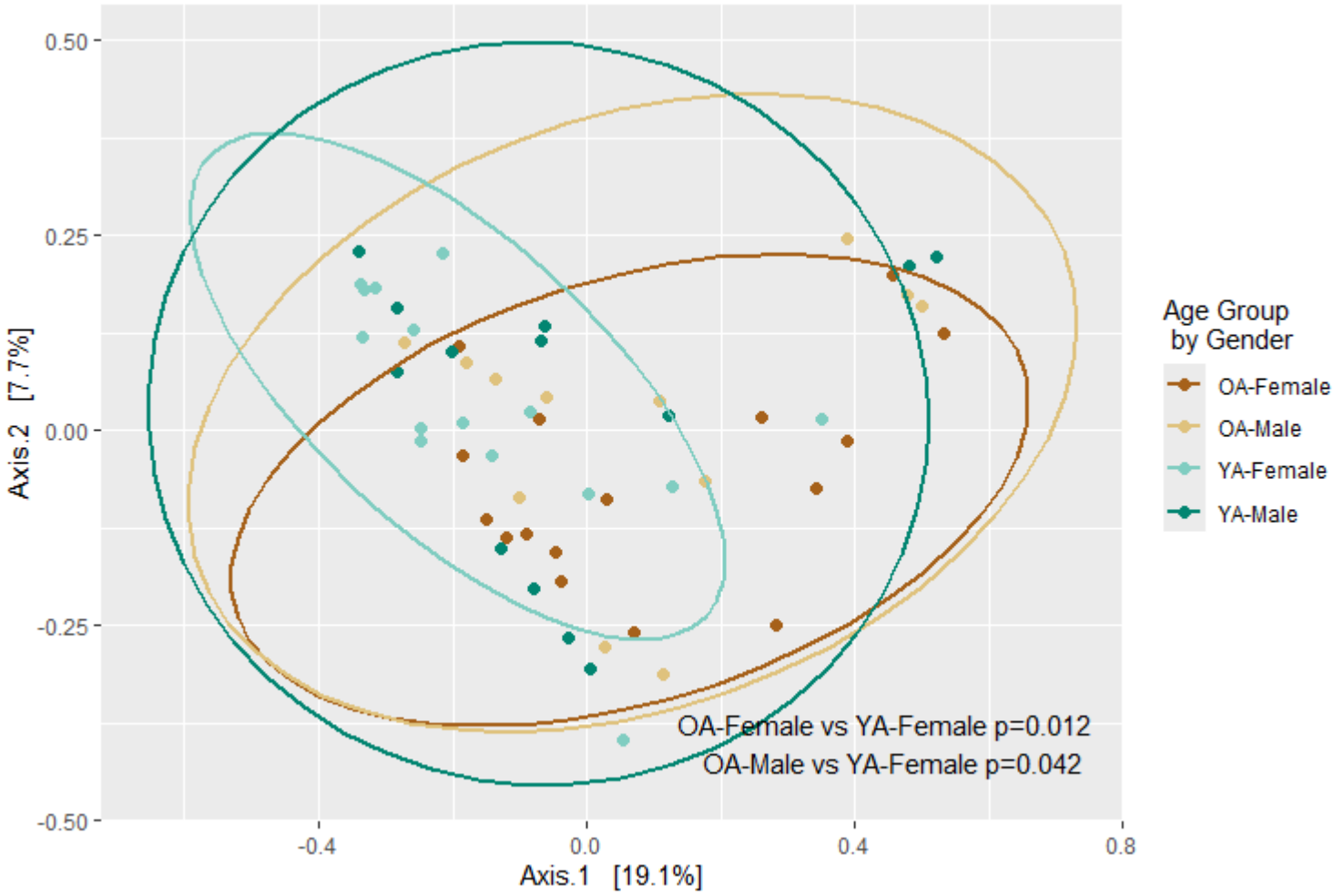


Figure 5.10 Bray Curtis dissimilarity measure of beta diversity as an ordination plot on arm samples. *Split by age group and gender, where a significant difference was observed in two interactions.*

Table 5.1 Beta diversity p-values for arm sample interactions between the age groups sub-divided into genders. *P-values below 0.05 are highlighted in green.*

	Beta diversity measure			
Interaction	Bray Curtis dissimilarity	Jaccard index	Unweighted UniFrac	Weighted UniFrac
OA-Female vs OA-Male	1.000	1.000	1.000	0.504

OA-Female vs YA-Female	0.012	0.012	0.006	0.006
OA-Female vs YA-Male	0.612	0.798	0.096	0.234
OA-Male vs YA-Female	0.042	0.078	0.018	0.588
OA-Male vs YA-Male	1.000	1.000	0.636	1.000
YA-Female vs YA-Male	1.000	1.000	1.000	1.000

The face samples alpha diversities did not show any significant differences when grouped by gender, however, similar to the arm, clustering can be observed. Unlike the arm the clustering here shows both the OA groups and the YA males as having tight clustering in Shannon and Simpson diversity, along with Pielou Evenness (Fig. 5.11). With observed features, the OA females have higher scores than the other groups, while the OA males and YA females look similar. Overall, the male groups appear similar in distribution across the metrics while the OA females tend to have a mean higher than the YA females. Again, a further break down of this grouping was not within the scope of this study, therefore beta diversity was the next option.

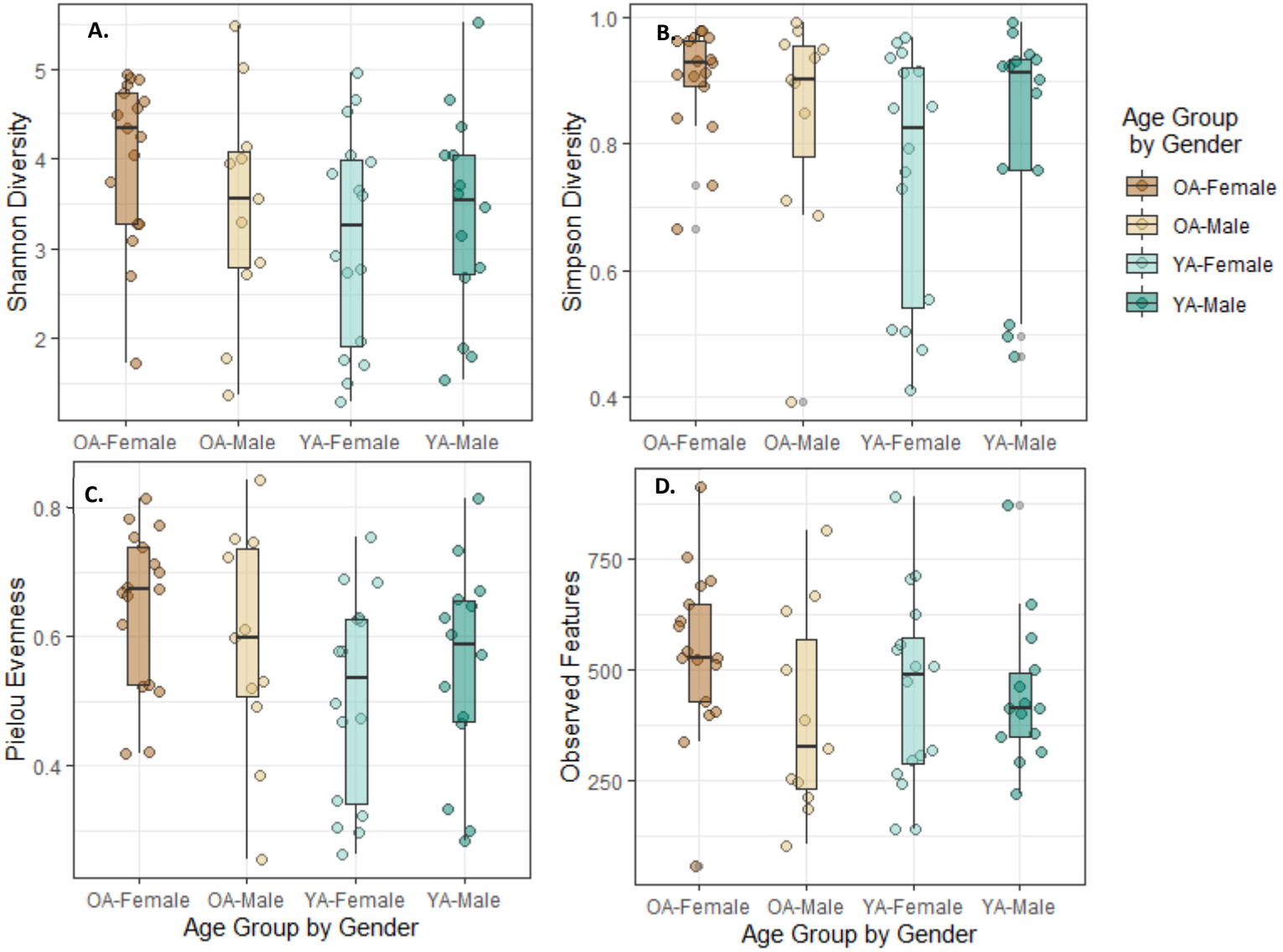


Figure 5.11 Alpha diversity metrics graphs for the face samples, split by age group gender. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

The beta diversity of the face samples when grouped by gender shows a significant difference between the OA and YA females when compared using Bray Curtis dissimilarity, Jaccard index, and unweighted UniFrac; with p-values of 0.012, 0.018, and 0.006 respectively (Fig.5.12, Table 5.3, and Fig. 9.4). Additionally, a significant difference is observed in unweighted UniFrac between the OA females and OA males with a p-value of 0.036, and the OA females and YA males with a p-value of 0.030. There is no significant difference observed in the weighted UniFrac for any interaction. Again, this results in distinct clusters of the significantly different interactions on the ordination plot but in clusters that have increased overlaps in the non-significantly different interactions. This change in these face diversities added to the consideration of gender-based differences throughout the skin analyses, however, different metadata groups still remained to be used in the analysis.

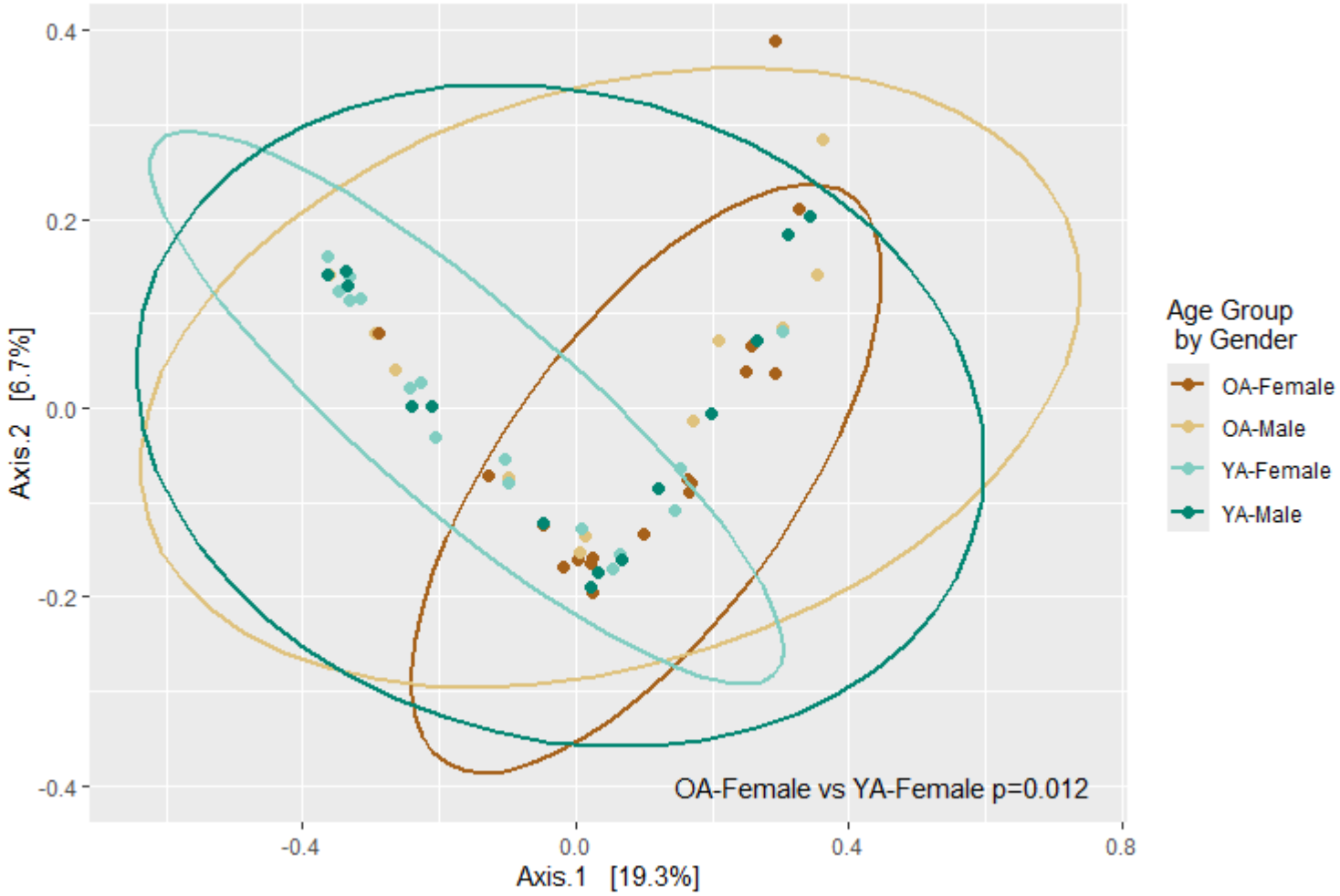


Figure 5.12 Bray Curtis dissimilarity measure of beta diversity as an ordination plot on face samples. *Split by age group and gender, where a significant difference was observed between OA-Females and YA-Females (p=0.012).*

Table 5.2 Beta diversity p-values for face sample interactions between the age groups sub-divided into genders. ***P-values below 0.05 are highlighted in green.***

	Beta diversity measure			
Interaction	Bray Curtis dissimilarity	Jaccard index	Unweighted UniFrac	Weighted UniFrac
OA-Female vs OA-Male	0.780	0.792	0.036	0.144
OA-Female vs YA-Female	0.012	0.018	0.006	0.102
OA-Female vs YA-Male	0.258	0.324	0.030	1.000
OA-Male vs YA-Female	0.576	0.492	0.174	0.102
OA-Male vs YA-Male	1.000	1.000	0.186	1.000
YA-Female vs YA-Male	1.000	1.000	1.000	1.000

5.3.2.3. Frailty Index

The FI that was used to score participants previously (Chapter 3), could also be used to group the sample data and view its diversity metrics. While no significant differences were seen in the metrics used clusters can be seen when comparing the groups, in Simpson and Shannon diversity, and Pielou evenness the samples from FI group OA-1 are tightly clustered in comparison to the other OA samples; meanwhile the YA-0 have a larger range than the YA-1 in these metrics (Fig.5.13). In observed features the YA groups are generally consistent with the range covered, while the OA-2 groups are clustered towards the lower end of the OA group with less than 100 features observed. Overall, the highest end of the FI scale, the OA-4 and YA-3 as single participant groups sit within the ranges of the other sample groups, preventing meaningful interpretation of any effect on the alpha diversity. As the FI scores cannot be easily broken down these cluster effects on the alpha diversity could not be sufficiently investigated in such a small-scale study beta diversity analysis was the next stage of analysis.

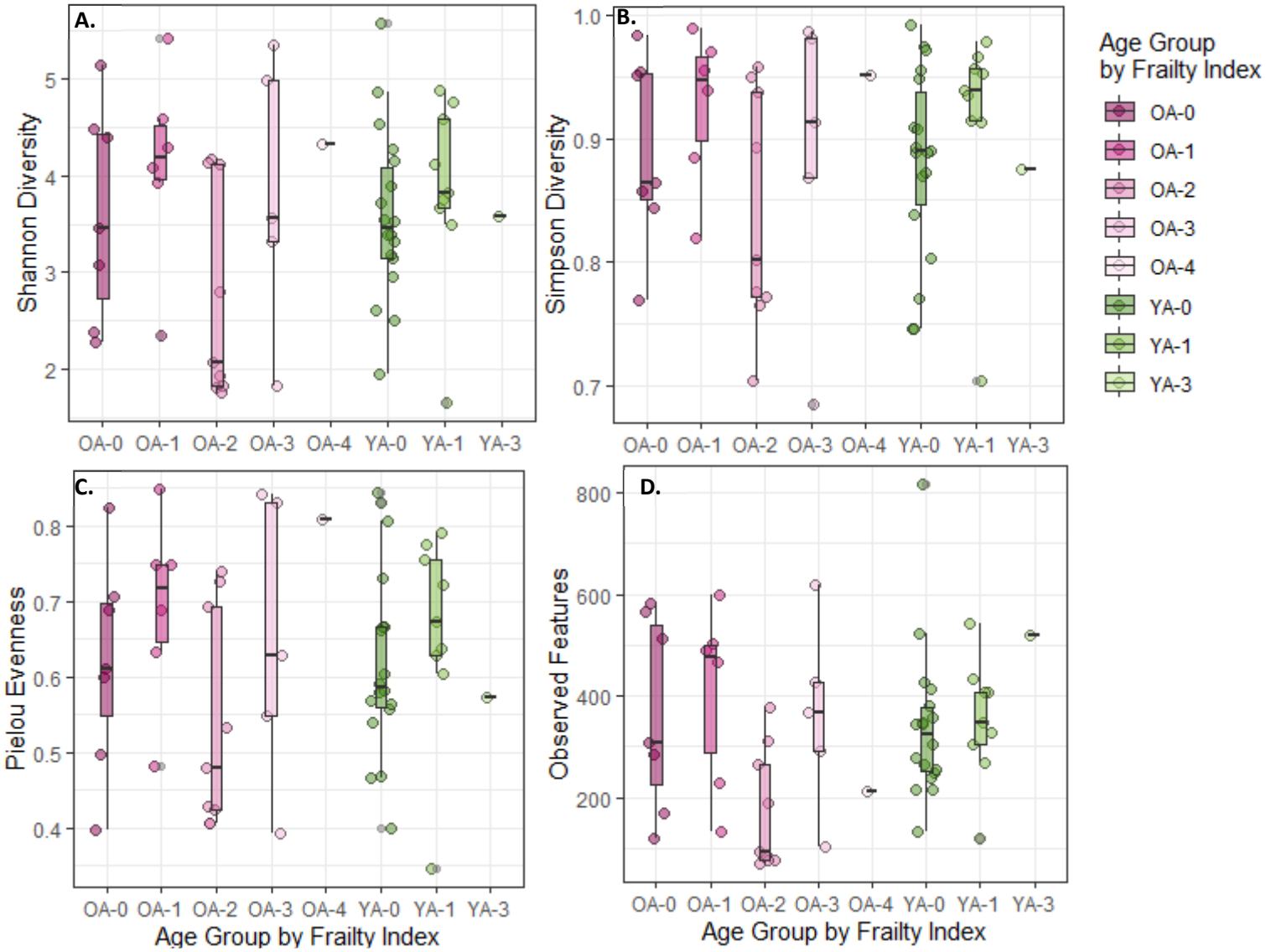


Figure 5.13 Alpha diversity metrics graphs for the arm samples, split by age group and frailty index. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

The beta diversity for the arm samples do not show any significant differences except for the interaction between OA-2 and YA-0 which has a p-value of 0.027 in the unweighted UniFrac metric, this is reflected by a general overlapping of the clusters in the ordination plots (Fig.5.14 and Fig. 9.5). Additionally performing statistical comparison between the OA-4 and YA-3 single participant groups proved impossible. Therefore, the only significant difference observed in the FI score in the arm microbiome using broad ecological terms is in the presence and absence of shared species between the OA-2 and YA-0 groups.

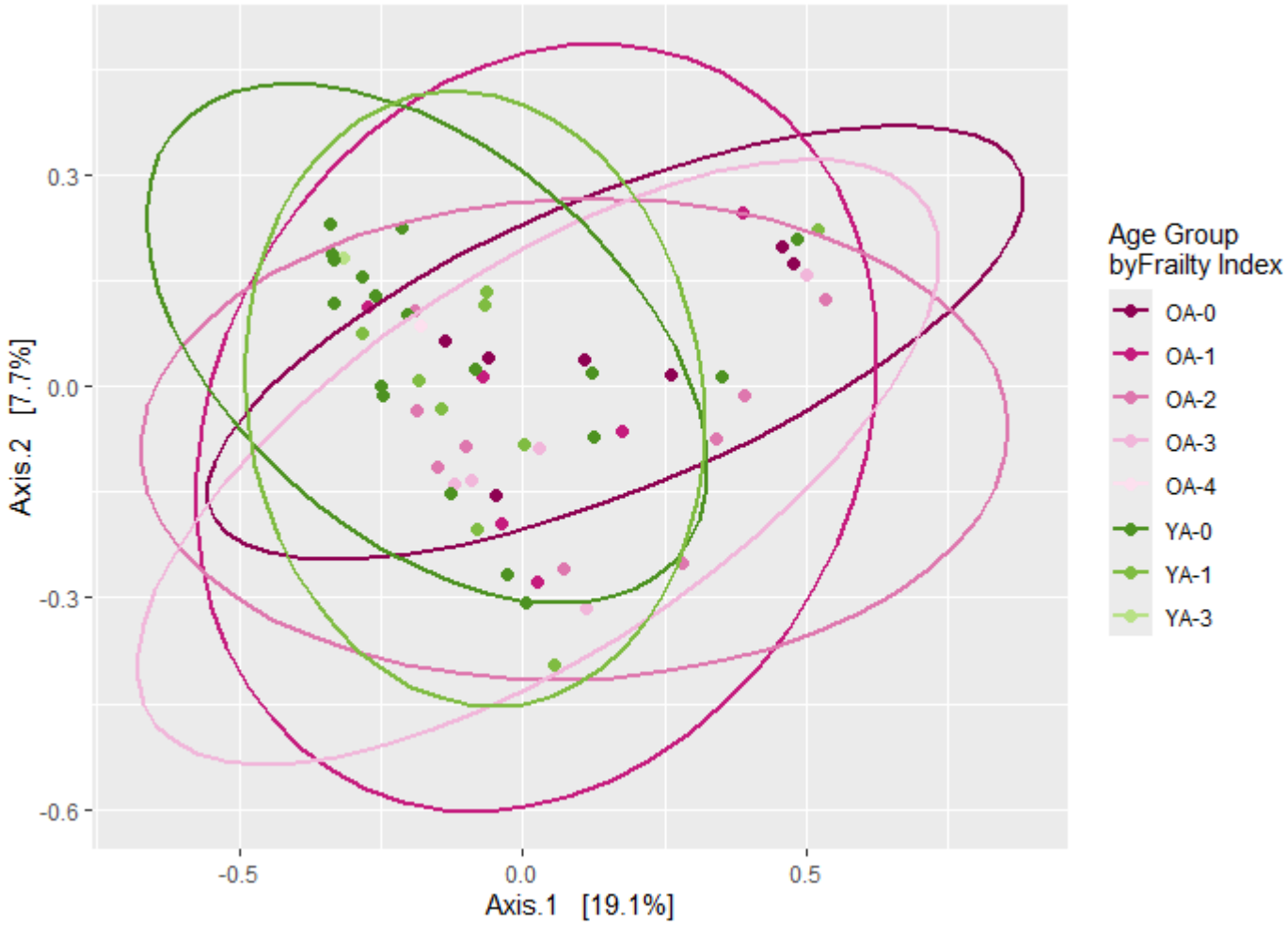


Figure 5.14 Bray Curtis dissimilarity measure of beta diversity as an ordination plot on arm samples. *Split by age group and Frailty Index score, where no significant difference was observed.*

Table 5.3 Beta diversity p-values for arm sample interactions between the age groups sub-divided into FI scores. ***P-values below 0.05 are highlighted in green.***

	Beta diversity measure			
Interaction	Bray Curtis dissimilarity	Jaccard index	Unweighted UniFrac	Weighted UniFrac
OA-0 vs OA-1	1.000	1.000	1.000	1.000
OA-0 vs OA-2	1.000	1.000	1.000	1.000
OA-0 vs OA-3	1.000	1.000	1.000	1.000
OA-0 vs OA-4	1.000	1.000	1.000	1.000
OA-0 vs YA-1	1.000	1.000	1.000	1.000
OA-0 vs YA-0	0.216	0.405	0.810	1.000
OA-0 vs YA-3	1.000	1.000	1.000	1.000
OA-1 vs OA-2	1.000	1.000	1.000	1.000
OA-1 vs OA-3	1.000	1.000	1.000	1.000
OA-1 vs OA-4	1.000	1.000	1.000	1.000
OA-1 vs YA-1	1.000	1.000	1.000	1.000
OA-1 vs YA-0	1.000	1.000	0.081	1.000
OA-1 vs YA-3	1.000	1.000	1.000	1.000
OA-2 vs OA-3	1.000	1.000	1.000	0.189
OA-2 vs OA-4	1.000	1.000	1.000	1.000
OA-2 vs YA-1	1.000	1.000	1.000	1.000
OA-2 vs YA-0	0.405	0.216	0.027	1.000
OA-2 vs YA-3	1.000	1.000	1.000	1.000
OA-3 vs OA-4	1.000	1.000	1.000	1.000
OA-3 vs YA-1	1.000	1.000	1.000	1.000
OA-3 vs YA-0	1.000	0.999	0.729	0.540
OA-3 vs YA-3	1.000	1.000	1.000	1.000
OA-4 vs YA-1	1.000	1.000	1.000	1.000
OA-4 vs YA-0	1.000	1.000	1.000	1.000
OA-4 vs YA-3	Invalid	Invalid	Invalid	Invalid
YA-1 vs YA-0	1.000	1.000	1.000	1.000
YA-1 vs YA-3	1.000	1.000	1.000	1.000
YA-0 vs YA-3	1.000	1.000	1.000	1.000

For the face samples similar to the arm samples no significant differences are seen in the alpha diversity metrics used when the FI grouping is used, additionally the groups all cluster within similar ranges to each other (Fig.5.15). Again, the FI scores cannot be easily broken down these cluster effects on the alpha diversity could not be sufficiently investigated in such a small-scale study beta diversity analysis was the next stage of analysis.

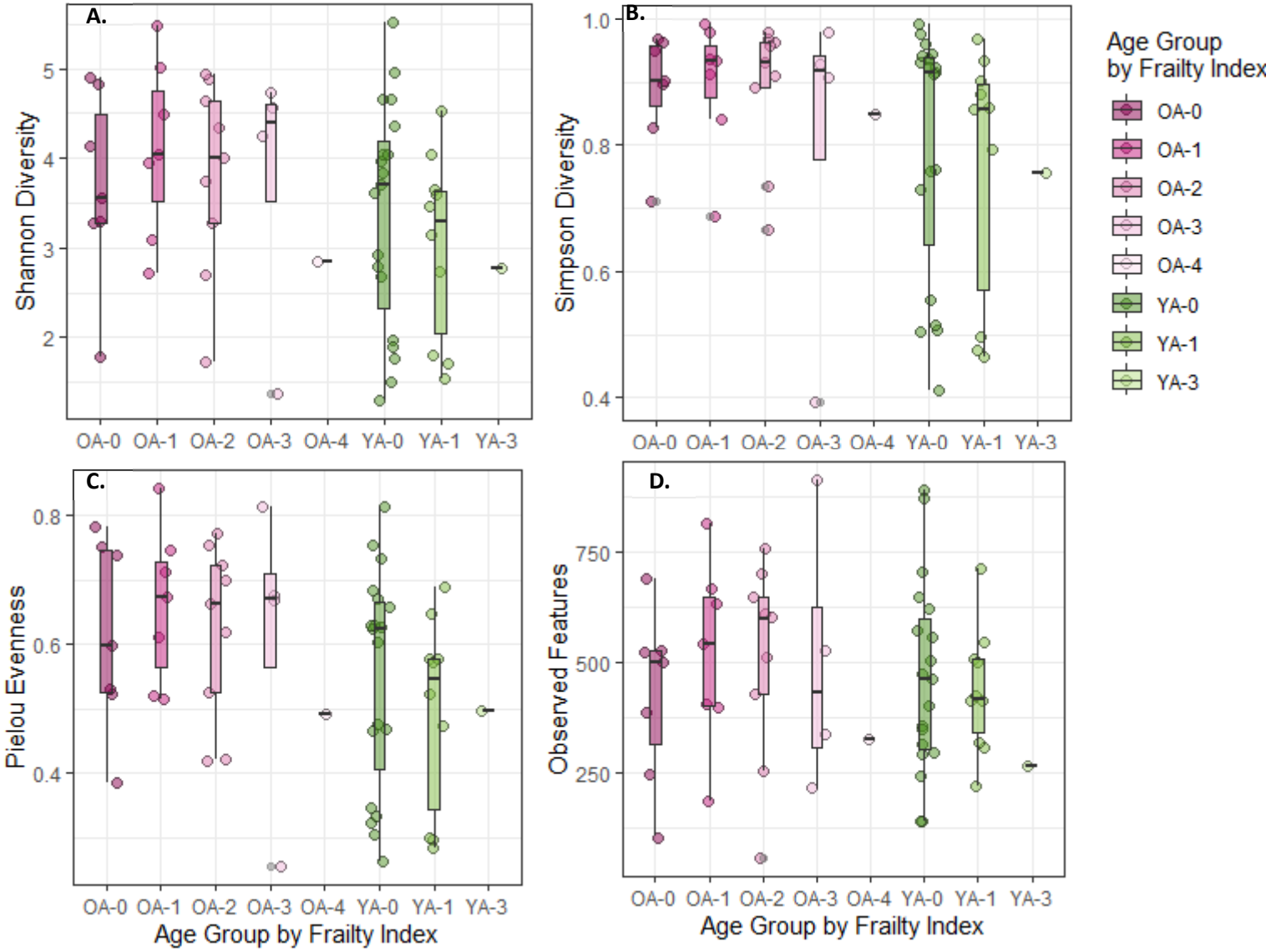


Figure 5.15 Alpha diversity metrics graphs for the face samples, split by age group and frailty index. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

In the beta diversity of the face samples show no significant differences when clustered by FI (Fig.5.16 and Supplemental Figure 1.6). This is reflected by a general overlapping of the clusters in the ordination plots, again performing statistical comparison between the OA-4 and YA-3 single participant groups proved impossible. Therefore, no significant differences were observed in the face broad ecological diversities when the FI scores were used to cluster the data, this is different from the arm which only saw one significantly different interaction.

This concluded the FI score analysis of the skin microbiome using alpha and beta diversity, additional potentially relevant subgroups remained still remained available for analysis.

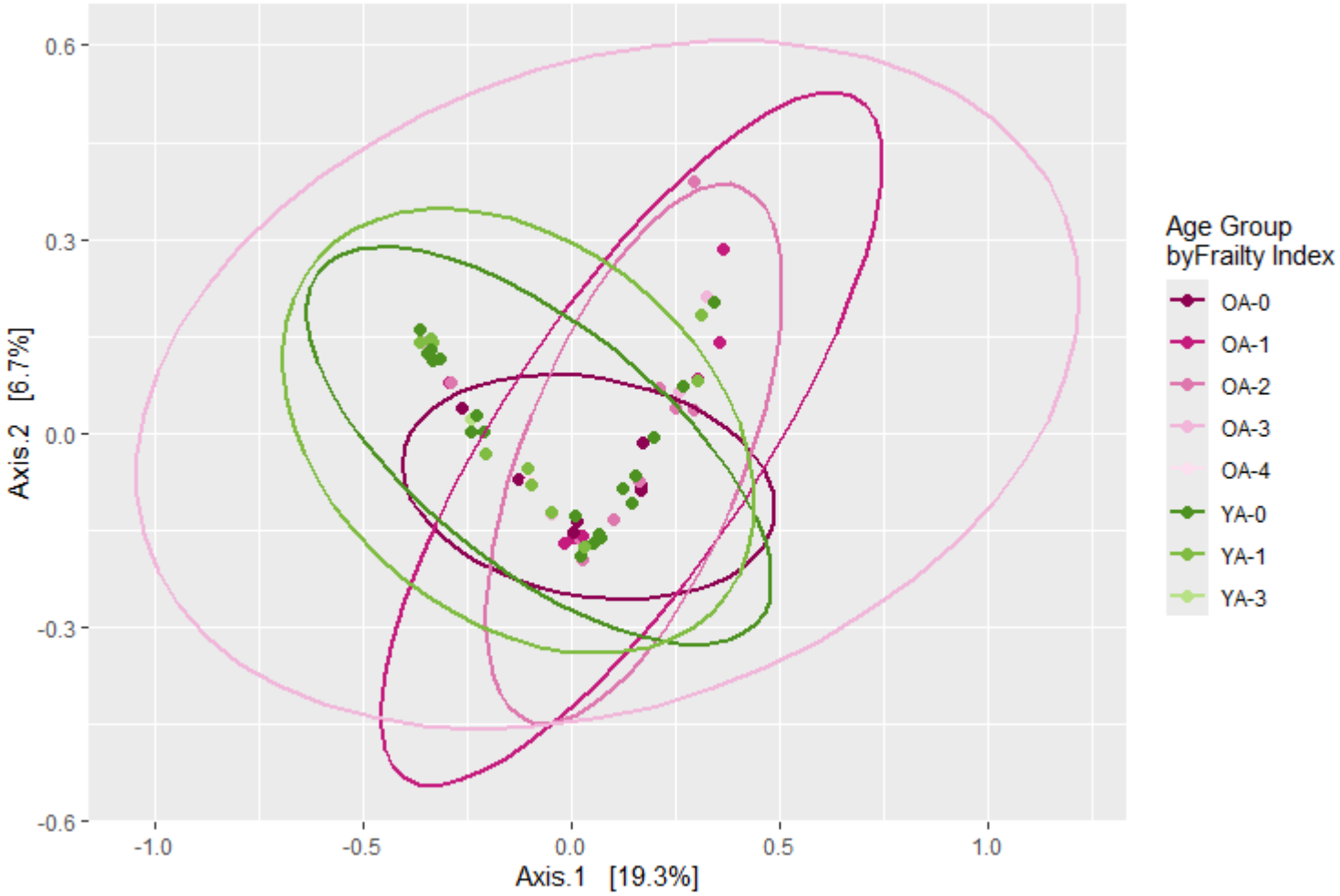


Figure 5.16 Bray Curtis dissimilarity measure of beta diversity as an ordination plot on face samples. *Split by age group and Frailty Index score, where no significant differences were observed.*

Table 5.4 Beta diversity p-values for face sample interactions between the age groups sub-divided into FI scores. *P-values below 0.05 are highlighted in green.*

	Beta diversity measure			
Interaction	Bray Curtis dissimilarity	Jaccard index	Unweighted UniFrac	Weighted UniFrac
OA-0 vs OA-1	1.000	1.000	1.000	1.000
OA-0 vs OA-2	1.000	1.000	1.000	1.000
OA-0 vs OA-3	1.000	1.000	1.000	1.000
OA-0 vs OA-4	1.000	1.000	1.000	1.000
OA-0 vs YA-1	1.000	1.000	0.648	1.000
OA-0 vs YA-0	1.000	1.000	1.000	1.000
OA-0 vs YA-3	1.000	1.000	1.000	1.000
OA-1 vs OA-2	1.000	1.000	1.000	1.000
OA-1 vs OA-3	1.000	1.000	1.000	1.000

OA-1 vs OA-4	1.000	1.000	1.000	1.000
OA-1 vs YA-1	1.000	1.000	1.000	1.000
OA-1 vs YA-0	1.000	1.000	1.000	1.000
OA-1 vs YA-3	1.000	1.000	1.000	1.000
OA-2 vs OA-3	1.000	1.000	1.000	1.000
OA-2 vs OA-4	1.000	1.000	1.000	1.000
OA-2 vs YA-1	1.000	1.000	1.000	1.000
OA-2 vs YA-0	0.864	1.000	0.594	1.000
OA-2 vs YA-3	1.000	1.000	1.000	1.000
OA-3 vs OA-4	1.000	1.000	1.000	1.000
OA-3 vs YA-1	1.000	1.000	1.000	1.000
OA-3 vs YA-0	1.000	1.000	1.000	1.000
OA-3 vs YA-3	1.000	1.000	1.000	1.000
OA-4 vs YA-1	1.000	1.000	1.000	1.000
OA-4 vs YA-0	1.000	1.000	1.000	1.000
OA-4 vs YA-3	Invalid	Invalid	Invalid	Invalid
YA-1 vs YA-0	1.000	1.000	1.000	1.000
YA-1 vs YA-3	1.000	1.000	1.000	1.000
YA-0 vs YA-3	1.000	1.000	1.000	1.000

5.3.2.4. Additional metadata groups

Previously in chapter 3 the additional metadata gathered from the participants was reviewed and analysed, identifying potential groups of interest or relevance to the microbiome in this study either through creating significantly different clusters or of specific sample relevance.

Firstly, the super area output code which groups the participants by their postcodes into convenient clusters of demographics, these showed a significant difference between the OA and YA groups in terms of the number of participants who lived within certain codes (3.3.2). When the arm alpha diversities were clustered using these metrics no significant differences were seen between the age groups or between any super area output code within any age group. However, across all codes the OA group participants tend to cover a larger range of diversities when compared to the YA group who except for 1A in Simpson diversity are more tightly clustered (Fig.5.17). Specifically in the YA group, participants with a code of 1A have lower diversities and evenness but similar observed features to those who live in code 7A. Whereas the in OA group 3B covers a wider range of all metrics when compared to those who are classified as 8A. however, further investigation of these groups and their effects on alpha diversity are beyond the scope of this study, therefore the next stage of analysing these groups was to use beta diversity.

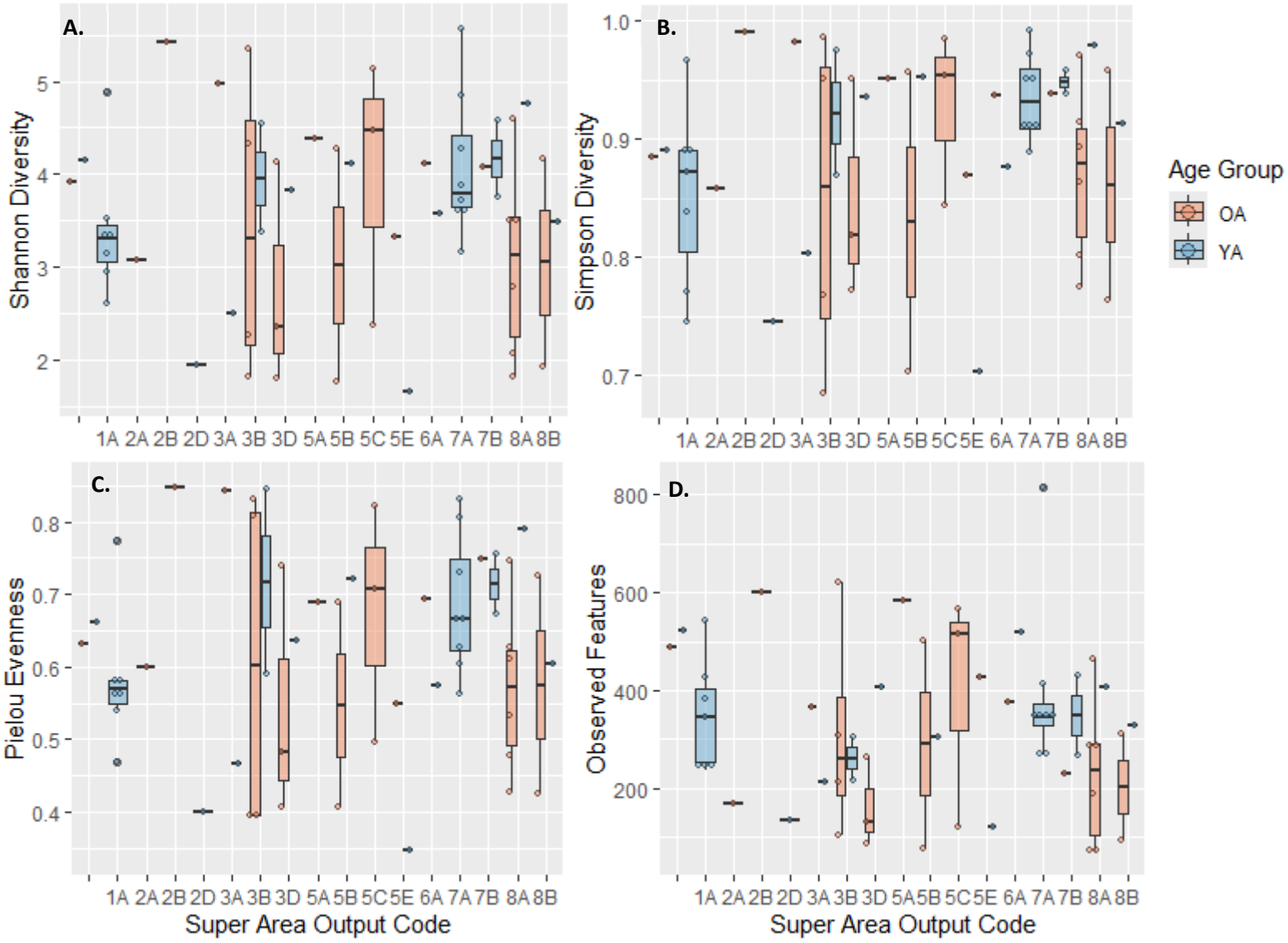


Figure 5.17 Alpha diversity metrics graphs for the arm samples, split by age group and Super area output code. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

The super area output code was then used to cluster for beta diversity, when the arm data is viewed in ordination plots only the YA groups 1A and 7A, and OA groups 3B and 8A are sufficient to generate clustered ellipses (Fig.5.18 and Fig. 9.7). Overall, there are no significant differences between any of the super area output groups in any of the beta diversity metrics. Therefore, super area output code did not have any effect on the participant arm microbiome diversities.

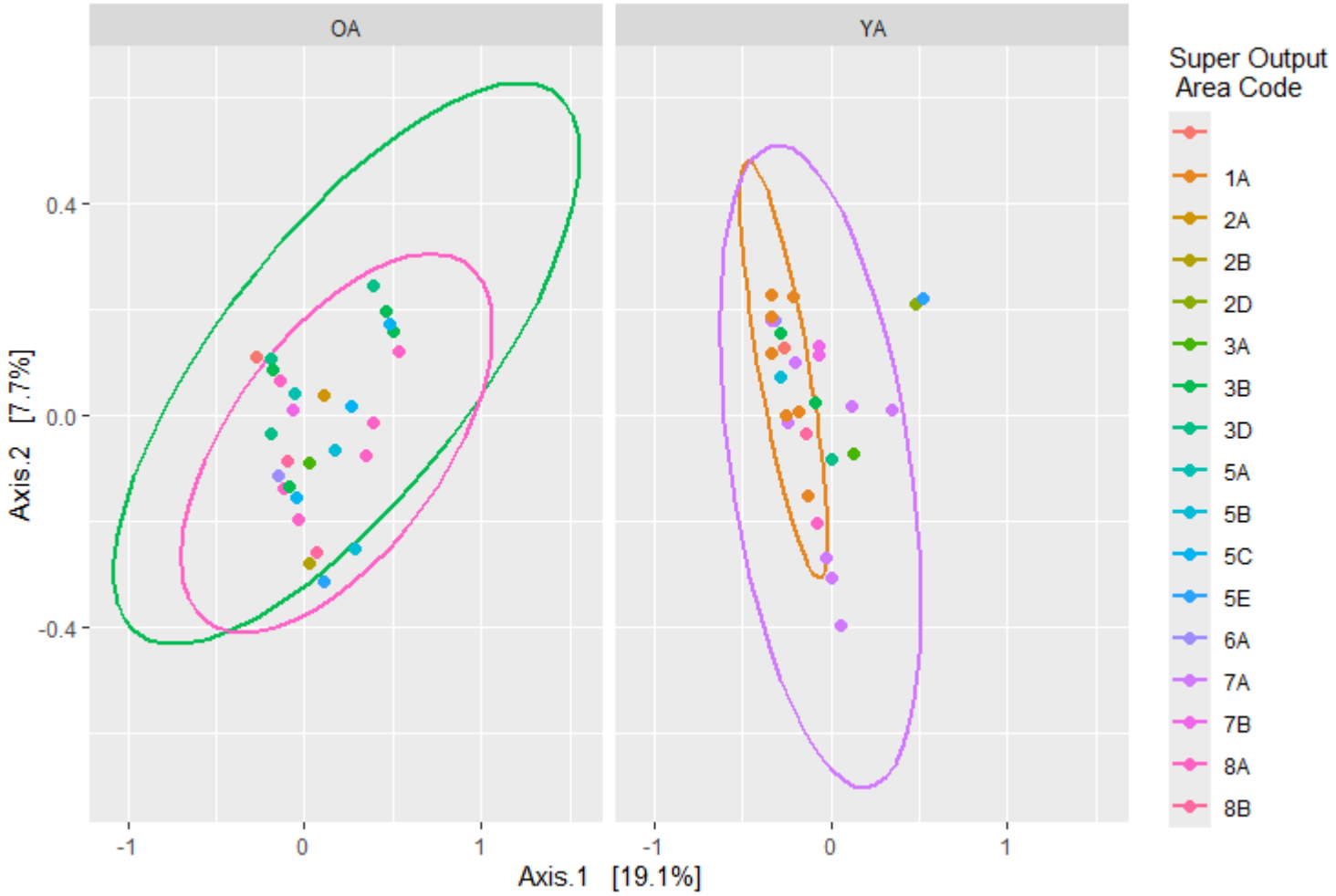


Figure 5.18 Bray Curtis dissimilarity measure of beta diversity as an ordination plot on arm samples. *Split by age group and super area output code, where no significant differences were observed.*

Viewing the face sample alpha diversity metrics with the super area output code clustering again no significant differences were observed either between or within the age groups and the super area codes. Overall, the YA 1A and OA 3B participants cover the largest ranges within all the metrics, while YA 7A and OA 8A are both higher than the diversity and evenness scores than the YA 1A and OA 3B respectively; OA 8A has a tightly clustered observed features which nestles within the OA 3B range (Fig.5.19). Again, clustering the face samples using the super area output code when applied to the beta diversity metrics revealed no significantly different interactions, this is despite the creation of clearly distinct groups of YA 1A and 7A, and OA 3B and 8A on the ordination plots (Fig.5.20 and Fig. 9.8). Therefore, super area output code did not have any effect on the participant face microbiome diversities.

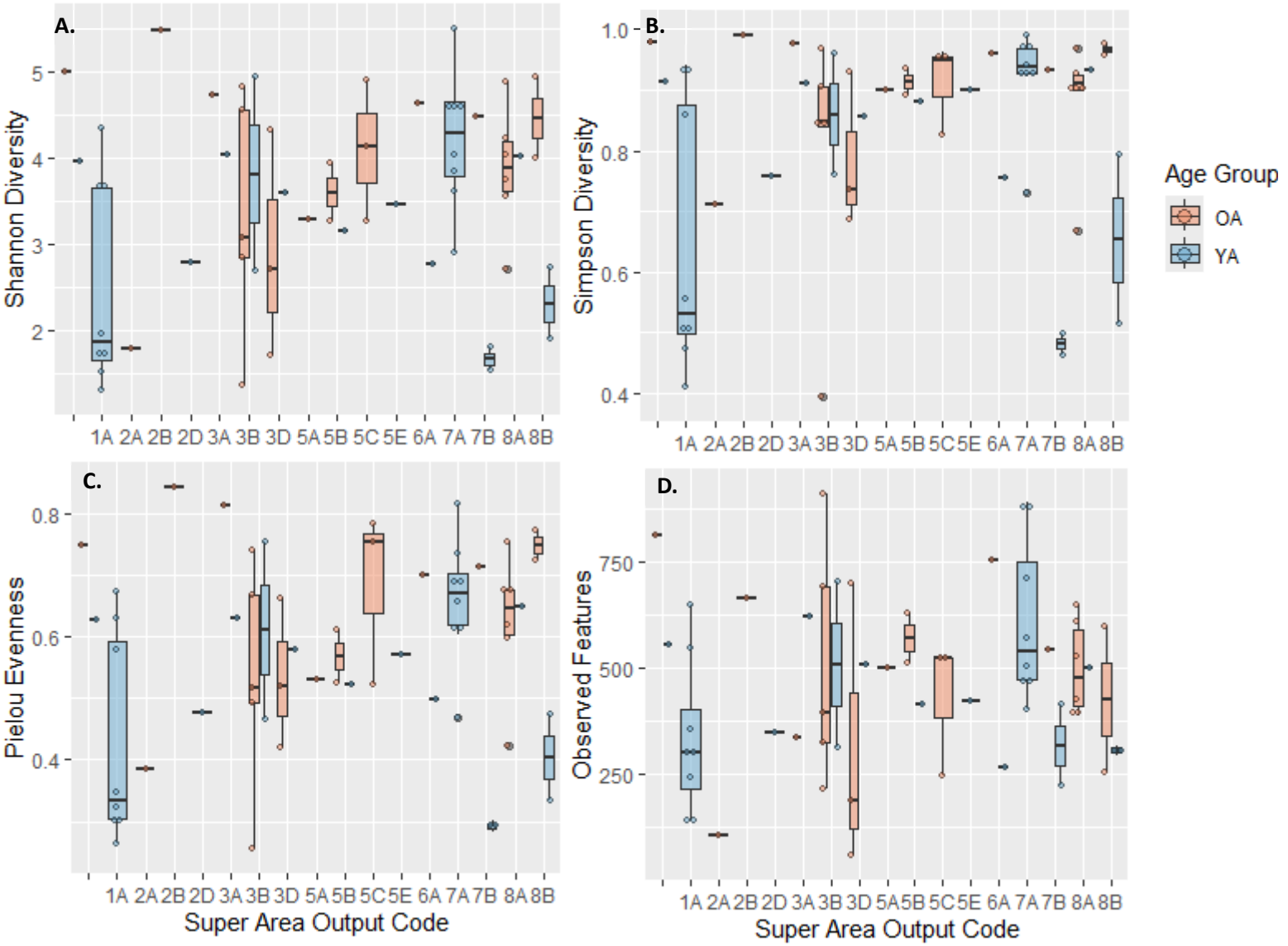


Figure 5.19 Alpha diversity metrics graphs for the face samples, split by age group and Super area output code. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

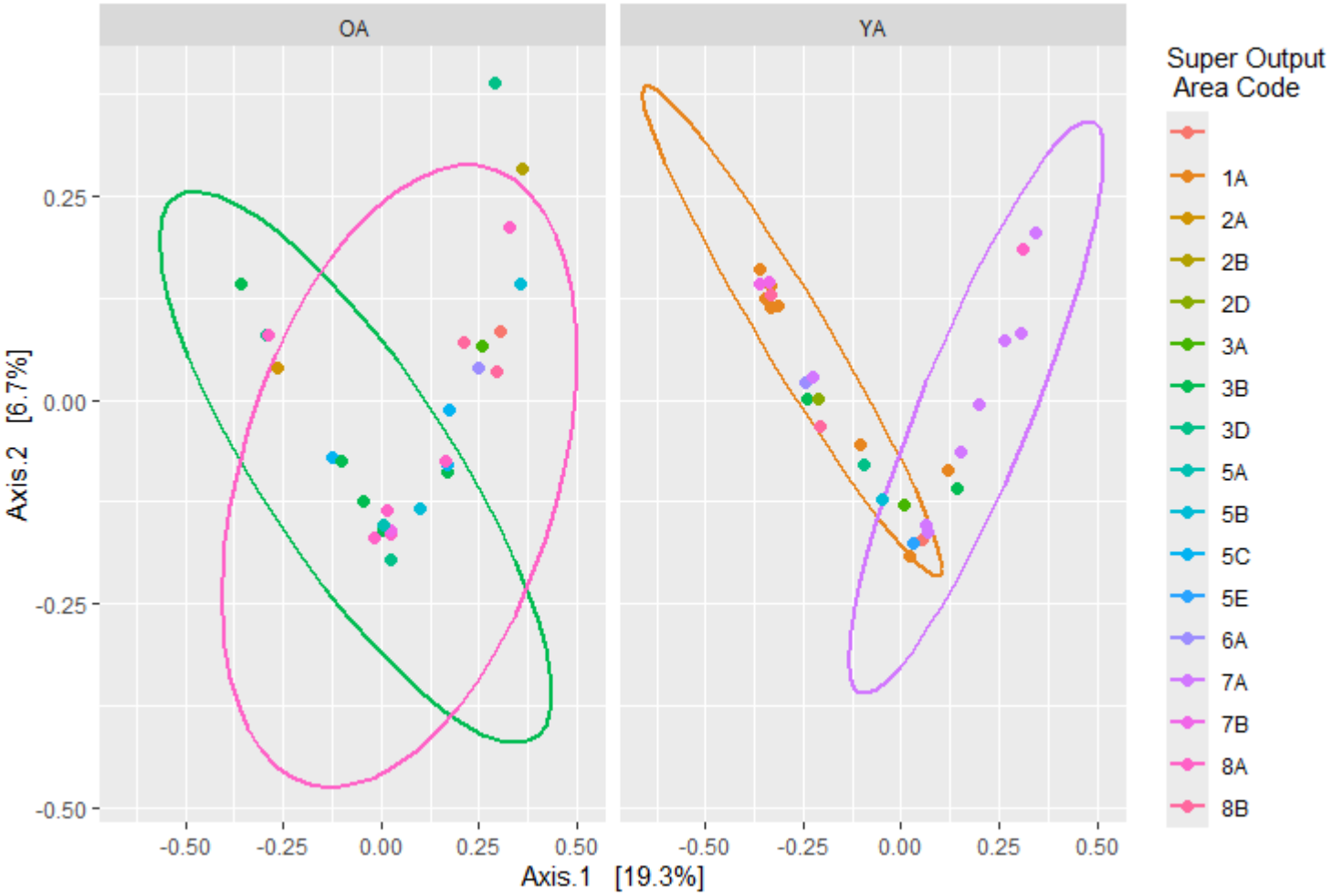


Figure 5.20 Bray Curtis dissimilarity measure of beta diversity as an ordination plot on face samples. *Split by age group and super area output code, where no significant differences were observed.*

Another sociodemographic tool that can be used to cluster the participants is the workplace zone code, this indicates the types of industry that operate within the area of the participants addresses. The participants showed a significant difference between the age groups when clustered with using this metric, additionally pollutant exposure has shown to affect the microbiome; therefore, it is pertinent to explore any contribution in this study.

Applying this metric to the arm alpha diversity data sees no significant differences between the age groups or within the workplace zone codes in the age groups (Fig.5.21). The YA B2 group is more tightly clustered than most of the other YA groups in Shannon and Simpson diversity, and Evenness, by these metrics it is also lower than most other YA groups; there is however little difference when comparing observed features. Comparatively all OA groups show a large range across all workplace zone codes and diversity metrics used. Overall, YA B2 and OA D2 are approximately similar across all metrics, but with no significant differences observed and the scope of this study being limited these clusters cannot be investigated further.

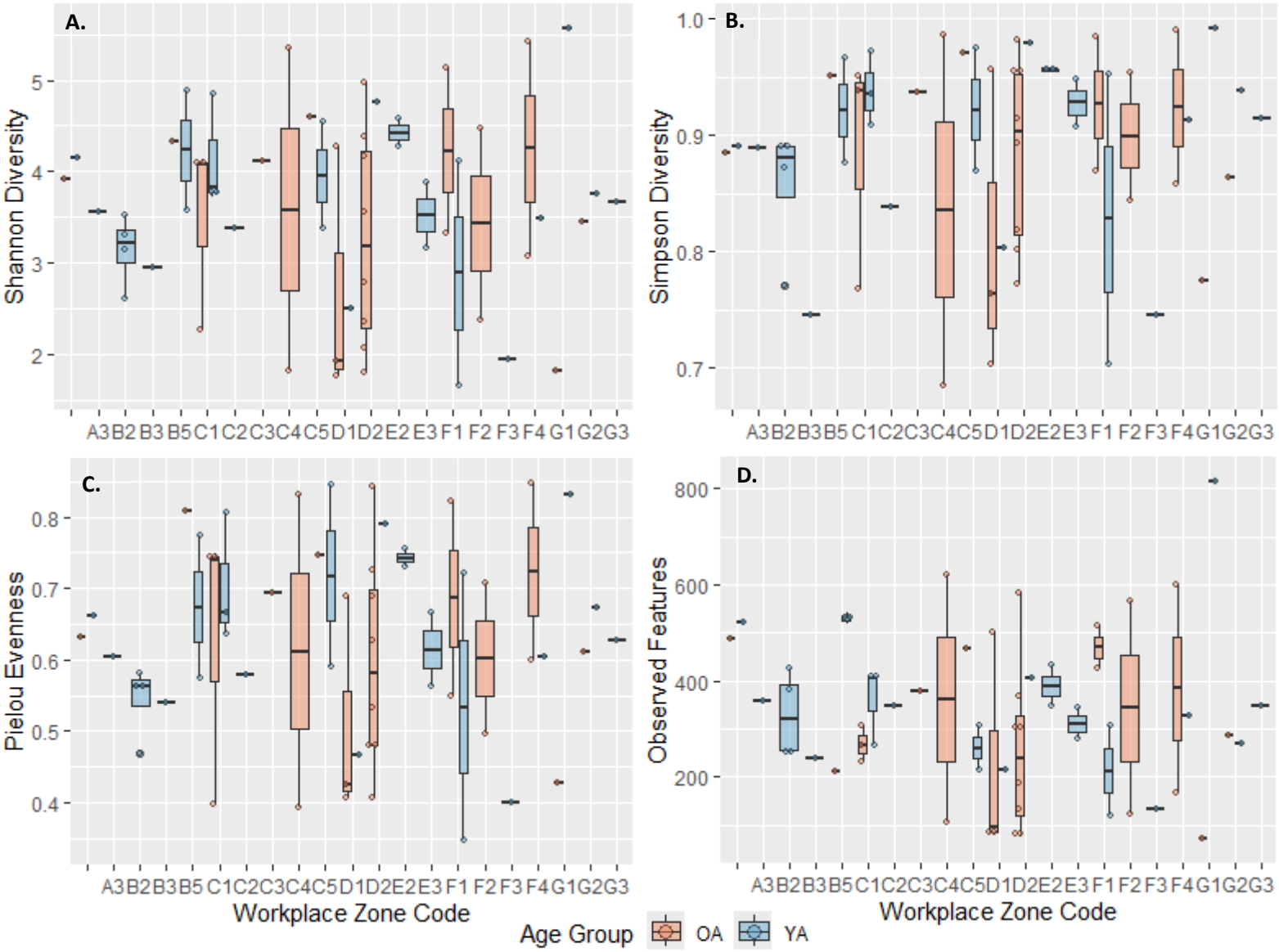


Figure 5.21 Alpha diversity metrics graphs for the arm samples, split by age group and workplace zone code. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

When used in the beta diversity clustering of the arm samples only the YA B2 and OA D2 contained enough participants for generation clustered ellipses in the ordination plots (Fig.5.22, and Fig. 9.9). Again, no significant differences were seen in any interaction across the metrics used, therefore, workplace zone code had no impact on arm alpha and beta diversities.

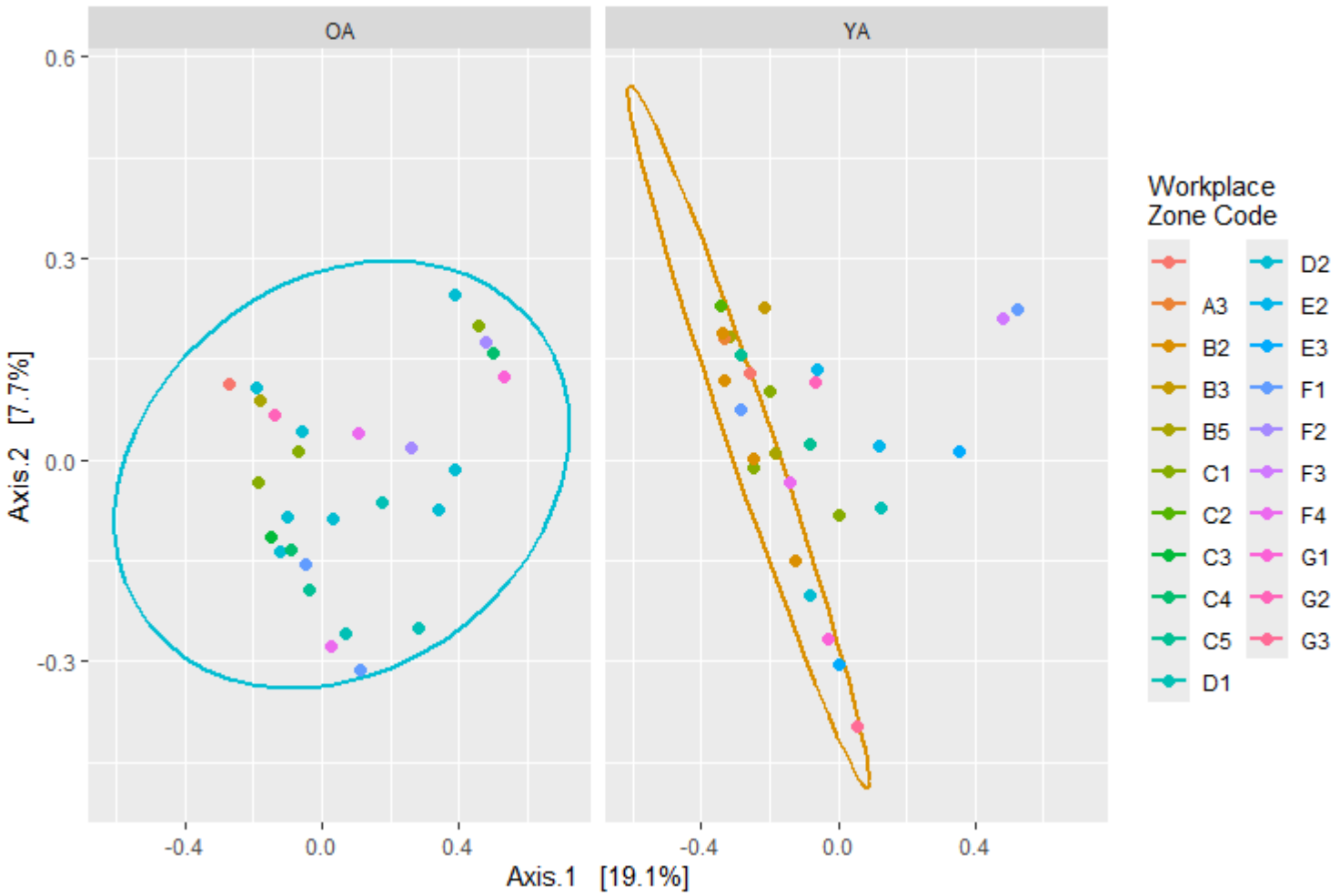


Figure 5.22 Bray Curtis dissimilarity measure of beta diversity as an ordination plot on arm samples. *Split by age group and workplace zone code, where no significant differences were observed.*

Using the workplace zone codes on the face samples shows no significant differences between the age groups or within the workplace zone codes in the age groups (Fig.5.23). Here the YA B2 group had a wider range than most of the other YA groups, but also averaging a lower score than them. Similarly, the OA D2 group had a wide range that sits within the ranges covered by some of the other OA groups and averages the middle of the metrics used here. Overall, the OA D2 has higher scores than the YA B2 across these metrics, but again the limited scope of this study prevented these from being investigated any further.

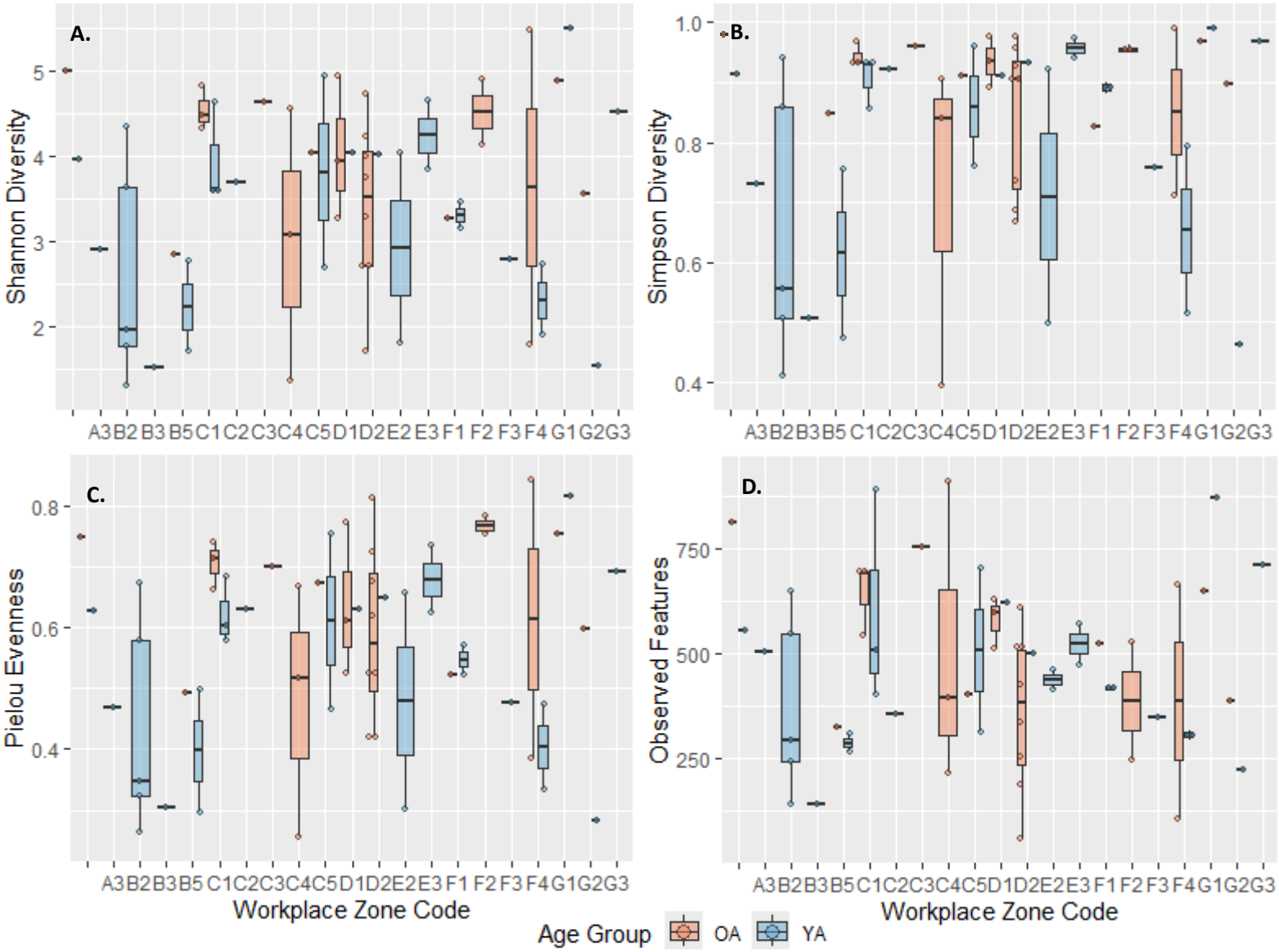


Figure 5.23 Alpha diversity metrics graphs for the face samples, split by age group and workplace zone code. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

When looking at the beta diversity metrics using the workplace zone codes on the face samples again no significant differences are seen in any interaction across the metrics used (Fig.5.24 and Fig. 9.10), despite some distinct clusters being made of YA B2 and OA D2 in the ordination plots. Therefore, workplace zone code had no impact on face alpha and beta diversities.

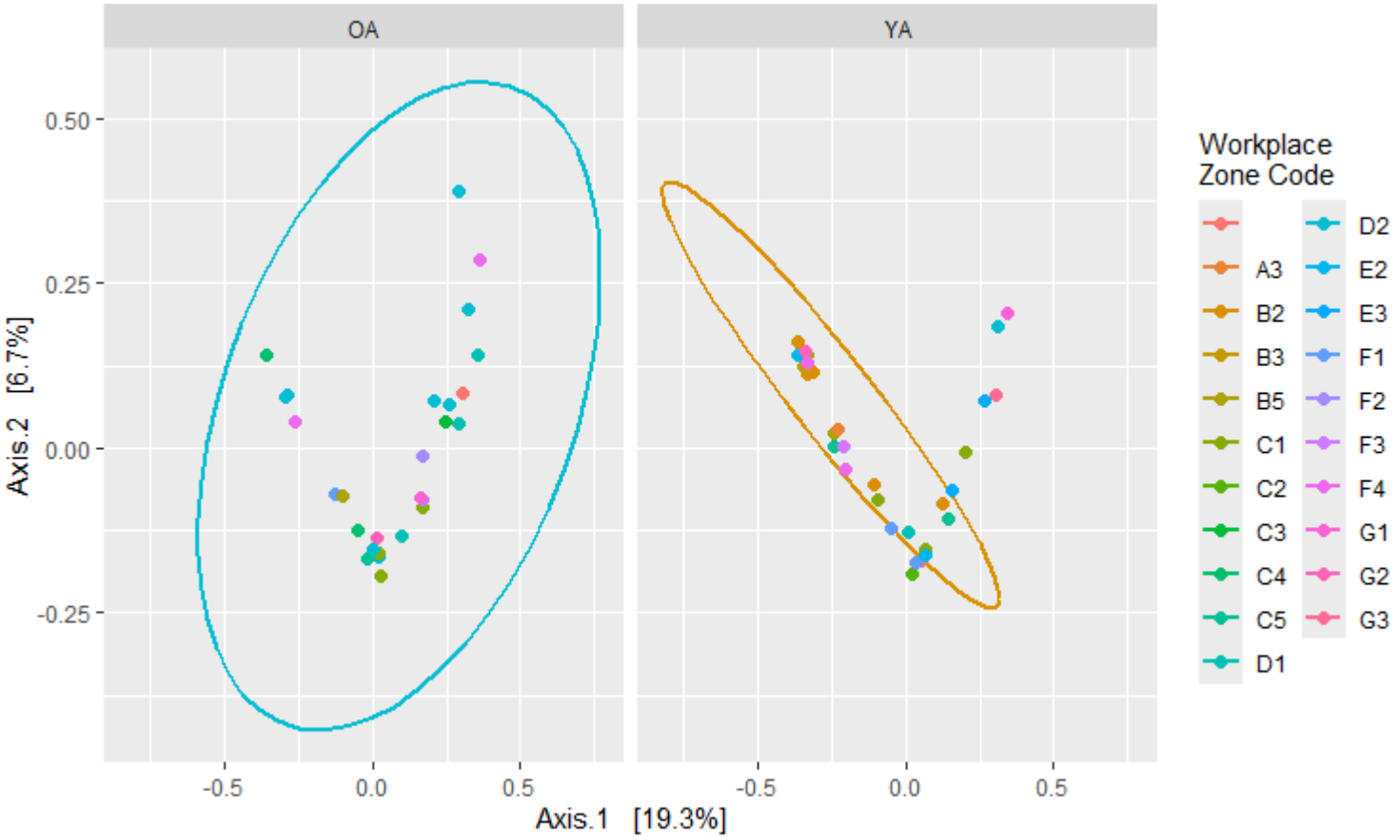


Figure 5.24 Bray Curtis dissimilarity measure of beta diversity as an ordination plot on face samples. *Split by age group and workplace zone code, where no significant differences were observed.*

The AUDIT-C score is a measure relating to participant alcohol consumption, which has an established affect the microbiome in other studies. Here the participants had significantly different scores between the YA and OA groups; with the OA scoring higher. This can then be further subdivided by gender where the YA-Females made up the majority of the YA low AUDIT-C scores.

The arm sample alpha diversity metrics when grouped using this alcohol consumption metric did not show any significant differences between the age groups or when grouped by age group and gender (Fig.5.25). Across Shannon and Simpson diversity and Pielou evenness the OA-Females and YA-Males contribute to the largest ranges of diversities at different AUDIT-C scores, whereas the OA-Males and YA-Females are comparatively tightly clustered. But the low number of participants across all these scores prevented the analysis of any of these clusters of alpha diversity being investigated further.

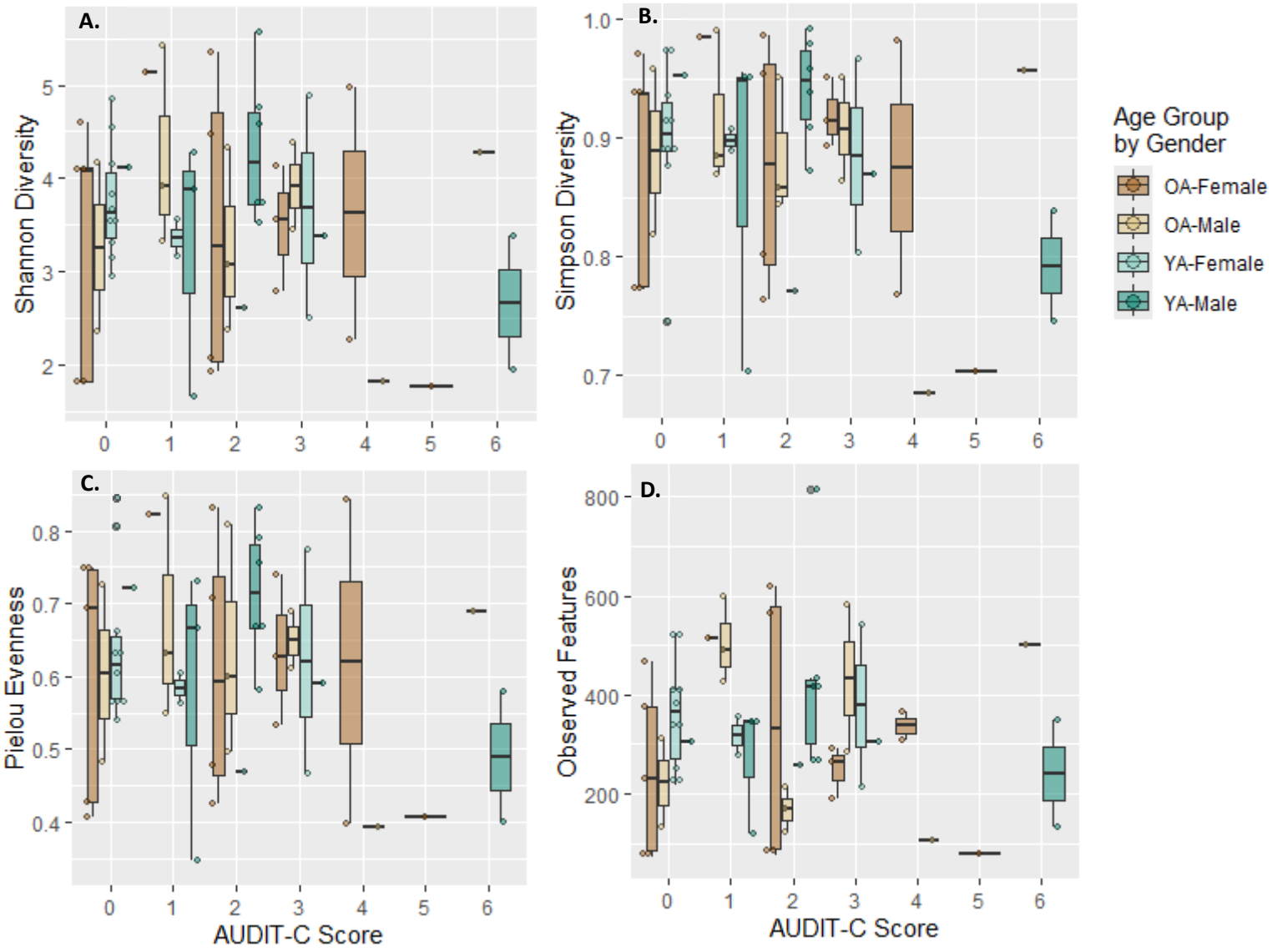


Figure 5.25 Alpha diversity metrics graphs for the arm samples, split by age group, gender, and AUDIT-C Score. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

Applying the AUDIT-C score to the beta diversity analysis creates clustered ellipses for those scores of zero and two in the OA-Females, and only zero or two in the YA-Females and YA-Males respectively (Fig.5.26 and Fig. 9.11). However, no significant differences are seen between any interactions of the various groups AUDIT-C scores. Therefore, AUDIT-C score does not have an effect on the arm alpha diversities in this study.

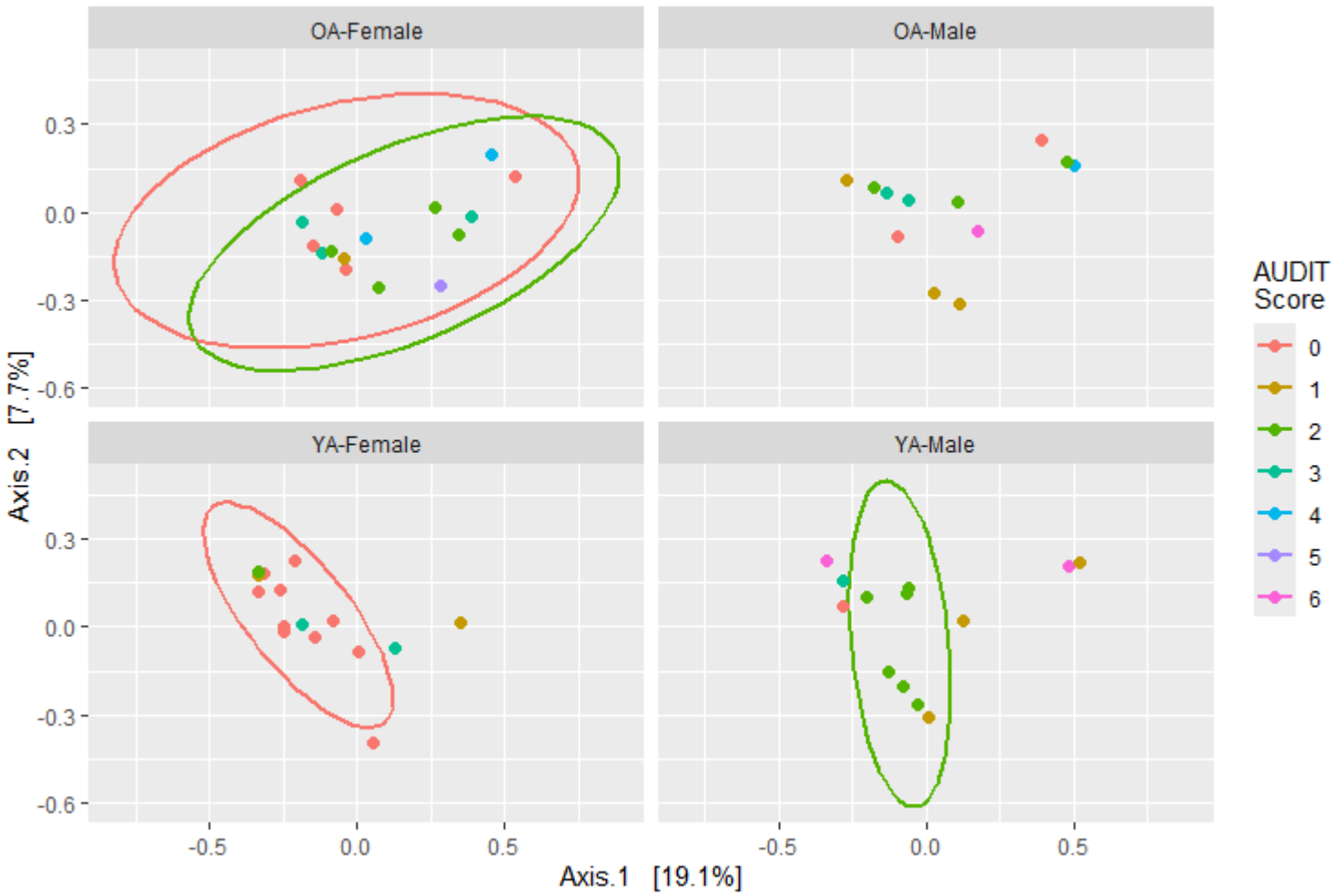


Figure 5.26 Bray Curtis dissimilarity measure of beta diversity as an ordination plot on arm samples. *Split by age group, gender and AUDIT-C Score, where no significant differences were observed.*

Face sample alpha diversity data again no significant differences were seen when clustered using the AUDIT-C scores (Fig.5.27). The ranges covered by the YA diversity and evenness scores at AUDIT-C zero are mirrored in scores two and three for the male and female groups respectively. In the OA groups the females showed consistently high diversities in scores zero, two, and four, but a lower diversity score in those who score three on the scale, across Shannon and Simpson diversities, and Pielou evenness; the OA-Males are consistent across these metrics. In observed features the OA-Females ranges overlap consistently across all scores whereas the other groups show more variation. But again, the low number of participants across all these scores prevented the analysis of any of these clusters of alpha diversity being investigated further.

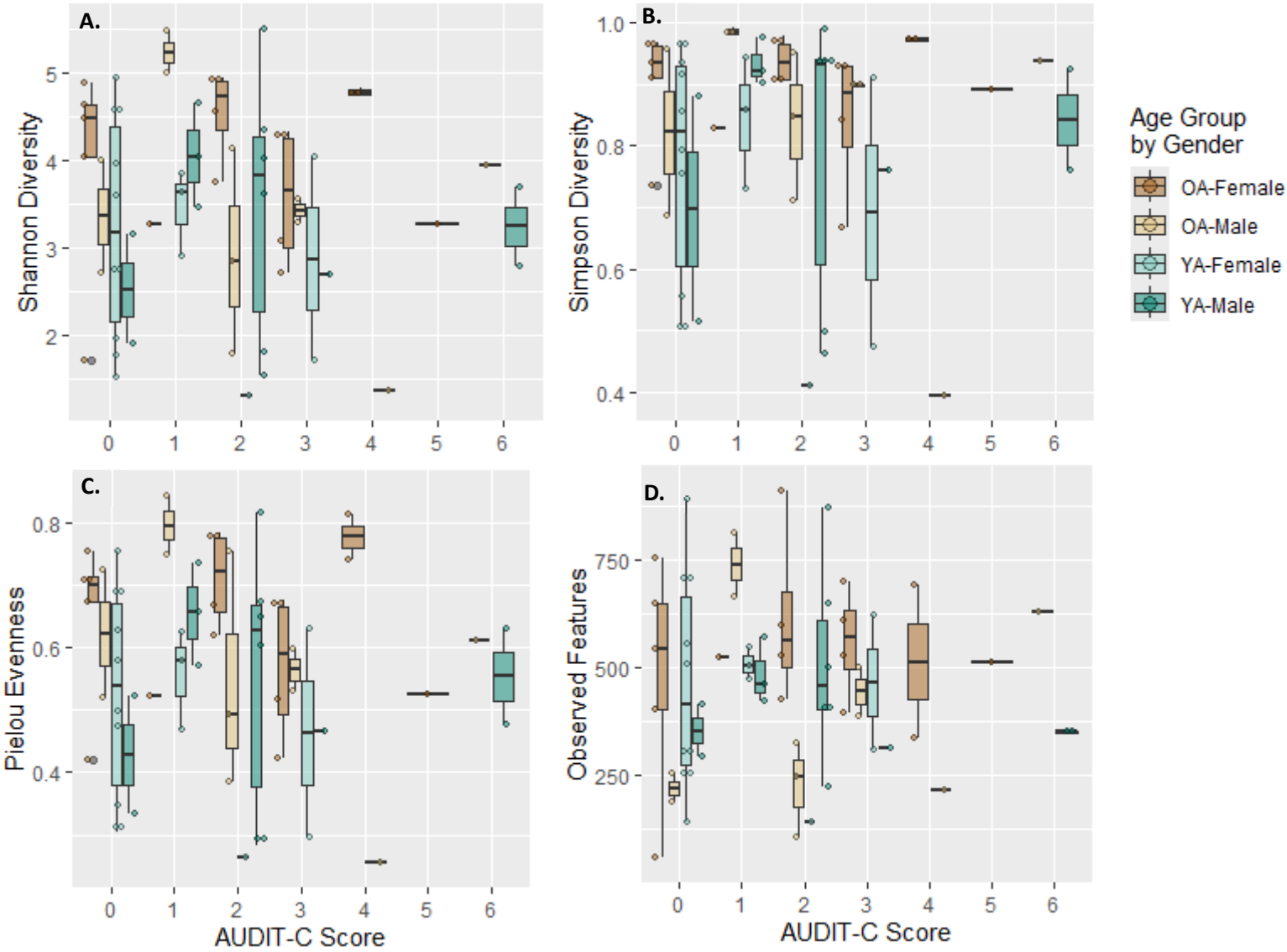


Figure 5.27 Alpha diversity metrics graphs for the face samples, split by age group, gender, and AUDIT-C Score. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

Beta diversity analysis using the AUDIT-C score on the face samples creates clustered ellipses for those scores of zero, two, and three in the OA-Females, and only zero or two in the YA-Females and YA-Males respectively (Fig.5.28 and Fig. 9.12). However, no significant differences are seen between any interactions of the various groups AUDIT-C scores. Therefore, the AUDIT-C score did not affect the alpha or beta diversities of the face microbiome samples within this study.

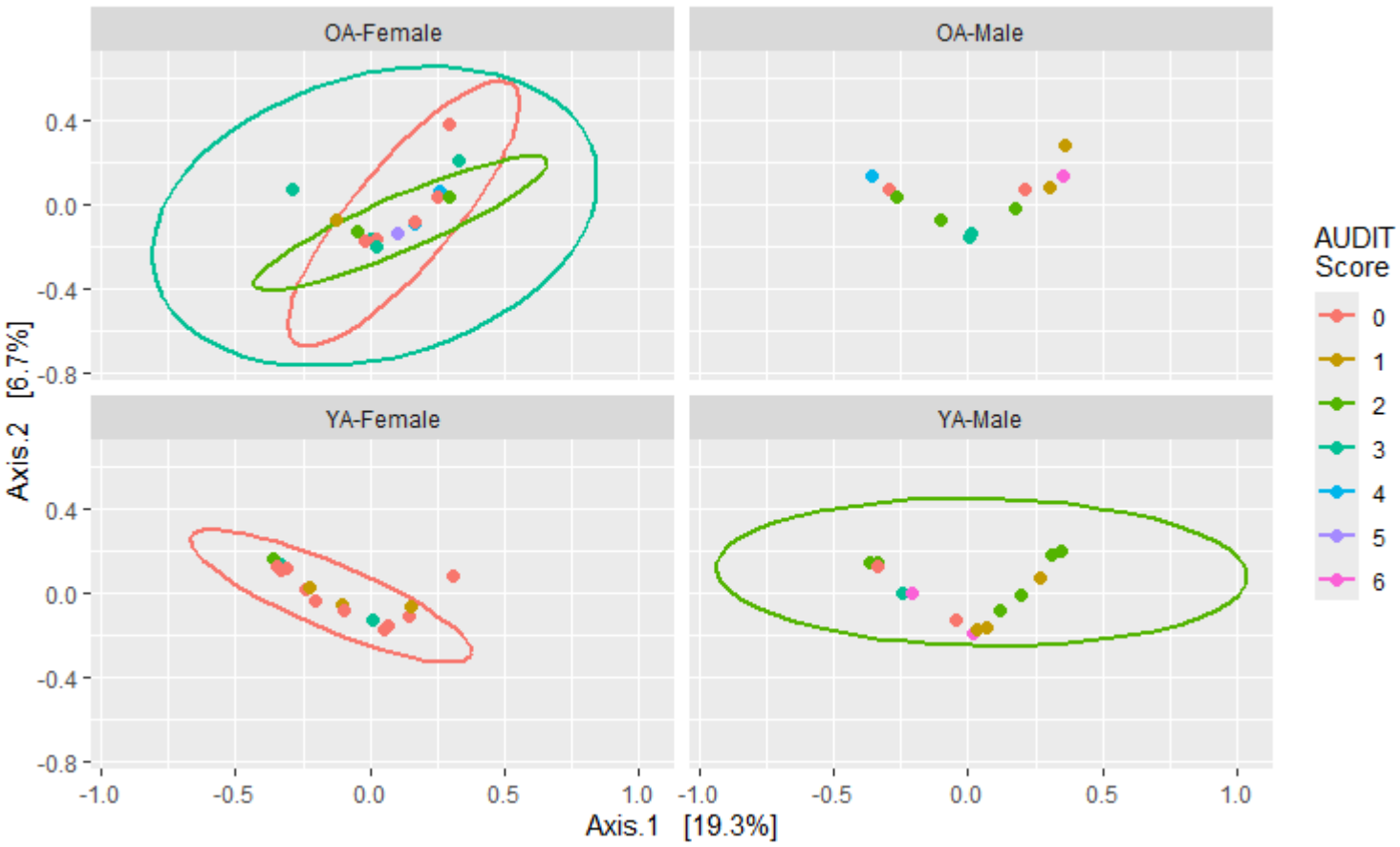


Figure 5.28 Bray Curtis dissimilarity measure of beta diversity as an ordination plot on face samples. *Split by age group, gender and AUDIT-C Score, where no significant differences were observed.*

Finally, potentially relevant participant metadata for analysing the skin included participant skin health and skin habits. Here while no significant groups of participants reported currently or in the past suffering from a skin condition, it was deemed appropriate to use the past skin condition metric as this may have led to changes in participant behaviour regarding skin health which could be significant (Fig. 3.8); these also generated the largest groups of participants for this. Of the skin hygiene and habit data, moisturiser usage showed significant differences when viewed using gender and through those who reported suffering from prior skin conditions (Fig. 3.13).

When applied to the arm sample alpha diversity data, past skin condition groups did not show any significant differences between them using the age groups or when further sub-divided by gender (Fig.5.29). However, those who featured in the no group have overlapping ranges when clustered by age group and gender, while those OA-Females and YA-Females who reported suffering other conditions in the past have diversity ranges below the no groups, meanwhile YA-Males who reported suffering from acne infections in the past had diversity ranges below the no groups while the YA-Females reporting acne infections are have similar or higher ranges. The reported moisturiser use does not show any significant differences in alpha diversity metrics,

generally there is a large overlap between the age groups and genders across these metrics regardless of moisturiser use. However, those OA-Males that reported using moisturiser have lower diversities than those who do not use moisturiser. While both metadata metrics showed clustering effects as there was no significant difference observed and given the scope of this study these cannot be taken any further.

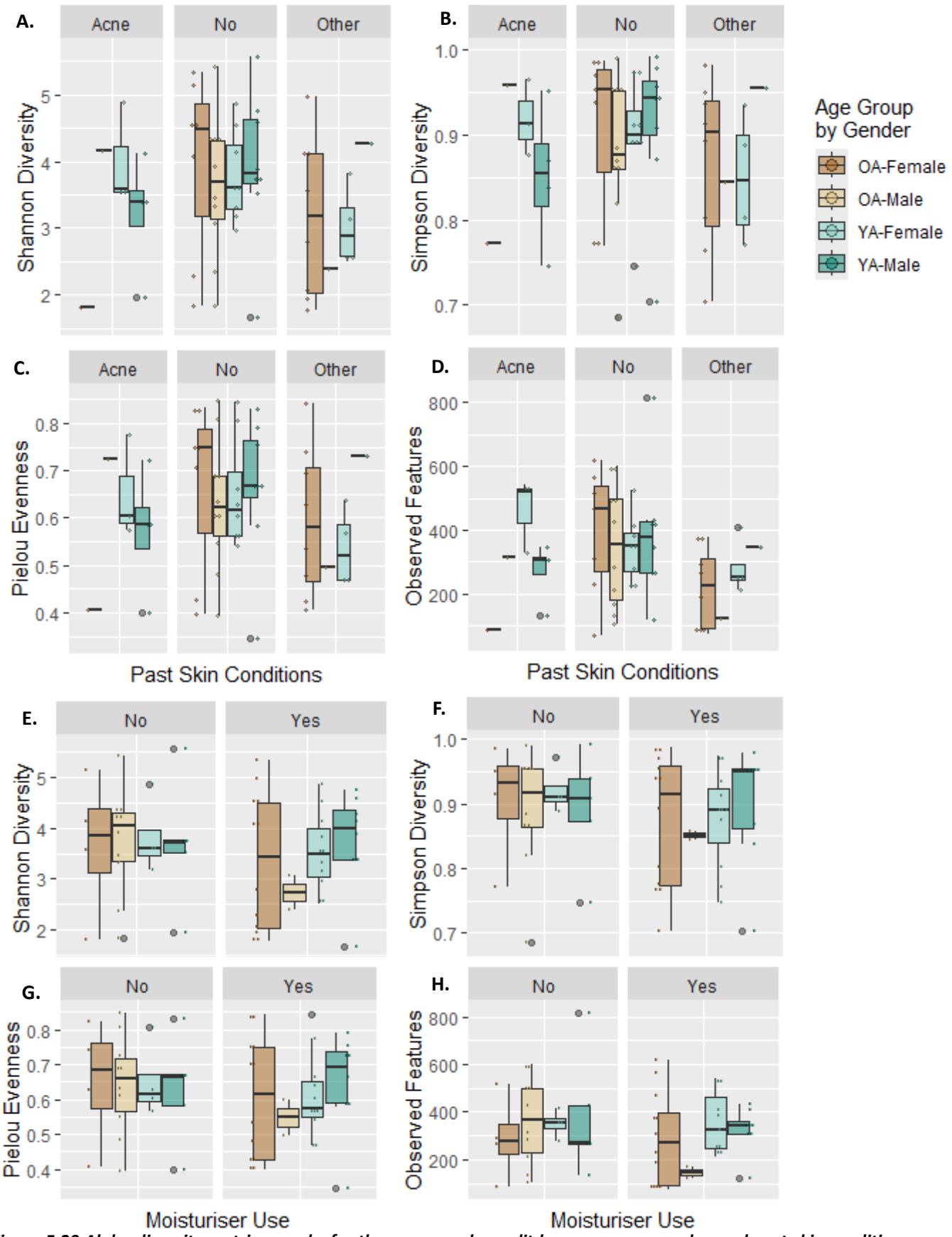


Figure 5.29 Alpha diversity metrics graphs for the arm samples, split by age group, gender, and past skin conditions or moisturiser use. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features. E. Shannon Diversity. F. Simpson Diversity. G. Pielou Evenness. H. Observed Features.*

Using the past skin condition metric on the arm samples beta diversity metrics, creates clusters for the no and other groups in the female participants, no and acne group in the YA-male participants, but only the no OA-Male participants. However, no significant differences are seen between these groups (Fig. 5.30, Fig. 9.13, and Fig. 9.14). Moisturiser use creates clusters for the yes and no responses in all bar the OA-Male group, there is also a significant difference between the YA-Female and OA-Females that responded yes, in both the Bray Curtis dissimilarity and Jaccard index but not in the UniFrac metrics. Therefore, those females who reported moisturiser use arm microbiomes had differences between the taxa observed and their abundances.

For the face samples alpha diversities using the past skin conditions no significant differences are seen (Fig.5.31). Like the arm sample data, the no responses had an overlapping clustering between the age group genders, the YA-Females who belong in the other and acne groups however had lower ranges than those in the no group; while the YA-Males and OA-Females in these groups overlap with the no group. The moisturiser use groups again showed no significant differences across the alpha diversity metrics, however, in all metrics the OA-Female yes participants have higher diversities than all the other participant groups in either the yes or no grouping. The YA-Females that are in the no group have higher diversity scores than those in the yes group, comparatively the male groups generally overlap regardless of moisturiser usage across nearly all metrics. The exception is in the observed features where the OA-Males in the yes group are lower than the no counterparts. While both metadata metrics showed clustering effects as there was no significant difference observed and given the scope of this study these cannot be taken any further.

The past skin groups when applied to the face samples beta diversities again create clusters for the no and other groups in the female participants, no and acne group in the YA-male participants, but only the no OA-Male participants (Fig.5.32, Fig. 9.15, and Fig. 9.16); again, no significant differences are seen. Again, moisturiser use creates clusters for the yes and no responses in all bar the OA-Male group, there is also a significant difference between the YA-Female and OA-Females that responded yes, only in the unweighted UniFrac metric but none is seen in any other metric. Therefore, those females who reported moisturiser use face microbiomes had differences between the taxa observed.

Overall, across all the groupings seen here significant differences are mostly observed using beta diversity metrics, additionally these tend to be seen in the age, and gender groups (Table 5.6). This may be due to these groups consisting of the largest number of participants enabling for better statistical comparisons.

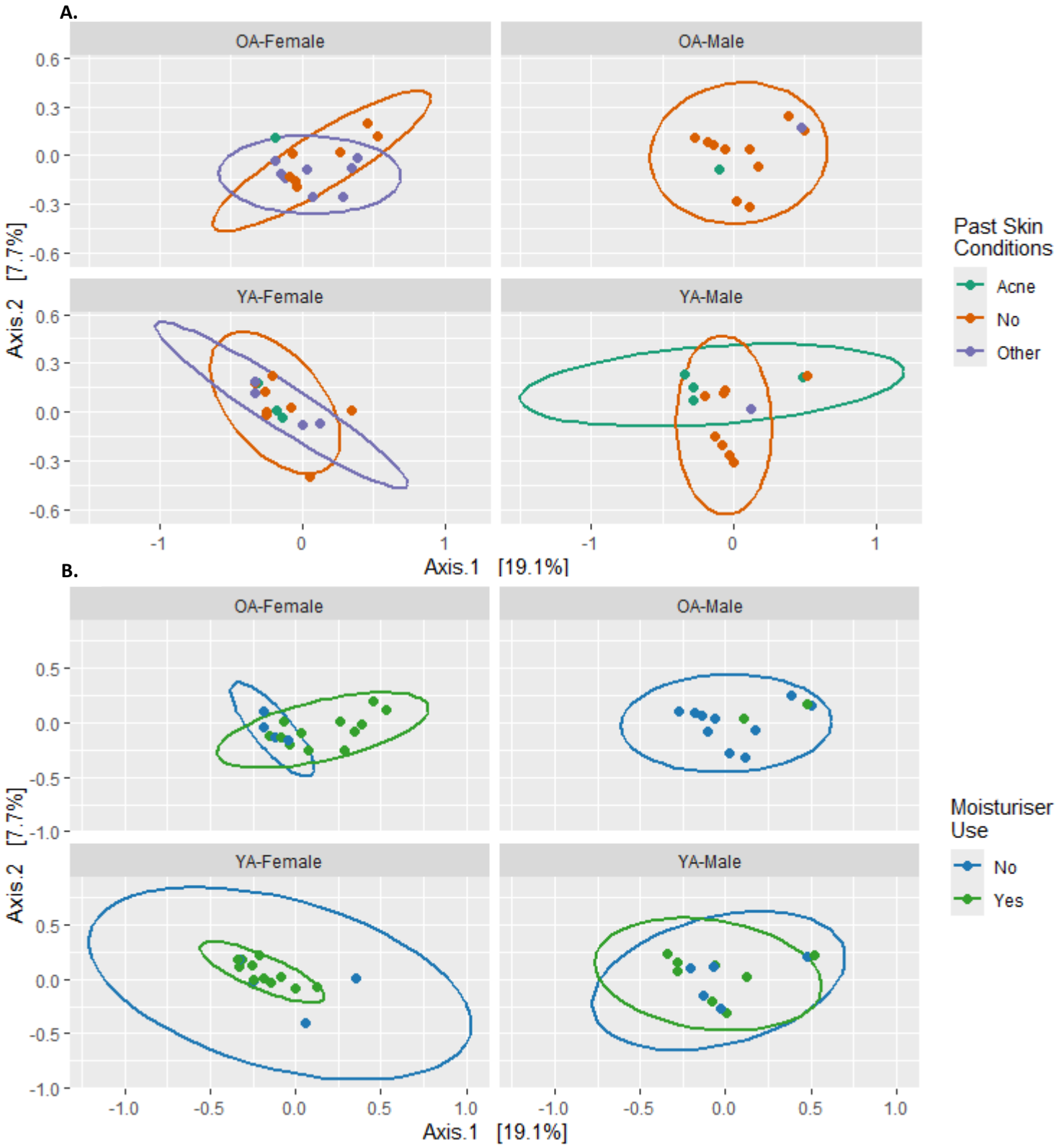


Figure 5.30 Bray Curtis dissimilarity measure of beta diversity as an ordination plot the arm samples. *Split by age group, gender, and past skin conditions or moisturiser use where a significant difference was observed. A. Past skin conditions. B. Moisturiser use a significant difference was observed between OA-Female vs YA-Female (Yes) (p=0.028).*

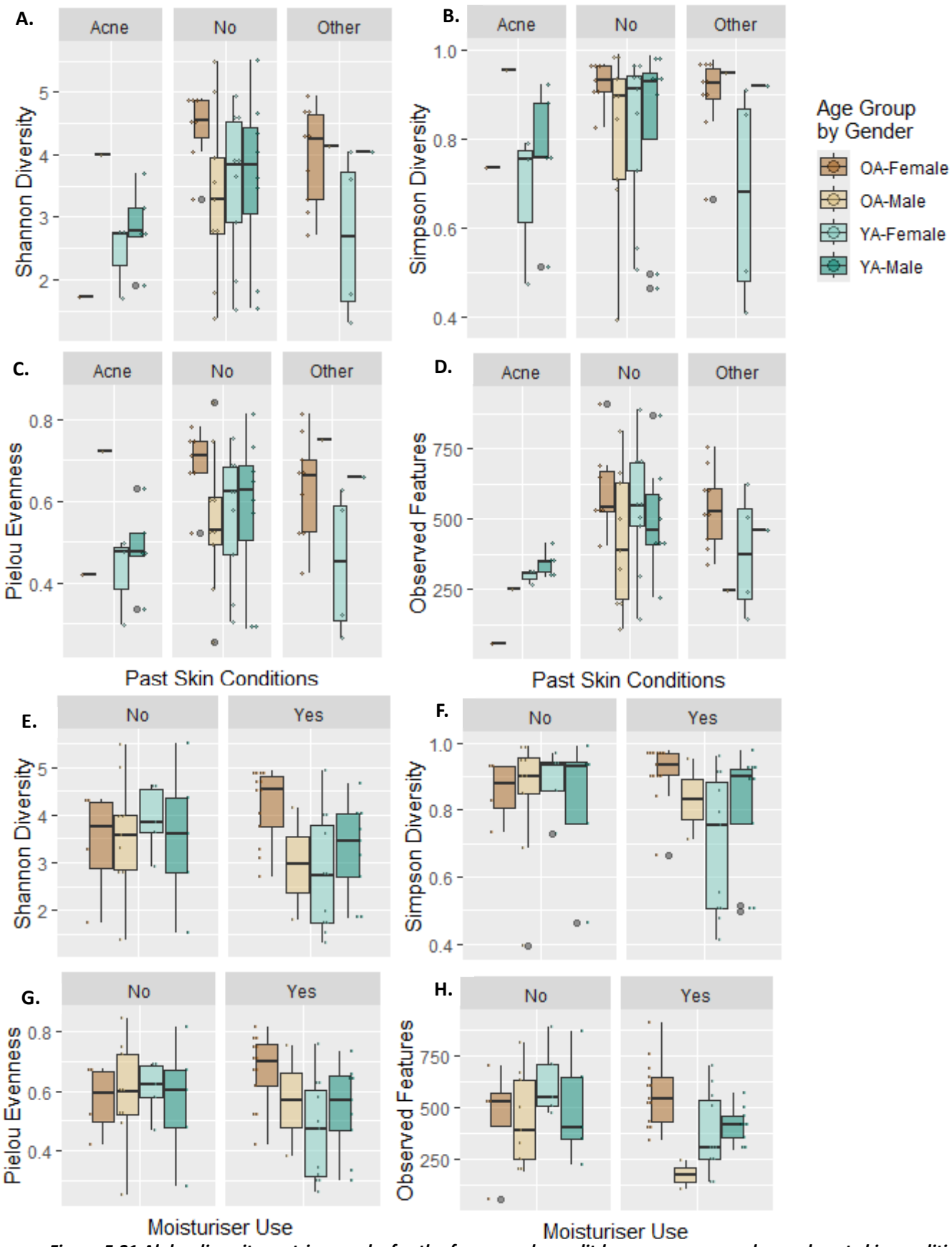


Figure 5.31 Alpha diversity metrics graphs for the face samples, split by age group, gender, and past skin conditions or moisturiser use. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features. E. Shannon Diversity. F. Simpson Diversity. G. Pielou Evenness. H. Observed Features.*

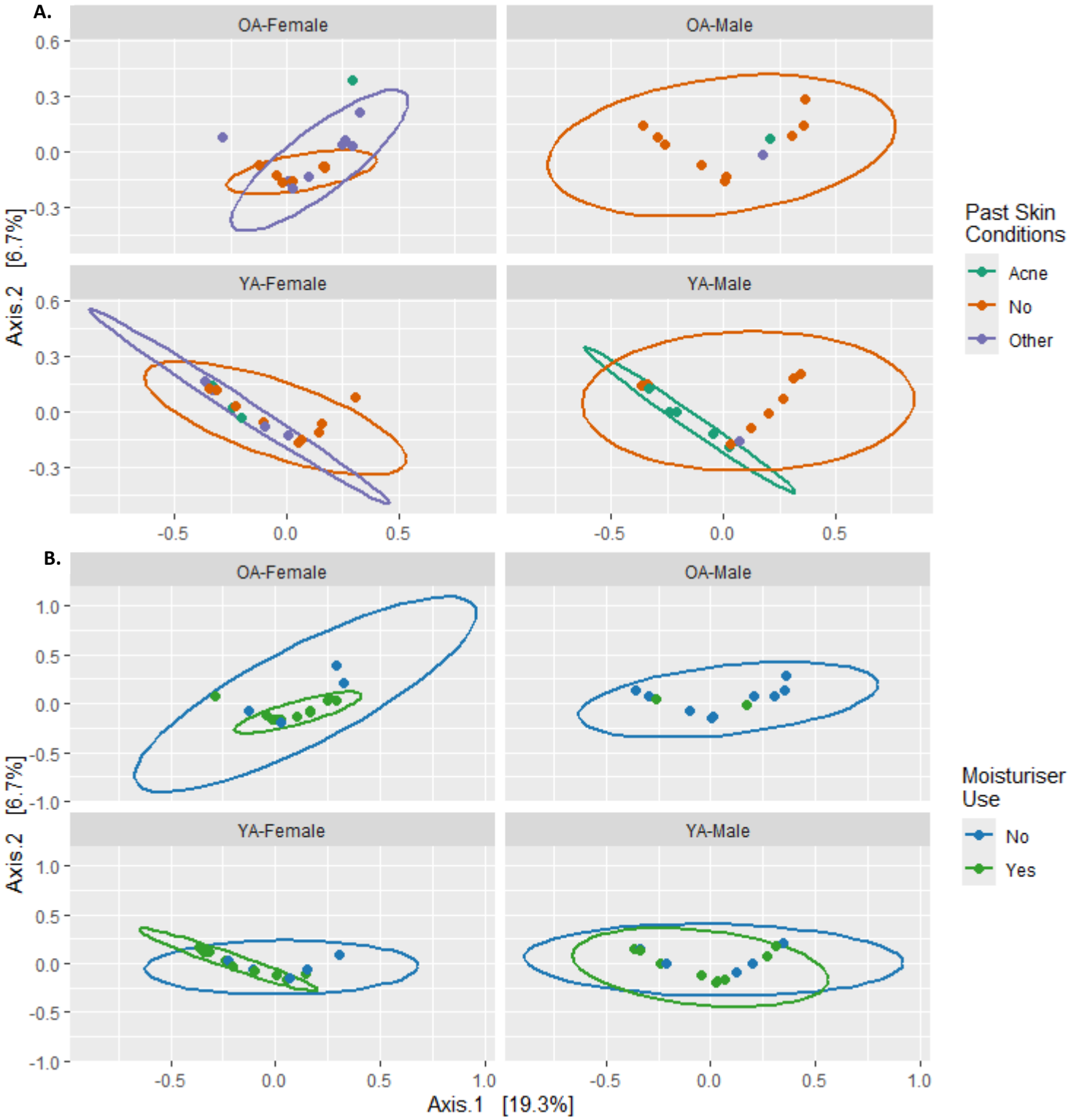


Figure 5.32 Bray Curtis dissimilarity measure of beta diversity as an ordination plot the face samples, split by age group, gender, and past skin conditions or moisturiser use. *Where a significant difference was observed. A. Past skin conditions. B. Moisturiser use a significant difference was observed between OA-Female vs YA-Female (Yes) (p=0.028).*

Table 5.5 Summary of alpha and beta diversity significant interactions for both the arm and face samples. ***Detailing the diversity metric and p-value where appropriate.***

	Arm			Face		
Grouping	Significant interaction	Diversity metric	p-value	Significant interaction	Diversity metric	p-value
Age Group	OA vs YA	Bray Curtis	0.002	OA vs YA	Pielou Evenness	0.024
		Jaccard Index	0.004		Bray Curtis	0.009
		Unweighted UniFrac	0.001		Jaccard Index	0.007
		Weighted UniFrac	0.007		Unweighted UniFrac	0.001
Gender	OA-Female vs YA-Female	Bray Curtis	0.012	OA-Female vs YA-Female	Bray Curtis	0.012
		Jaccard Index	0.012		Jaccard Index	0.018
		Unweighted UniFrac	0.006		Unweighted UniFrac	0.006
		Weighted UniFrac	0.006	OA-Female vs OA-Male		0.036
	OA-Male vs YA-Female	Bray Curtis	0.042	OA-Female vs YA-Male		0.030
		Unweighted UniFrac	0.018			
FI	OA-2 vs YA-0	Unweighted UniFrac	0.027	No significant differences seen in any interaction		
Super Area Output Code	No significant differences seen in any interaction					
Workplace Zone Code						
AUDIT-C score						
Past Skin conditions						
Moisturiser use	OA-Female vs YA-Female (Yes)	Bray Curtis	0.028	OA-Female vs YA-Female (Yes)	Unweighted UniFrac	0.028
		Jaccard Index	0.028			

5.3.3. Differential Abundance

To identify specific organisms associated with the participant groups differential abundance methods were employed, these were performed on the raw un-normalised datasets.

5.3.3.1. Effects of filtering

The datasets did require some further filtering to aid the differential abundance methods and to ensure that only frequently occurring/core organisms were included along with those that had been identified to an appropriate taxonomic rank. When filtering those taxa that occurred less than 10 times in a sample a reduction

of approximately 10 taxa compared to the raw was seen, but this did not change the identified taxa when ran through the Deseq method (Table 5.7). Increasing the filtering to 50 reduced the total taxa by approximately half along with a drastic reduction in identified taxa to the OA in the arm samples and YA in the face samples; interestingly the identified taxa increased in the YA arm samples. A further near half reduction in total taxa occurred from the 50 to 100 filtering steps, this also saw a change in the identified taxa with a reduction to all bar the arm OA who increase. Finally, the 200 filter saw a similar half reduction compared to the 100 filter, in the arm samples only a minor reduction in identified taxa occurs whereas in the face samples larger reduction happened to the OA group but a doubling has occurred in the YA group. After this comparison the filter removing taxa that were present less than 100 times in in the samples was used for the next stages of analysis as a suitable compromise due to lacking an appropriate statistically unbiased method to aid the comparison.

Table 5.6 Demonstrating the effects of various per sample filtering thresholds in both the arm and face samples. ***Detailing the total number of taxa present and number of taxa identified within each age group using the Deseq method.***

Filter	Arm			Face		
	Total taxa	Number of identified Taxa		Total Taxa	Number of identified Taxa	
		OA	YA		OA	YA
Raw	7895	19	32	9464	55	37
10	7871	19	32	9442	55	37
50	4149	4	43	4052	45	11
100	2427	16	36	2489	36	8
200	1277	15	34	1392	28	19

5.3.3.2. Age Groups

After the raw arm dataset was filtered to remove per sample taxa that occurred less than 100 times and subsequently analysed using five different differential abundance methods, 12 individual taxa were identified (Table 5.8), only *Aerococcus viridans* and *Enterococcus cecorum* were identified by all five methods. Two of the identified taxa were not successfully classified fully to the species level, these are a *Gemella spp.* and *Staphylococcus spp.* Additionally of the 12 taxa eight were identified to be abundant within the YA group, while only 4 were within the OA group.

When these 12 taxa were plotted using their relative abundance within each sample a large variation was seen in terms of proportion of abundance (Fig. 5.33). *Acinetobacter parvus, Facklamia hominis,* and *Gemella spp*. were all seen to only be present within the OA group and only at very low levels, most OA samples they are present in less than 1% of. Meanwhile the *Aerococcus viridans, Bifidobacterium spp., Sphingomonas colocasiae,* and *Sphingomonas desiccabilis* were only present within the YA group, again these are mostly present in less than 1% of the samples. Comparatively *Corynebacterium cystitidis, Enterococcus cecorum,* and *Mobilicoccus*

pelagius were shown to be present in both groups but in drastically different proportions and then only within some samples. For example, *C. cystitidis* is present in 40% of one OA sample and 25% of another, but then only within less than 5% of all other samples. Finally, *Cutibacterium acnes,* and *Staphylococcus spp.* are present in large abundances within both groups, but the OA tended to have a lower abundance of them compared to the YA group.

Table 5.7 Taxa identified as differentially abundant in the arm dataset. ***Only those species identified in more than three of the five abundance tests carried out on the dataset, when grouped based on age group.***

Species	Abundant Group	Methods Identified by				
		Ancombc	Deseq	Edger	Lefse	Limma-Voom
Acinetobacter parvus	OA		✓	✓	✓	✓
Aerococcus viridans	YA	✓	✓	✓	✓	✓
Bifidobacterium spp.	YA		✓	✓	✓	✓
Corynebacterium cystitidis	OA	✓	✓		✓	✓
Cutibacterium acnes	YA	✓	✓		✓	✓
Enterococcus cecorum	YA	✓	✓	✓	✓	✓
Facklamia hominis	OA		✓	✓		✓
Gemella spp.	OA		✓	✓	✓	✓
Mobilicoccus pelagius	YA	✓	✓	✓	✓	
Sphingomonas colocasiae	YA		✓	✓	✓	✓
Sphingomonas desiccabilis	YA		✓	✓	✓	✓
Staphylococcus spp.	YA	✓	✓			✓

After filtering the raw face dataset to remove per sample taxa that occurred less than 100 times and subsequently analysed using five different differential abundance methods, 11 individual taxa were identified (Table 5.9), only *Acinetobacter spp.* was identified by all five methods, this was one of the two taxa identified that were not successfully classified fully to the species level. The other was a *Citrobacter spp.* only one of the 11 taxa eight were identified to be abundant within the YA group, while the remaining 11 were within the OA group.

When these 11 taxa were plotted using their relative abundance within each sample a large variation was seen in terms of proportion of abundance (Fig.5.34). *Acinetobacter parvus, Campylobacter lanienae*, *Citrobacter spp.,* and *Fusobacterium animalis* were all present only within the OA sample group but not within the YA group, these are present in less than 1% of the majority of samples. The *Acinetobacter spp., Corynebacterium cystitidis, Pseudomonas versuta, Streptococcus gordonii,* and *Streptococcus oralis* while all present within both sample groups have a relative abundance of less than 1% in most samples, but in some like the *Pseudomonas versuta* is present in approximately 50% of one OA participant. Again *Cutibacterium acnes* has the largest abundance within both age groups, however, here the mean abundance within the OA group is 25%

and the YA 30%, these are both higher than in the arm dataset; further none of the YA samples do not contain *C. acnes*. These species are abundant when viewed through the age grouping, but as previously established with the alpha and beta diversity results differences in these taxa were also observed in different metadata groupings.

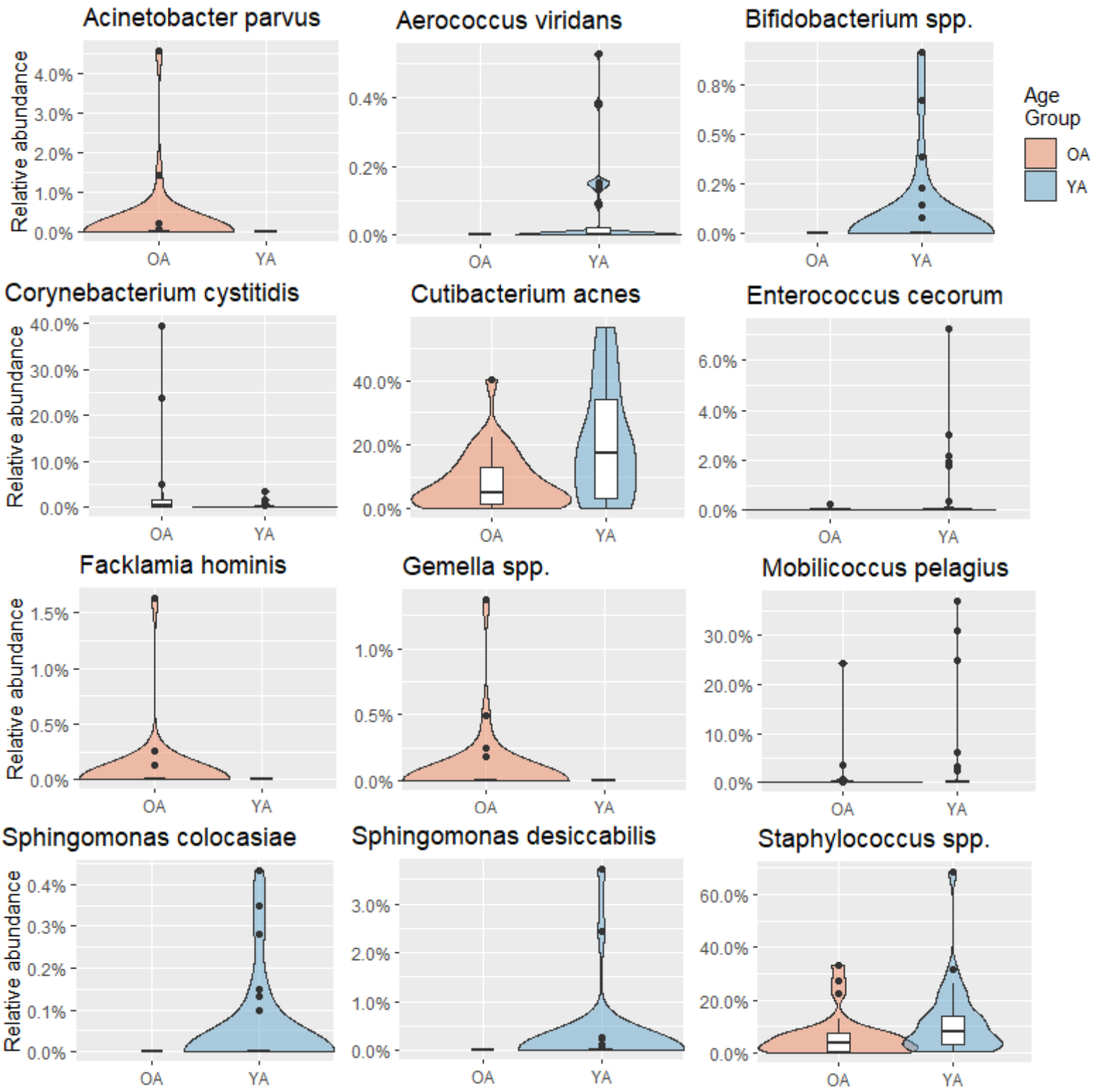


Figure 5.33 Relative abundance plots of the 12 taxa in the arm dataset. *Identified through differential abundance as being abundant in either the OA of YA age group.*

Table 5.8 Taxa identified as differentially abundant in the face dataset. ***Only those species identified in more than three of the five abundance tests carried out on the dataset, when grouped based on age group.***

Species	Abundant Group	Methods Identified by				
		Ancombc	Deseq	Edger	Lefse	Limma-Voom
Acinetobacter parvus	OA		✓	✓		✓
Acinetobacter spp.	OA	✓	✓	✓	✓	✓
Campylobacter lanienae	OA		✓	✓	✓	✓
Citrobacter spp.	OA		✓	✓	✓	✓
Corynebacterium cystitidis	OA	✓			✓	✓
Cutibacterium acnes	YA	✓	✓		✓	
Fusobacterium animalis	OA		✓	✓	✓	✓
Porphyromonas gingivalis	OA		✓	✓		✓
Pseudomonas versuta	OA		✓	✓		✓
Streptococcus gordonii	OA	✓			✓	✓
Streptococcus oralis	OA		✓		✓	✓

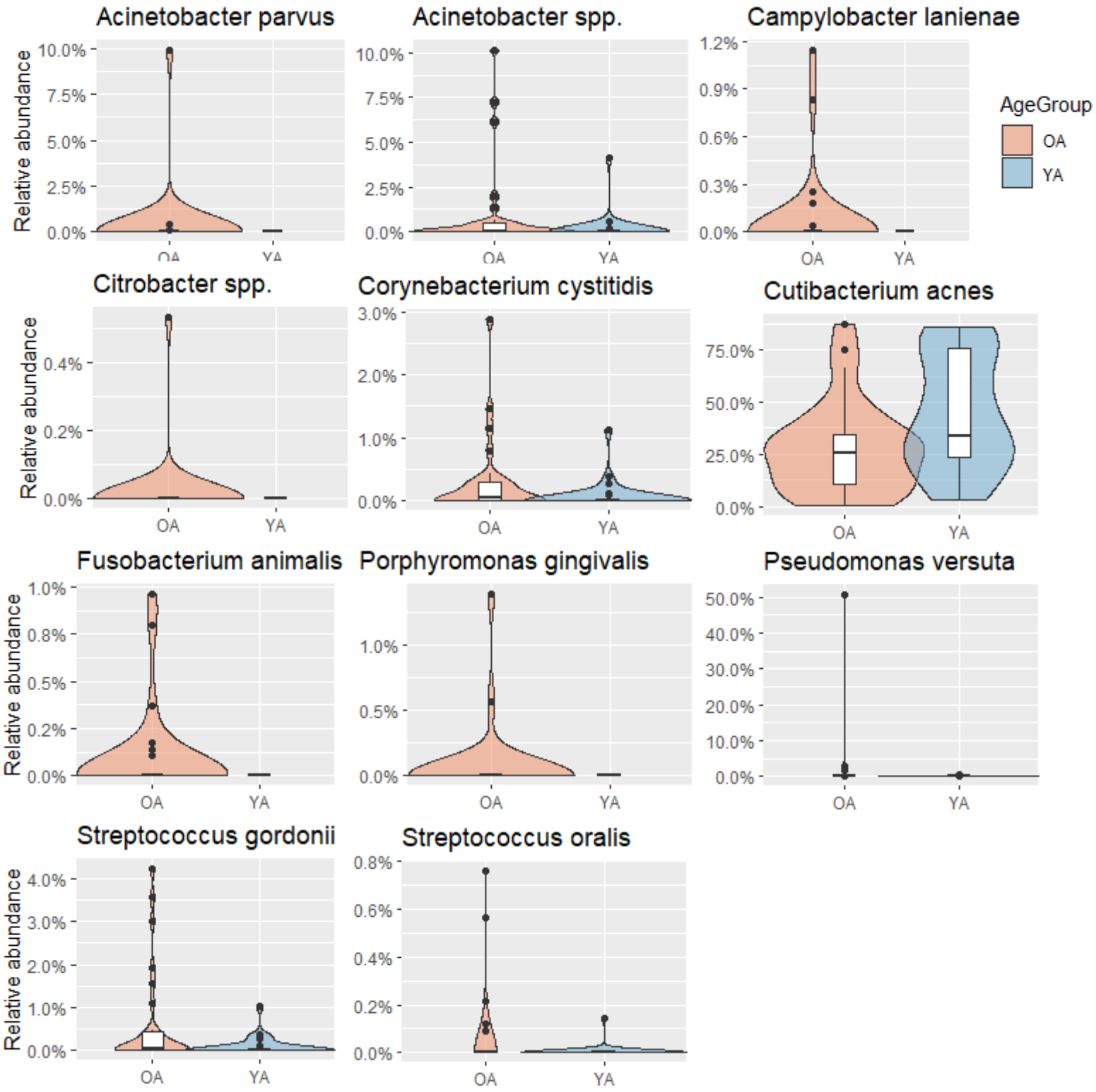


Figure 5.34 Relative abundance plots of the 11 taxa in the face dataset. *Identified through differential abundance as being abundant in either the OA of YA age group.*

5.3.3.3. Gender

The differential abundance methods were also applied to the age groups broken down by gender, this was used to follow the taxa identified in the previous age group differential abundance analysis. In the arm dataset the Ancombc method did not identify any taxa as differentially abundant, and some methods disagreed in which group was abundant for the taxa, while the Limma-Voom method identified the most taxa in this grouping breakdown (Table 5.10). Here most of the abundant groups are the OA and YA females, with only the OA and YA male identified in *Acinetobacter parvus* and *Enterococcus cecorum* respectively. Finally, the *Staphylococcus spp.* was not identified by any method.

Only the eight taxa that were identified by multiple methods were plotted as relative abundance broken down by age group and gender (Fig.5.35). Here *F. hominis* was seen to only be present within the OA-Female group and not in any other. *A. parvus, Bifidobacterium spp.,* and *S. deciccabilis* are seen to be split between the genders in the age groups they were associated with. In the *C. cystitidis* the OA-Female had the largest abundance while the other groups have similarly low abundance, this is inverted somewhat in the *E. cecorum* where the YA-Females have the largest abundance while the other groups have lower abundances. The *C. acnes* had similar abundances in both male groups, but appeared different in the females, with the OA-Female having lower abundance than the YA-Female. This was again slightly inverted in the *M. pelagius* where the OA-Males have low abundance compared to the YA-Males, while the female groups have similar abundances.

Table 5.9 Taxa identified as differentially abundant in the arm dataset grouped based on age group and gender. ***Only those species identified in more than three of the five abundance tests carried out on the arm dataset, when grouped based on age group.***

Species	Abundant Group	Methods Identified by				
		Ancombc	Deseq	Edger	Lefse	Limma-Voom
Acinetobacter parvus	OA-Male			✓		
	OA-Female					✓
Aerococcus viridans	YA-Female					✓
Bifidobacterium spp.	YA-Female			✓		✓
Corynebacterium cystitidis	OA-Female		✓		✓	✓
Cutibacterium acnes	YA-Female		✓		✓	
Enterococcus cecorum	YA-Female			✓		
	YA-Male					✓
Facklamia hominis	OA-Female			✓		✓
Gemella spp.	OA-Female					✓
Mobilicoccus pelagius	YA-Female		✓	✓		✓
Sphingomonas colocasiae	YA-Female					✓
Sphingomonas desiccabilis	YA-Female			✓		✓
Staphylococcus spp.	NA					

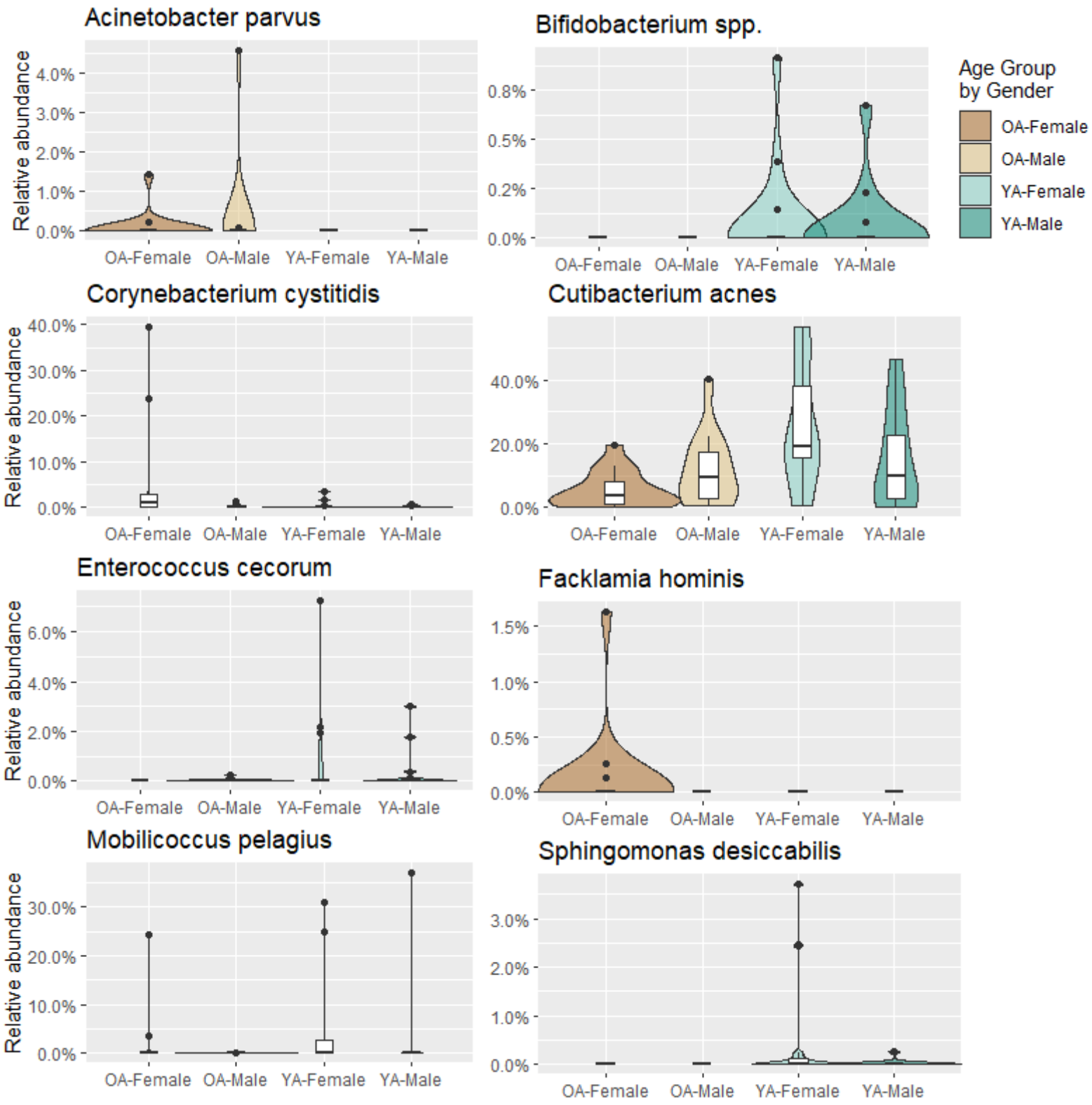


Figure 5.35 Relative abundance plots of the 8 taxa in the arm dataset identified through differential abundance grouped by the age groups by gender. *That were previously identified as abundant in the age groups alone.*

This gender grouping also effected the face results, however here none of the taxa identified in the age grouping methods were identified with either the Ancombc or Deseq methods, additionally two previously identified taxa were not identified through these groupings (Table 5.11). Most identified taxa were abundant within the OA-Female group with only two in the OA-Male group, none were identified within the YA gender groups.

Only one taxa was identified by multiple methods when the face dataset was grouped by age group and gender, this was *A. parvus* which was shown to be in higher abundance within the OA-Males compared to less than 1% in the OA-Females (Fig.5.36). Comparing both the arm and face taxa identified by age groups then traced through the gender group shows that despite similar beta diversity differences the resulting taxa differences this present are likely similar in the arm but different in the face microbiomes.

Table 5.10 Taxa identified as differentially abundant in the face dataset grouped based on age group and gender. ***Only those species identified in more than three of the five abundance tests carried out on the face dataset, when grouped based on age group.***

Species	Abundant Group	Methods Identified by				
		Ancombc	Deseq	Edger	Lefse	Limma-Voom
Acinetobacter parvus	OA-Male			✓		✓
Acinetobacter spp.	OA-Female					✓
Campylobacter lanienae	OA-Male					✓
Citrobacter spp.	OA-Female					✓
Corynebacterium cystitidis	NA					
Cutibacterium acnes	NA					
Fusobacterium animalis	OA-Female			✓		
Porphyromonas gingivalis	OA-Female			✓		
Pseudomonas versuta	OA-Female					✓
Streptococcus gordonii	OA-Female				✓	
Streptococcus oralis	OA-Female				✓	

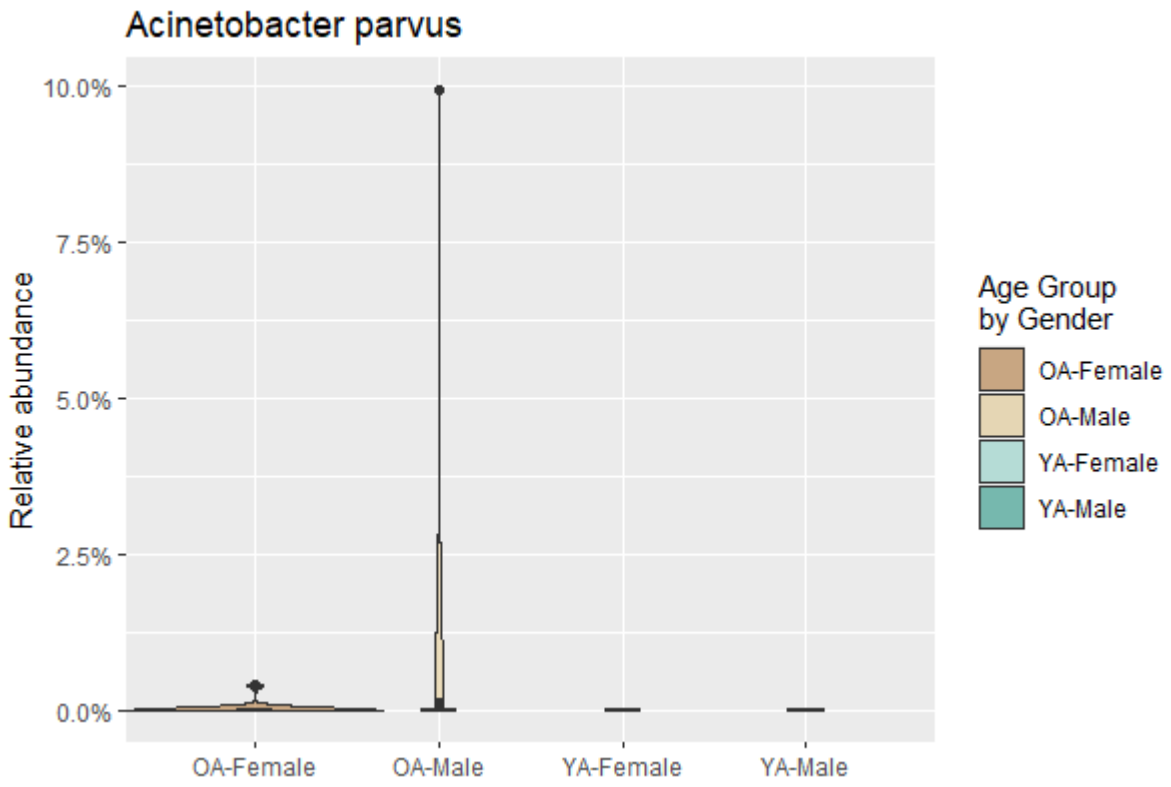


Figure 5.36 Relative abundance plots of the taxa in the face dataset identified through differential abundance grouped by the age groups by gender. *That were previously identified as abundant in the age groups alone.*

5.3.3.4. Frailty index

Finally, the differential abundance methods were also applied to the age groups broken down by FI score. In the arm dataset the Ancombc method did not identify any taxa as differentially abundant, while the Limma-Voom method identified the most based on the previously found taxa additionally four taxa were not to be abundant by any method used (Table 5.12). Most of the identified taxa are abundant within the YA-0 group, while the OA-3 group was the only OA group with abundant taxa.

When the four taxa identified by multiple methods were plotted as relative abundance, they show some variation in abundance across the FI scores (Fig.5.37). Here *A. parvus* was seen in high abundance within the OA-2 and OA-3 groups compared to the OA-0, while *S. desiccabilis* was in high abundance within the YA-0 group and low in the other YA groups. *E. cecorum* is shown to only be present within the OA-0 group, while it was present across all the YA groups it is predominant in the YA-0. Finally, *M. pelagius* is seen in high abundance in the OA-1 and YA-0 groups and relatively lowly in the other groups. Across all four of these identified taxa the high points of relative abundance appeared to be outliers within each group, this warranted further study of these datasets which was beyond the scope of this study.

Table 5.11 Taxa identified as differentially abundant in the arm dataset grouped based on age group and FI. *Only those species identified in more than three of the five abundance tests carried out on the arm dataset, when grouped based on age group.*

Species	Abundant Group	Methods Identified by				
		Ancombc	Deseq	Edger	Lefse	Limma-Voom
Acinetobacter parvus	OA-3			✓		✓
Aerococcus viridans	YA-1					✓
Bifidobacterium spp.	NA					
Corynebacterium cystitidis	NA					
Cutibacterium acnes	YA-0		✓			
Enterococcus cecorum	YA-0			✓		✓
Facklamia hominis	OA-3					✓
Gemella spp.	NA					
Mobilicoccus pelagius	YA-0			✓		✓
Sphingomonas colocasiae	YA-0					✓
Sphingomonas desiccabilis	YA-0			✓		✓
Staphylococcus spp.	NA					

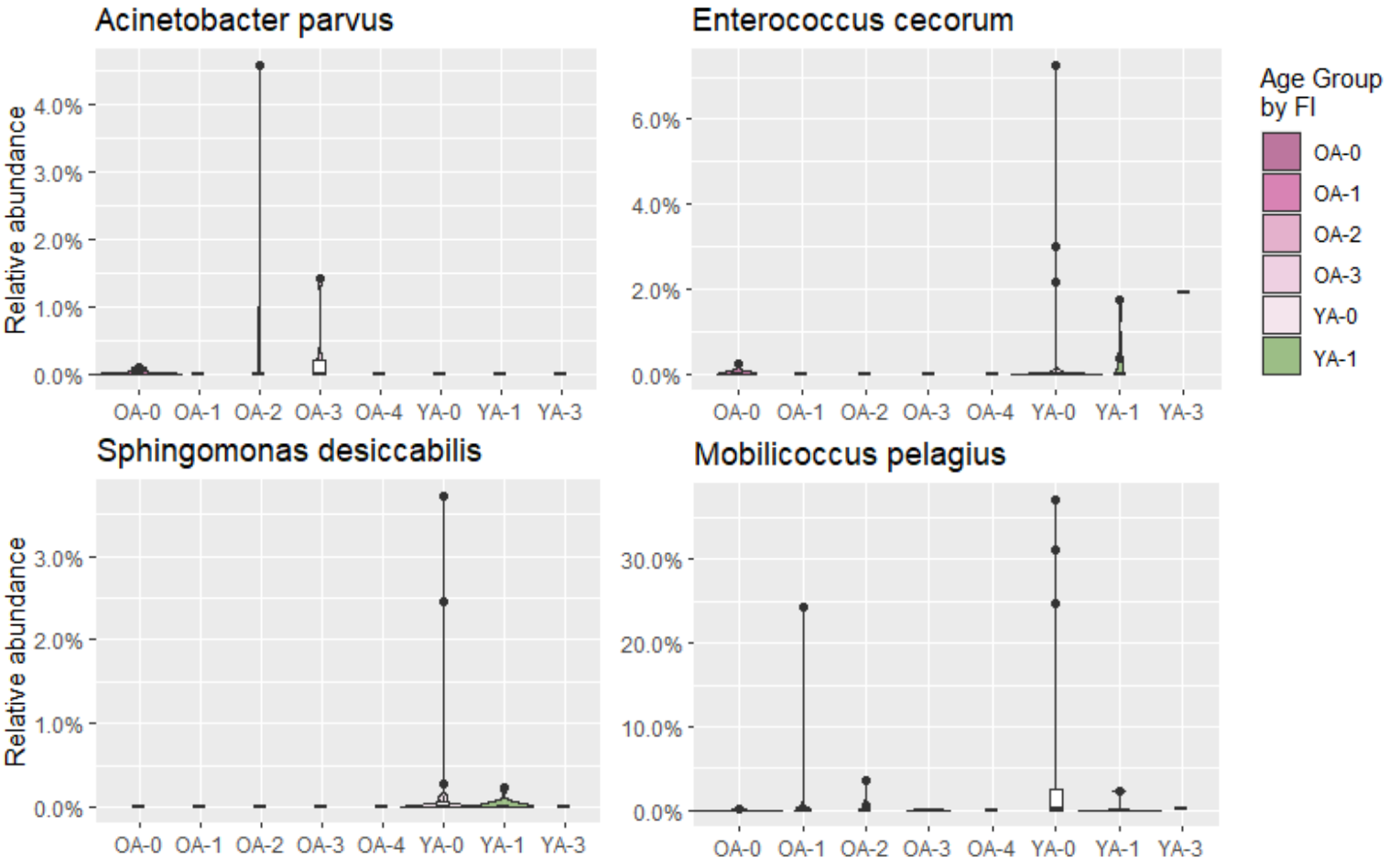


Figure 5.38 Relative abundance plots of the 4 taxa in the arm dataset. identified through differential abundance as being abundant in the age groups by FI score. *That were previously identified as abundant in the age groups alone.*

When this FI score grouping was applied to the face dataset only the Edger and Limma-Voom methods identified any of the taxa found when grouped by age group alone, additionally four of the taxa were not identified by any method (Table 5.13). All taxa were abundant only in the OA age group, with OA-1 having the most attributed to it. Only one taxa was identified through multiple methods, this was *F. animalis* which when plotted using relative abundance is shown to be the most abundant in the OA-1 group (Fig. 5.38). These conclude the differential abundance investigations into the skin microbiomes, further investigations into the microbiomes was carried out using co-occurrence network analysis.

Table 5.12 Taxa identified as differentially abundant in the face dataset grouped based on age group and FI. ***Only those species identified in more than three of the five abundance tests carried out on the face dataset, when grouped based on age group.***

Species	Abundant Group	Methods Identified by				
		Ancombc	Deseq	Edger	Lefse	Limma-Voom
Acinetobacter parvus	OA-2			✓		
Acinetobacter spp.	OA-1					✓
Campylobacter lanienae	OA-1					✓
Citrobacter spp.	NA					
Corynebacterium cystitidis	NA					
Cutibacterium acnes	NA					
Fusobacterium animalis	OA-1			✓		✓
Porphyromonas gingivalis	OA-3					✓
Pseudomonas versuta	OA-2			✓		
Streptococcus gordonii	NA					
Streptococcus oralis	NA					

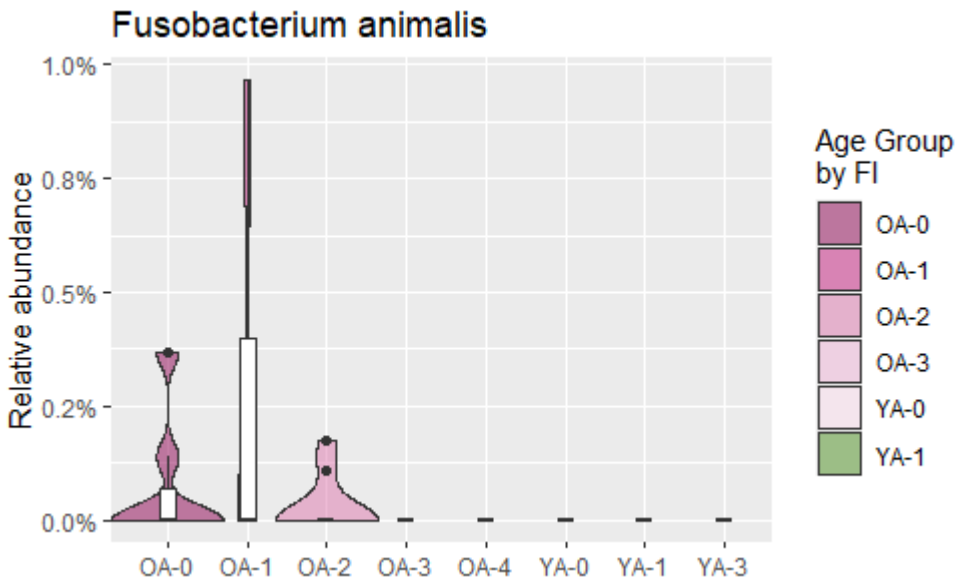


Figure 5.39 Relative abundance plots of the taxa in the face dataset. identified through differential abundance as being abundant in the age groups by FI score. *That were previously identified as abundant in the age groups alone.*

5.3.4. Network analysis

Co-occurrence networks were constructed to infer and analyse potential ecological relationships within the age groups microbiomes. The datasets were filtered to remove low abundance taxa and ensure only shared core taxa were considered for this computationally intensive analysis, additionally the networks were only compared at the age group level. After filtering a total of 156 taxa remained in the arm dataset and 214 for the face samples.

The arm networks showed 151 and 144 connected nodes for the OA and YA groups respectively, seven of the taxa identified in the differential abundance analysis are found in the OA network, while five are in the YA network (Fig.5.39). When viewed in more detail the largest connected component (LCC) in the OA group is 148 taxa and 140 in the YA, however the YA networks and LCC had a higher positive estimated correlation of around 74% whereas the OA have around 70% (Table 5.13).

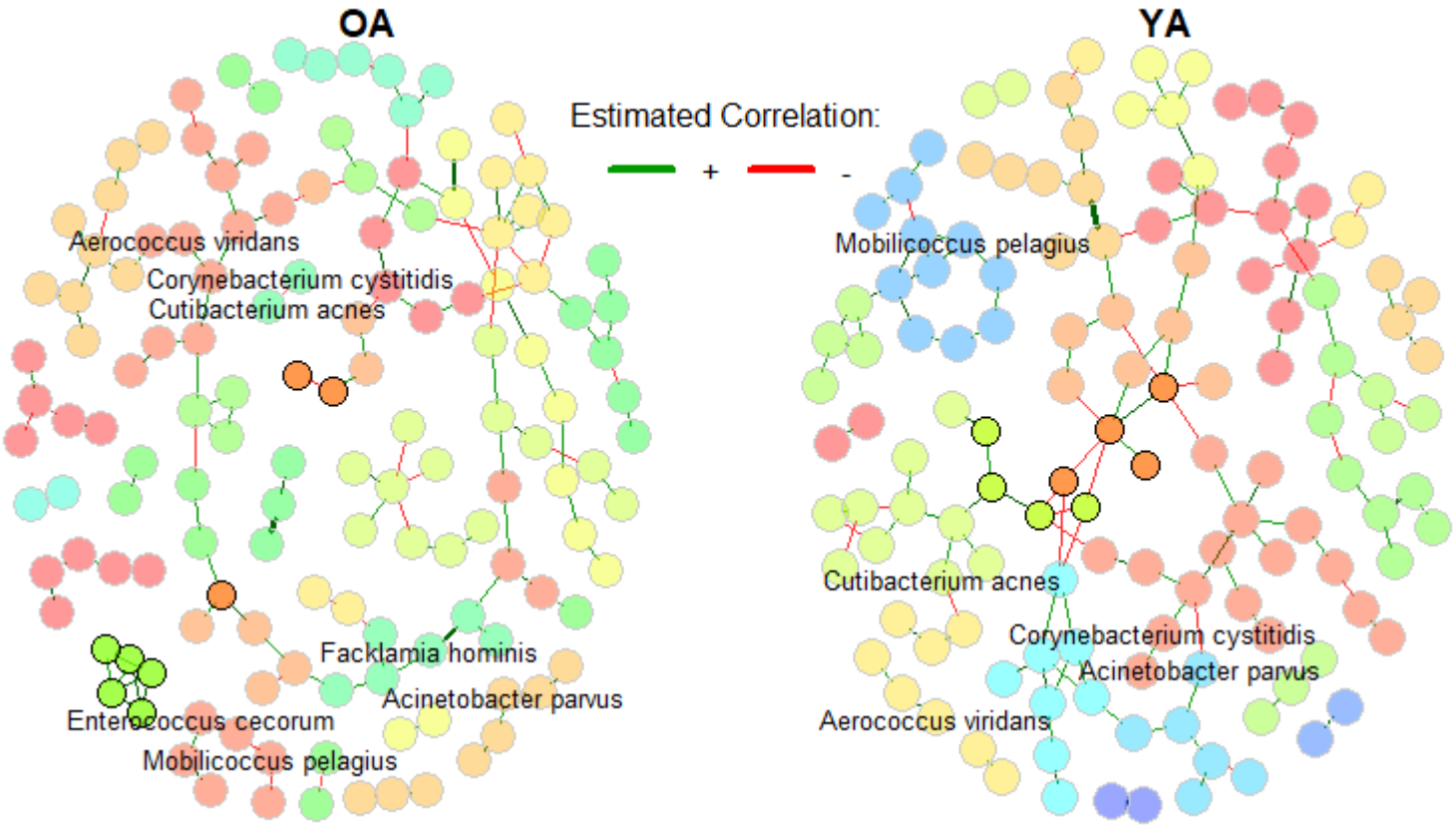


Figure 5.40 Co-Occurrence networks for the arm dataset comparing the OA and YA age groups. *The networks were constructed using SPIEC-EASI statistical method in the Netcomi R package, each node represents a unique taxa (either at the species level or there originally assigned level where speciation unavailable), within the dataset, nodes are coloured based on clustering, unconnected nodes have been removed. Edges are representative of association with nodes, split into estimated positive and negative association by the colours green and red. Nodes representing taxa identified as differentially abundant are labelled where present.*

Additionally, the statistically identified hub taxa were not shared with any of the taxa identified through differential abundance, both networks however, share a *Glutamicibacter* but no other taxa. These *Glutamicibacter* are the most common hub taxa in the YA network, comparatively the OA have *Staphylococcus* as the most common taxa. A detailed statistical comparison of the networks reveals no significant difference in the overall descriptive statistics, but significant differences are observed when a Jaccard index is used on the centrality measures degree, betweenness, eigenvector, and hub taxa; with p values of 0.002, 0.0006, 0.002, and 0.02 respectively. This showed significant dissimilarity rather than significant similarity between the age group arm networks.

Table 5.13 A summary of the descriptive information and statistical summary for the arm co-occurrence network. ***P-values below 0.05 are highlighted in green.***

Descriptive information		OA			YA		
Components	Size	148	3	1	140	2	1
	Number	1	1	5	1	2	12
Network overview	Number of components	7.00000			15.00000		
		Whole network	LCC		Whole network	LCC	
	Clustering coefficient	0.09095	0.09170		0.09022	0.09022	
	Modularity	0.73455	0.73569		0.72493	0.72121	
	Positive edge percentage	70.85427	70.55838		74.11168	73.84615	
	Edge density	0.01646	0.01811		0.01629	0.02004	
	Natural connectivity	0.00747	0.00792		0.00747	0.00843	
	Relative LCC size	NA	0.94872		NA	0.89744	
	Vertex connectivity		1.00000			1.00000	
	Edge connectivity		1.00000			1.00000	
	Average dissimilarity		0.99436			0.99373	
	Average path length		4.62786			4.26172	
Hub taxa		*Actinobacteria 6314*			*Bacillales_Incertae Sedis XI*		
		Corynebacterium 7566			*Fusobacterium periodonticum*		
		Glutamicibacter 6370			*Glutamicibacter 6348*		
		Micrococcales 6315			*Glutamicibacter 6353*		
		Paracoccus solventivorans			*Glutamicibacter 6370*		
		Staphylococcus 4753			*Glutamicibacter 6905*		
		Staphylococcus 4848			*Microbacteriaceae 6573*		
		Staphylococcus hominis			*Sphingobacterium lactis*		

Statistical summary		Absolute difference		p-value	
Network overview		Whole network	LCC	Whole network	LCC
	Clustering coefficient	0.001	0.001	0.991009	0.880120
	Modularity	0.010	0.014	0.850150	0.941059
	Positive edge percentage	3.257	3.288	0.711289	0.807193
	Edge density	0	0.002	0.951049	0.980020
	Natural connectivity	0	0.001	1.000000	0.993007
	Relative LCC size	NA	0.051	NA	0.713287
	Vertex connectivity		0		1.000000
	Edge connectivity		0		1.000000
	Average dissimilarity		0.001		0.989011
	Average path length		0.366		0.796204

Jaccard index	Measure	Jacc	P(<=Jacc)	P(>=Jacc)
	degree	0.122	0.001888	0.999524
	betweenness centr.	0.149	0.000573	0.999813
	closeness centr.	0.238	0.067940	0.962095
	eigenvec. centr.	0.164	0.001572	0.999427
	hub taxa	0.067	0.019411	0.997716

The face networks showed 155 and 211 connected nodes for the OA and YA groups respectively, four of the taxa identified in the differential abundance analysis were found in the OA network, while six were in the YA network (Fig.5.40). When viewed in more detail the largest connected component (LCC) in the OA group is 98 taxa and 211 in the YA, however the OA networks and LCC have a higher positive estimated correlation of around 74-78% whereas the YA had around 71% (Table 5.14). Only one of the statistically identified hub taxa is shared with the taxa also identified through differential abundance this is the *S. gordonii* which is a hub in the YA network. A total of 11 taxa are hub nodes within these networks, this is 3 more than in the arm networks. The most abundant hub taxa in the OA group were *Micrococcales* whereas in the YA group it was *Acinetobacter*. A detailed statistical comparison of the networks reveals significant differences across many of the overall descriptive statistics across both the whole network and in LCC including natural connectivity, modularity, and clustering coefficients. Additionally significant differences are observed when a Jaccard index is used on the centrality measures degree, betweenness, closeness, eigenvector, and hub taxa; with p values of 0.008, 0.0004, 0.0005, 0.004, and 0.002; this shows significant dissimilarity rather than significant similarity between the face networks. Therefore, while the arm and face networks were visibly distinct from each other when comparing the networks between the age groups they were significantly dissimilar, this showed that there are community differences between both age groups in the skin microbiomes studied here.

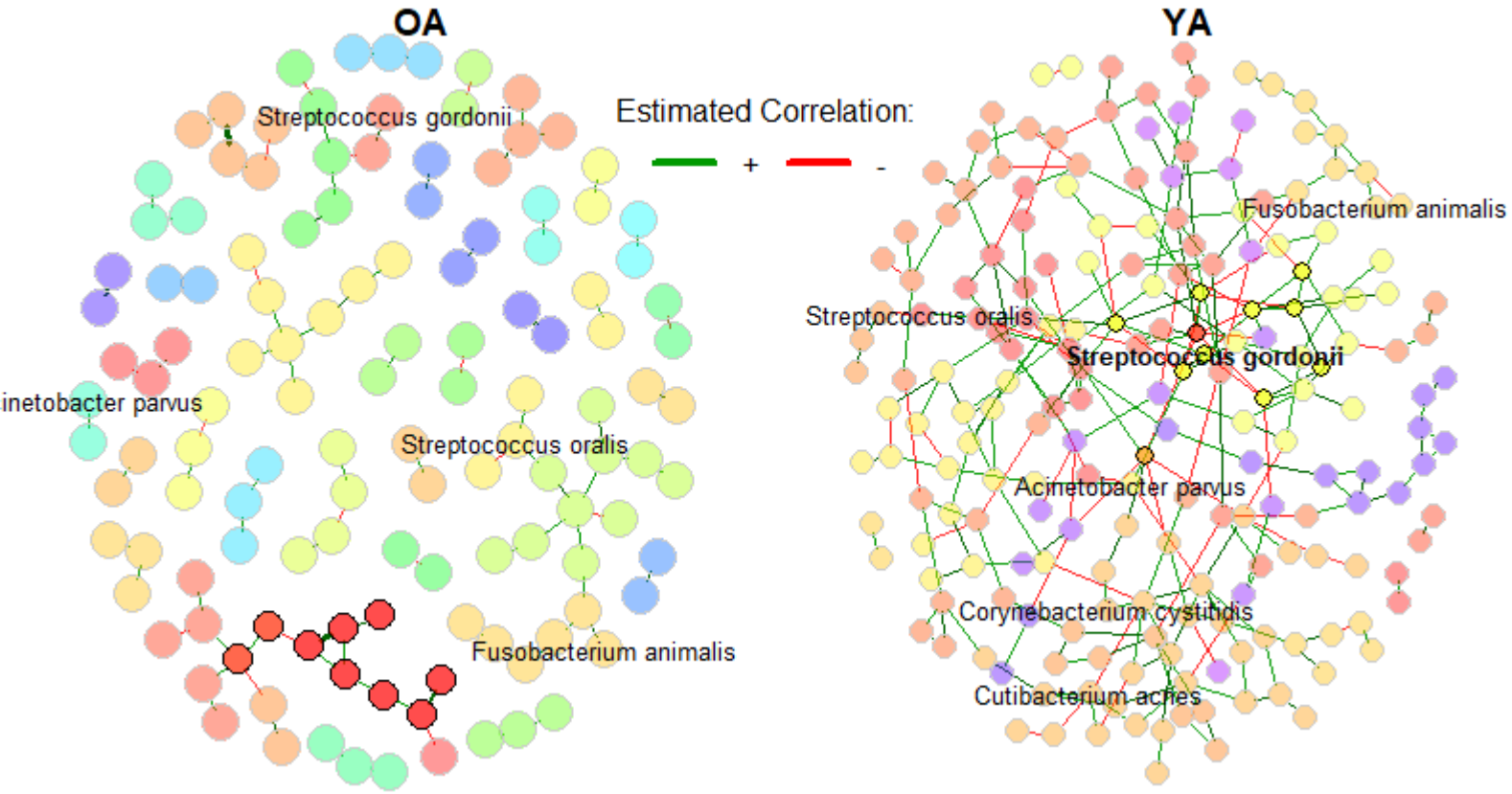


Figure 5.41 Co-Occurrence networks for the face dataset comparing the OA and YA age groups. *The networks were constructed using SPIEC-EASI statistical method in the Netcomi R package, each node represents a unique taxa (either at the species level or there originally assigned level where speciation unavailable), within the dataset, nodes are coloured based on clustering, unconnected nodes have been removed. Edges are representative of association with nodes, split into estimated positive and negative association by the colours green and red. Nodes representing taxa identified as differentially abundant are labelled where present.*

Table 5.14 A summary of the descriptive information and statistical summary for the face co-occurrence network. *P-values below 0.05 are highlighted in green.*

Descriptive information		OA						YA	
Components	Size	98	9	4	3	2	1	211	1
	Number	1	1	2	6	11	59	1	3
Network overview	Number of components	80						4	
		Whole network				LCC		Whole network	LCC
	Clustering coefficient	0.01225				0.01495		0.08639	0.08639
	Modularity	0.84680				0.77033		0.62333	0.62333
	Positive edge percentage	78.41727				74.50980		71.83463	71.83463
	Edge density	0.00610				0.02146		0.01698	0.01747
	Natural connectivity	0.00510				0.01176		0.00567	0.00576
	Relative LCC size	NA				0.45794		NA	0.98598
	Vertex connectivity					1.00000			1.00000

	Edge connectivity		1.00000		1.00000
	Average dissimilarity		0.99339		0.99455
	Average path length		4.92409		3.27189
Hub taxa		*Actinobacteria 7826*		*Acinetobacter 6929*	
		Capnocytophaga sputigena		*Acinetobacter 6931*	
		Haemophilus parainfluenzae		*Actinobacteria 8801*	
		Lautropia mirabilis		*Bacillales_Incertae Sedis XI*	
		Micrococcales 7793		*Glutamicibacter 7870*	
		Micrococcales 7807		*Gordonia 8849*	
		Micrococcales 7808		*Granulicatella adiacens*	
		Micrococcales 7828		*Haemophilus parainfluenzae*	
		Micrococcales 7830		*Pseudomonas 6574*	
		Micrococcales 7836		*Stenotrophomonas rhizophila*	
		Roseomonas mucosa		*Streptococcus gordonii*	

Statistical summary		Absolute difference		p-value	
Network overview		Whole network	LCC	Whole network	LCC
	Clustering coefficient	0.074	0.071	0.011988	0.017982
	Modularity	0.223	0.147	0.035964	0.033966
	Positive edge percentage	6.583	2.675	0.209790	0.599401
	Edge density	0.011	0.004	0.000999	0.069930
	Natural connectivity	0.001	0.006	0.000999	0.032967
	Relative LCC size	NA	0.528	NA	0.029970
	Vertex connectivity		0		1.000000
	Edge connectivity		0		1.000000
	Average dissimilarity		0.001		0.082917
	Average path length		1.652		0.079920

Jaccard index	Measure	Jacc	P(<=Jacc)	P(>=Jacc)
	degree	0.173	0.008079	0.996930
	betweenness centr.	0.165	0.000397	0.999853
	closeness centr.	0.174	0.000503	0.999800
	eigenvec. centr.	0.200	0.003845	0.998206
	hub taxa	0.048	0.002306	0.999800

5.3.5. Predictive metabolomics

The raw datasets were used in the Qiime2 Picrust2 plugin to create the predicted metabolomic pathway data. This was normalised through conversion into relative abundance data before broad analysis using beta diversity. In the arm datasets no significant differences were observed between the age groups using either Bray Curtis or Jaccard index, while the in the face datasets there was a significant difference with a p-value of 0.018 and 0.009 for these metrics respectively (Fig. 5.41 and Fig. 9.17). This sees some distinct clustering in the face sample ordination plots but a large level of overlap in the arm sample plots.

These data were also analysed using the same differential abundance methods as used previously, but without any filtering. In the arm data 35 pathways were identified as abundant in three or more of the various methods, however none were identified by all five methods (Table 9.4). Comparatively in the face data 162 pathways were identified as abundant in three or more of the various methods, but five were identified by all five methods (Table 9.5). Additionally in both datasets different methods found some the same pathways abundant but in different groups.

When these five abundant pathways were plotted using relative abundance it demonstrated that they constitute a relatively small proportion of the overall number of pathways as they all are present in all samples at levels lower than 0.5%, these were also only abundant within the OA age group (Fig.5.42). The 4-methylcatechol degradation, L-arginine degradation II, nitrate reduction I, and toluene degradation III pathways all had single samples with the OA group that are higher than the majority of other OA samples compared to the YA who are more compact. In comparison the phenylacetate degradation I pathway had several samples in the OA group spread throughout the abundance and similarly with the YA, however the OA still have a higher abundance. This demonstrated a shift in the microbial metabolome in the face microbiome, but not the arm associated with the age groups. This concluded the skin characterisation methods and analysis.

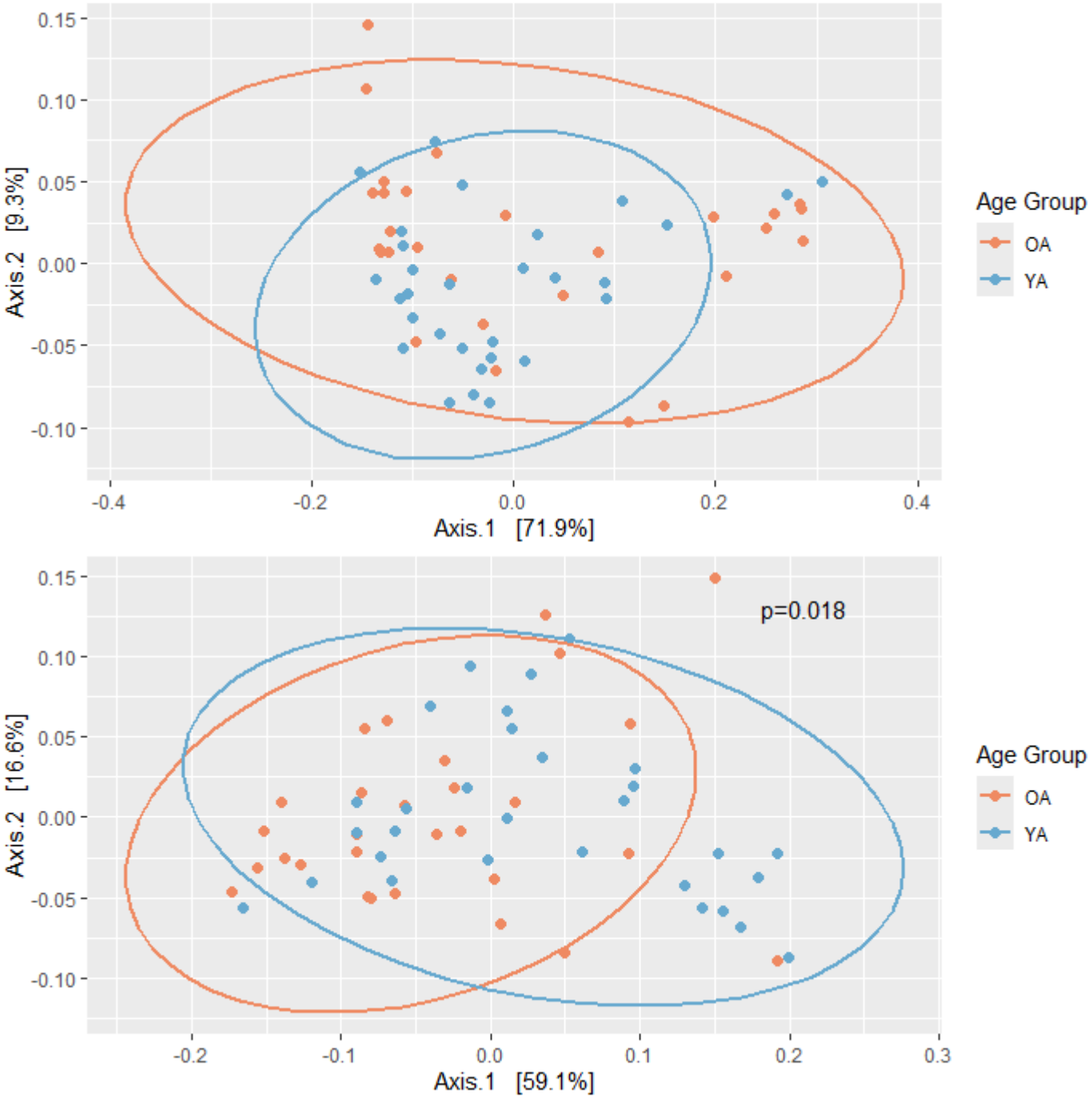


Figure 5.42 Bray Curtis dissimilarity measure of beta diversity as an ordination plot of the predicted metabolomic pathways. *Where a significant difference was observed. A. In the arm dataset. B. In the face dataset, where a significant difference was observed (p=0.018).*

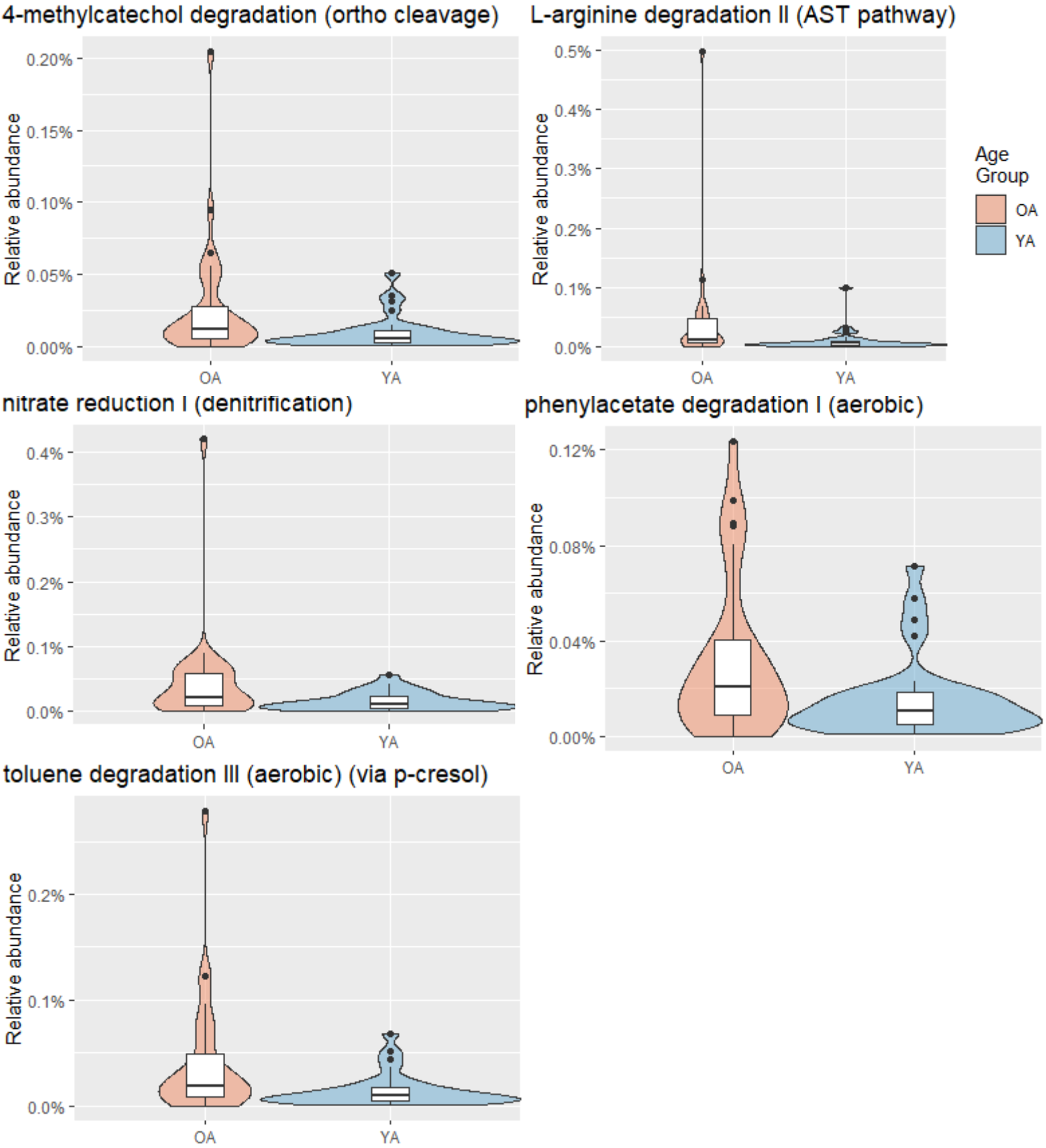


Figure 5.43 Relative abundance plots of metabolic pathways identified as abundant in the face dataset.

5.4. Discussion

The focus of this chapter has been the characterisation of the skin microbiome using the datasets previously finalised in chapter 4 with a focus on the age groups the samples were collected from. As this is one of the main themes of this thesis it required a considered and comprehensive analysis covering microbial ecology, mathematical modelling, and post taxonomic toolkits to enable detailed characterisation and comparison. Additionally previous analysis of the participant metadata conducted in chapter 3 identified other groups within the participant panel using some of these to breakdown and compare the data to better understand any differences or similarities seen between the age groups. While no direct statistical comparison between the arm and face data is made here, combining them in a single chapter does allow for them to be compared through side-by-side observation; still enabling a complete characterisation of the skin samples.

The broad overview of the top 10 species (Fig. 5.2) identified within the sample sets shows expected skin organisms in varying proportions across the samples, although *P. versuta* for example is not associated with skin colonisation, instead it is associated with soil (See-Too *et al.*, 2017). In the OA arm samples *Staphylococcus spp.* can been seen to be dominant while this is less in the YA samples, similarly *C. acnes* have a lower presence in the OA whereas it is higher in the YA. The recovery of more unique sequences and taxa in the face samples compared to the skin is expected as the face is more exposed to environmental deposits, similarly some of these organisms are oral in nature which can easily be acquired through speaking and other normal behaviours. Interestingly the samples that were removed during the normalisation step show similar top species content and distribution to the other samples within their age groups, the exception is OA-20 who's face sample consists of only two species, this suggests that while there removal may have initially been interpretable as a loss of data this was not the case as there similarity to the other samples ensured that the data was still broadly similar. While this broad overview analysis is interesting it is limited without more in-depth statistical comparisons.

Using alpha diversity measures on the arm samples showed no significant differences when clustered by age group, but when the ages were stretched out and viewed using Shannon diversity the YA are clustered together enough for a significantly confident regression line to be drawn whereas the regression line is not significantly confident in the OA (Fig. 5.4). This maybe a sign of an actual difference between the age groups or individual ages if more participants were present within the study or more potentially transient/contaminant sequences were filtered from the dataset. Comparatively with the face samples a significant difference is seen but only using the Pielou evenness metric, which shows that the OA face microbiomes tend to have more similar sequence abundances than the YA who are more diverse (Fig. 5.6). This however may not be linked to age but may be due to the increased diversity of the YA population and the differences in face skin care regime therein. This level of OA diversity is supported when the ages were stretched out and viewed using Shannon diversity where a nearly linear regression line can be seen, however this support is limited as both age groups regression

lines are not significantly confident (Fig. 5.7). Again, there is the possibility that significant differences may become apparent between the age groups if more participants were present within the study or more potentially transient/contaminant sequences were filtered from the dataset, this could be more likely in these face samples as there are distinct differences in mean alpha diversity on the graphs. However, a study comparing multiple skin studies that focus on age notes inconsistent results when performing this type of data analysis (Myers *et al.*, 2024), although this may be due to various differences between the studies the inability to accurately cross compare makes conclusions difficult. Beyond this when taxonomy is considered using beta diversity metrics significant differences are seen across multiple metrics in both samples. This shows that overall, the age groups are distinct from each other while being internally similar, this is supported when looking back at the top 10 organisms where a distinct shift in the populations can be seen; mostly this is a loss in *C. acnes* in the species graphs. Together this shows that differences in the skin microbiome do occur in ageing and that these changes are more at the level of changes of specific species rather than in the broader diversity metrics, however the changes are different depending upon the area of skin being analysed.

Breaking down the age groups into their gendered sub populations is a logical step as these are both the largest subgroups within the population pool but also gender can play a large role in the ageing process both in terms of biology and socioeconomically (House, Kessler and Herzog, 1990; Kane and Howlett, 2021). At the alpha diversity level for both the arm and face data no significant differences were seen between any of the age group genders, however similarities were seen in the male populations where similar ranges of diversities were seen, while the females one age group tended to have a more compact range than the other. It maybe that an actual significant difference is seen in these female groups with a larger population of participants or with a different refinement to the dataset. Using beta diversity metrics, a consistent significant difference was observed between the female groups in both the arm and face samples (Table 5.1 and Table 5.2), but inconsistent significant differences were observed between other OA gender and YA gender groups. This suggests that the main driver seen in the age groups significant differences is in these female populations as they are the larger groups, additionally this difference may be due to the YA female population being pre-menopausal and the OA females being post-menopausal, as the menopause is a change in hormonal output which has an established effect on the microbiome, this change has begun to be established particularly in the gut and vaginal microbiome (Peters *et al.*, 2022; Park, Cho and Oh, 2023). In the face samples the additional significant differences in beta diversity are only seen in the Unweighted UniFrac metric, suggesting that there are taxonomic differences between the OA-Female and both male groups. Meanwhile the arm samples display a significant difference between the OA-Male and YA-Female in Bray Curtis dissimilarity and Unweighted UniFrac suggesting not only taxonomically based differences but also some alterations at the abundance level, which could be due to various lifestyle or hormonal differences between these two groups. This expands upon the previously established changes in the skin

microbiome in ageing by showing the potential importance gender can play in the ageing microbiome, which should be an area for further study to enable better treatment and monitoring systems to be created.

Participants were scored on a frailty index to provide a quantitative measure of their ageing, this showed a significant difference between the two populations where the YA mostly score zero or one while the OA population are spread over the scale (Fig 3.5), this index can also be used to review the diversities of the skin microbiome samples. In both the arm and face samples no significant differences were observed in alpha diversity when clustered by FI score and the only difference in beta diversity was seen in the arm Unweighted UniFrac between OA-2 and YA-0, this beta diversity difference is in the largest two groups of the FI (Table 5.3). This does mean that the FI as applied in this study is not related to the diversity of the skin microbiome, this can be demonstrated in the alpha diversity graphs where there is a strong overlap in diversity metric across the FI scores for each age group. However, this study did not use frailty as a recruitment mechanism, this can be seen in how the OA group mostly score a two on an index capable of scoring 10 or more, similarly there are limitations in taking the complexities of frailty into a neutral scoring system. If this study was conducted on a different population where frailty was a metric used in recruitment or on even older participants, then it may reveal some link between frailty and the skin microbiome, as already established in this USA based study (Larson *et al.*, 2022).

In this chapter the two sociodemographic measures used to review the diversity analysis were the super area output code and workplace zone code, these showed significant differences between the participant age groups in the metadata analysis (Fig. 3.3). Across both these metrics and both arm and face samples there was no significant difference observed in any metric, meaning that in this study shows no association with these sociodemographic metrics and skin diversity (Table 5.5). However, the trends seen in the clusters between the age groups does suggest that some more nuanced study that explores sociodemographic factors in its participants may reveal more information about the links between these factors and the skin microbiome, as some corelations were observed in a USA based skin study (Dimitriu *et al.*, 2019). This would be beneficial here as this study is limited in its participant numbers, which creates small groups in this level of focus. Similarly, while the workplace codes are more a measure of sociodemographic area and using them here to infer pollutant exposure could be considered inappropriate, it is still showing some promise as it is limited by the size of the study and its recruitment goals and could become a valid method in a more appropriate study, particularly if pollutant levels were also directly considered with this coding system.

The AUDIT-C measure of alcohol consumption showed promise in the metadata with a significant increase in score for the OA group (Fig. 3.6), as alcohol consumption has body wide effects and is proven to impact the microbiome structure investigating any links between consumption and the skin microbiome ageing was a valid consideration. However, no significant differences were observed here across the alpha and beta diversities and the AUDIT-C scores. This was particularly interesting with the female populations as the YA are

mainly non-drinkers whereas the OA are mixed so there was a possibility for a link, however, the OA-Females beta diversity ellipses heavily overlap (Fig. 5.26), this demonstrates how alcohol consumption is not linked with ageing in this study and that the gendered difference previously seen is not affected by alcohol consumption. There may be a potential for alcohol consumption to influence the skin microbiome in ageing (Liu and Chen, 2023), but this maybe at the extreme ends of the spectrum and is not within the scope of this thesis as the effect may also be masked by other lifestyle choices of the host (Moitinho-Silva *et al.*, 2021).

Our participants skin health and what products they regularly expose their skin will have an impact on the content of their skin microbiomes (Cooper *et al.*, 2015; Wallen-Russell, 2018), the past skin conditions and moisturiser use created the largest subgroups of the participants and therefore were the most appropriate to investigate how they affect the skin microbiome in this study. Past skin conditions did not show any significant differences in either alpha or beta diversity in both skin sample groups, while some trends were seen they are not contributing towards any differences in skin ageing. This is again likely limited by the study not recruiting based on any skin conditions and the study was lacking a dermatologist to design or implement a detailed skin health assessment for the participants during sampling which prevents a more nuanced picture of skin health being used here. In the moisturiser use significant differences were seen in the female populations, however, most female participants reported using moisturiser meaning that these results complicate the gendered differences previously discussed. As the location of moisturiser use was not included, we cannot confirm if the differences seen here are due to different branded moisturiser use between the participants or truly the result of menopausal hormonal change. In the arm samples Bray Curtis ordination plot (Fig. 5.30), there is an overlap in the OA-Females who report no and the YA-Females that report yes but this and the lack of significant difference between these groups could demonstrate an association, but there is a reduced number of participants in this no group. However, the inner elbow is an area where moisturiser use would be considered unusual, so this does indicate how complicated this issue is. Regardless the area of gendered differences in the skin microbiome in ageing warrants further study and these results regarding moisturiser use show the additional factors that would need to be considered in such a study. Overall, the moisturiser usage and skin health may have an impact on the ageing skin microbiome but again this study is limited in its ability to investigate this properly, for example a study has shown use of prebiotic and postbiotic moisturising regimes can stabilise the face microbiome in women (Iglesia, Kononov and Zahr, 2022), suggesting that a detailed consideration of skincare regimes would be required in such a study. Additional work has outlined up to six different combinations of personal care products and there effect on the skin microbiome (Mim *et al.*, 2024), this also summarises the effects on specific bacteria such as moisturisers increasing *Bifidobacteria* while galacto-oligosaccharides reduce *Staphylococcus* which further demonstrates the breakdown of information required from participants while indicating it to be a worthwhile area for future study.

Using differential abundance testing several species were identified in both the arm and face samples as being associated with either the YA or OA age groups (Table 5.7 and Table 5.8). Of those identified in the arm samples *A. parvus, M. pelagius, S. colocasiae*, and *S. desiccabilis* are associated with various environments including plants and soil as such there presence within the samples is expected however it would be inappropriate to draw conclusions about their contribution to the human microbiome or the ageing process (Nemec *et al.*, 2003; Reddy and Garcia-Pichel, 2007; Hamada *et al.*, 2010; Lin *et al.*, 2018). The enterically associated *E. cecorum* is associated with domestic and agricultural animals meaning that it could have transferred to the participants skin during the handling of raw meat or domestic animals (Jung *et al.*, 2018; Lake *et al.*, 2020), again it would be inappropriate to suggest this plays any roles in the skin microbiome; however, the genera has some associations with the gut microbiome so there may be the possibility of cross talk across the gut-skin axis. Both the *A. viridans* and *F. hominis* are associated with pathogenicity rather than serving as commensal organisms but they are capable of being recovered from various human sources (Christensen and Ruoff, 2015; Abat, Garcia and Rolain, 2016; Parrey *et al.*, 2016; Rahmati *et al.*, 2017), again this limits any validity of conclusions about the effect on the skin microbiome. The un-speciated *Bifidobacterium spp.* and *Gemella spp.* belong to genera that are known human and mammalian commensals so while they are not fully classified, they are likely valid components of the participant microbiomes even if they are not normally associated with the skin (O'Callaghan and van Sinderen, 2016; Torres-Morales *et al.*, 2023); however, without further information discussion on their role here would be speculative at best. Similarly, *C. cystitidis* while associated with cattle is a member of a genus that does form part of the normal human microbiome content (Yanagawa and Honda, 1978; Jensen *et al.*, 2023), suggesting that either this occurs here due to transfer from handling meat or an error in the taxonomic classification and it should be a different *Corynebacteria* species. Additionally, the *Staphylococcus spp.* again while not fully classified is likely representative of some normal skin or human microbiome associated species, so more valid conclusions may be drawn about the impact on the ageing microbiome. Finally, *C. acnes* is an accepted member of the human skin microbiome with documented behaviour that is regularly studied for its role in infections (Corvec *et al.*, 2019; Iqbal *et al.*, 2024).

The face samples share the differentially abundant *A. parvus, C. cystitidis,* and *C. acnes*, this is acceptable as both sites are the skin and the top species in this study have demonstrated that they both share some species. Those not shared in terms of differential abundance include the unclassified *Citrobacter spp.* and *Acinetobacter spp.* which along with *P. versuta* are again often isolated in the environment rather than being associated with the human microbiome (See-Too *et al.*, 2017, 2017; Sheck *et al.*, 2023), this prevents any valid conclusions about the role in the microbiome during ageing. *C. lanienae* is considered a normal part of the mammalian microbiome and could be a natural part of the participants microbiome or have been acquired from interaction with a domestic or agricultural animal (Logan *et al.*, 2000). Interestingly *F. animalis* has a complex history and can be considered either a normal part of the microbiome or a potential pathogen which was also considered a sub-

species of *Fusobacterium nucleatum* (Han, 2015; Ye *et al.*, 2017), this makes its potential role in the skin microbiome difficult to determine as the literature around this species is conflicting. Finally, *P. gingivalis*, *S. gordonii* and *S. oralis* are all considered commensals associated with the oral microbiome where they also can be linked to periodontal diseases (How, Song and Chan, 2016; Park *et al.*, 2020; Ren *et al.*, 2024), some research indicates a role played by *Streptococcus* within the skin (G. Kim *et al.*, 2021) while additional research confirms that both genera (along with other oral organisms) can be found on the face across various individuals in various countries (Brandwein *et al.*, 2021; Lee *et al.*, 2021; Ghimire *et al.*, 2024), this suggests that they could be considered a normal part of the face microbiome and play a valid role in the ageing microbiome. Overall, this shows that only some of these species identified as differentially abundant in the arm and face datasets could be contributing to the skin microbiome in ageing, these are, *Bifidobacterium spp., C. acnes*, *C. lanienae, Gemella spp., P. gingivalis*, *Staphylococcus spp.*, *S. gordonii,* and *S. oralis*.

When looking at the arm differentially abundant organisms plotted by relative abundance (Fig. 5.33) only the *C. acnes* and *Staphylococcus spp.* are present in both age groups, additionally these take up a large share of the species present within them. This is contrasted by the *Bifidobacterium spp.* and *Gemella spp.* which are only present within one age group each, along with this their presence is very low in these groups where on average participants do not have these species present. Comparatively the *C. acnes, S. gordonii,* and *S. oralis* identified as differentially abundant in the face dataset are present within both age groups, where the *C. acnes* again take up a large share of the species present within the groups. While the *C. lanienae* and *P. gingivalis* are only present within the OA age group. This allows us to view the changes in the species that occur in both age groups over ageing, opening areas for interpretation and future research into these species and the role they may play. While those that only occur in very low numbers in one group but not the other could just be a result of statistical noise and could require further investigation to clarify if they are genuinely differentially abundant, before conducting any further analysis on the role in the ageing skin microbiome. Additionally, the non-human associated species *A. Parvus, A. viridans*, *C. lanienae, C. cystitidis, E. cecorum, F. hominis, M. Pelagius,* and *P. versuta* are littered with outliers where individual samples have high relative abundances compared to the rest of the samples. Which are also present in the oral associated organisms *P. gingivalis*, *S. gordonii,* and *S. oralis* this suggests that there identification as differentially abundant organisms in this study could be due to these outliers in these samples effecting the statistical analysis.

When these species were tracked through conducting the differential abundance methods on the gender and frailty groups only two species that are in the list previously discussed still occur in the results when filtered by the same criteria used (Table 5.9 and Table 5.11). These are *Bifidobacterium spp.,* and *C.* acnes which are only shown in the arm dataset, this is interesting as C. *acnes* was identified in the face dataset but is not identified by any method in the gender or frailty group. Therefore, it is likely that *C. acnes* has a role in the gendered shift in the skin microbiome previously observed in the arm but not the face, this can be seen in the *C. acnes* abundance

plot broken down by gender where the male groups have similar shaped abundances but the OA-Female have majority low abundance while the YA-Female have higher average abundance than all the other groups. Shifts in *C. acnes* population in age have been reported by other studies but have not necessarily been limited to one gender (Jung *et al.*, 2024; Niedźwiedzka *et al.*, 2024). Meanwhile the *Bifidobacterium spp.* share the similar distributions in both YA gender groups even though they are identified as abundant in the female population. These gender driven differences seen reflect established literature in the significant differences gender has on ageing skin, including skin thickness and hydration level (Keaney, 2016; Dąbrowska *et al.*, 2018). Of additional interest in this study the failure to identify differentially abundant species for the face microbiome when using gender as other studies have (J.-H. Kim *et al.*, 2021), this may be due to this studies limitations in its methods. None of these species track successfully through to the frailty groups, this adds weight to the previous diversity analysis where the results demonstrated that the FI scoring system used here shows no links to the ageing microbiome.

Moving away from specific species roles in the microbiome and into the overall community structure required the construction of co-occurrence networks for each age group and sample type. The arm networks see a larger connected component in the OA group than the YA, but the YA have more estimated positive correlated connections than the OA (Table 5.13). However, as there is no significant difference between these two elements, we cannot conclude anything from them. Instead, when looking at overall centrality metrics a significant difference is seen, demonstrating that they are significantly dissimilar form each other. This can be seen for example in the statistically calculated hub taxa where the OA relies on various *Staphylococcus spp.* and the YA on *Glutamicibacter spp.*, this suggests that the linking co-courant factors in the YA are focused on an understudied genus associated with environmental isolation rather than one associated with human habitation (Ghimire *et al.*, 2024). However, such inferences are limited due to the statistical nature of this identification method and the quality of the dataset used, this difference could equally be due to the increased diversity in the YA population compared to the OA. Looking at the species identified through differential abundance organisms more are present in the OA network when compared to the YA, however those that are in both have different connections, particularly the *C. acnes* which is connected as part of the YA LCC but not in the OA network. Overall, these changes in network structure show that the reliance on some species as key parts of the microbiome has changed, this is represented in the shift away from *C. acnes* and an increase in *Staphylococcus spp.* presence in the OA group.

Moving onto the face co-occurrence network, the YA group has a larger LCC which consists of most of the taxa used to construct the networks, this results in a single large network which is significantly larger than the OA LCC and whole network; while the YA network does contain more negatively estimated correlated connections this is not significantly so (Table 5.14). Again, a significant dissimilarity between the networks was seen when reviewed using a Jaccard index. In hub taxa terms this sees the OA more reliant on a variety of

Micrococcales, which as an order in taxonomic terms limits any meaningful conclusions. Instead, here both groups use *H. parainfluenzae* as a hub organism, this is a normal part of the human microbiome often associated with the oral cavity, this suggests that more than just typical skin species can play a role in the skin microbiome (van Rensburg *et al.*, 2015; Tseng *et al.*, 2025). When looking at the species identified through differential abundance a reversal of the arm network is seen, and in the face more species are present in the YA group compared to the OA. Given the nature of the OA network this also sees these species disconnected and less likely to co-occur, while the YA sees these linked, especially the *S. gordonii* which has also been statistically identified as a hub taxa but this was not abundant in the YA group. Which suggests that key parts of the community structure may not be associated with being prominent members. Overall, the face networks suggest that the YA have a more connected and established community than the OA, which may make the OA face microbiomes more susceptible to colonisation of atypical species. These results across the skin microbiome have been similarly identified in Asian based studies into the skin microbiome during age (Kim *et al.*, 2019, 2022), suggesting that the loss of community structure in the skin microbiome is universal in older age and that this loss is due to similar species loss as these studies also identified a reduction in *C. acnes* population in their older populations; however, this comparison is limited to the face microbiomes as no other data on arm co-occurrence networks and ageing is currently available.

Using the post taxonomic predictions of metabolomic pathways allows for the investigation of how the microbiomes may be functioning on the host and identify any differences between the age groups, this is important to demonstrate if the identified population differences are impacting ageing or if the changes are the result of the host age. When looking at the beta diversity of the pathway abundances only significant differences were observed in the face pathways (Fig. 5.40), this is different from the population diversities where differences are seen in both sample sets. This demonstrates that there are differences in the pathways present in the age groups in the face samples but not in the arm samples. After this an in-depth analysis of the pathways using the same differential abundance methods as used on the taxonomic data only five pathways were consistently identified across all methods (Table 9.5), and these were within the face dataset. This result combined with the beta diversity results suggest that within these datasets the pathways present in the face microbiome are contributing more to differences than in the arm microbiome. Specifically of the pathways identified through differential abundance they are all associated with breaking down various molecules. Three of the these have an established role in the host microbiome interface, 4-methylcatechol is associated with gut health along with cellular recovery and increased immune function (Walker and McKinney-Freeman, 2019; Behera *et al.*, 2020; Y. Guo *et al.*, 2022), arginine is associated with immune function modulation, skin cell differentiation, and antimicrobial production (Nüse *et al.*, 2023; Szondi *et al.*, 2025), finally nitrate reducing bacteria are considered to be beneficial to various diseases (Rosier *et al.*, 2020; Liu *et al.*, 2023). Of the remaining two phenylacetate is associated as an intermediary metabolite for other aromatic compounds in microbial metabolism including the

production of toluene (Luengo, García and Olivera, 2001; Teufel *et al.*, 2010; Zargar *et al.*, 2016), the toluene itself is considered a pollutant on the skin where exposure can cause long lasting effects including premature ageing and it is also considered linked to changes in the microbiome structure (Leung *et al.*, 2020). From these results it is unclear how or why these have increased in the OA group, it is possible that the OA have been exposed to pollutants for longer which have caused a shift in the microbiome to increase detoxification of the skin, this is countered by a possible increase in the production of similar compounds by the microbiome. Similarly there is no immediate explanation for the increase in degradation of compounds that may be potentially beneficial to the host, however it may simply be that more of these compounds are present in the OA or that the shifts in microbiome structure in the OA have resulted in a more "selfish" microbiome that is prioritising these compounds for its own purposes. Similar studies have reported and concluded various different results in changes to the skin microbiome metabolome through ageing (Alkema *et al.*, 2021; Russo *et al.*, 2023; Zhou *et al.*, 2023), suggesting more research is needed. Regardless of why these are identified here, there very identification allows for future research studies to be conducted either focusing on these kinds of pathway interactions or investigating population changes based on pollutant or similar compound exposure; additionally, these results indicate that the changes in population within the OA face microbiome are also associated with changes in microbiome functionality.

During the preparation, analysis and construction of this chapter this studies funder Colgate-Palmolive carried out their own independent analysis carried out by Mary Hannah Swaney under the supervision of Min Li. While they used the same raw sequence output, they operated a different and more conservative pipeline in Qiime2, this featured a higher level of sequence input quality and a finer range of final sequence lengths, additionally a different custom classifier was used. While both this thesis and their analysis follow similar lines in terms of diversity, mathematical modelling, and post taxonomic investigations, for example this thesis relied on phyloseq objects in R whereas the Colgate-Palmolive analysis does not use this package. This results in different results seen in both analyses, for example more significant differences are observed in alpha diversity measures in their analysis and some different organisms are identified in the differential abundance tests, which is partially expected as different differential abundance methods have been used. Specifically, there analysis identifies *Staphylococcus capitis* as a differentially abundant organism, which could conceivably be in the *Staphylococcus spp.* identified here. Additionally, this analysis has not done the Dirichlet multinomial mixtures analysis to create community types, this was due to time constraints, however, their analysis has only focused on the age groups and has not covered any of the additional metadata which has been explored here, this allowed the development of a further understanding of the changes the microbiome undergoes in ageing particularly around gender. While both analyses show differences in their respective results, they also both have similarities such as identifying beta diversity differences and the changes in *C. acnes* during ageing, this means

that despite having different valid reasons for the pipeline differences, these do not matter as they result in the same conclusions.

This investigation into the skin microbiome in ageing does suffer from a range of limitations. Initially there have been problems in recovering the DNA from the samples, but these have been outlined previously. One issue that does carry forward from the pipeline optimisation chapter is from the classifier not fully speciating some of the taxa, this has prevented possibly meaningful conclusions from being drawn as only generalisations can be made from genera or other higher taxonomic rank information, this could be improved by creating a better classifier in a future study as multiple factors effect classifier quality which has been proven to effect microbiome study analysis (Bokulich *et al.*, 2018; Walsh *et al.*, 2018). Broadly the whole analysis is limited by only having one sample set from each participant, this prevents the creation of averaged microbiomes for the participants which would allow for the confident elimination of transient contaminant species that could be creating noise in the diversity statistics and in the differential abundance results. Additionally multiple samples over time would enable the investigation into the resilience and alterations of the skin microbiome in the different age groups, the importance of this has been identified in similar studies (Hillebrand *et al.*, 2021). Similarly, more participants across all age ranges would enable the investigation into the changes in the skin microbiome over ageing totally rather than just at the specific age ranges used here.

This chapter has covered several additional breakdowns of the dataset using other aspects of the participant metadata than just the age groups, while this has enabled an exploration of the differences or similarities seen in the age groups these have been limited by the study only being powered for the age groups. In future studies the number of participants recruited should be increased or the age groups alone should not be the sole factor for recruitment, this for example would enable larger gendered subgroups to be created enabling a better investigation into the differences seen here. Similarly, while this study shows no links between frailty and the skin microbiome it maybe that the health of the OA participants was too good and that if frailty was a metric used during recruitment, then links maybe established. Along with this no components of skin health were assessed in any detail, such as thickness or elasticity measurements, if these were included it may have been possible to understand more about the health of the participant population and how their skin was ageing. Finally, some of the difficulties experienced in this chapter have revolved around decisions to do with the selection of data filtering and statistical methodologies, in the future this could be made easier by seeking advice from more experienced bioinformaticians or statisticians as required.

5.5. Conclusion

In summary this chapter analysed the arm and face datasets created previously in terms of diversity, mathematical modelling, and predictive metabolism, where possible this was also broken down not just in terms of the age groups but in terms of participant gender, FI, and other potentially relevant metadata subgroups. In

doing so it demonstrates that there are significant differences in the skin microbiome during ageing, some of which can be traced through to gender specific differences. In broad diversity terms there is a significant difference in beta diversity seen in both the arm and face samples between the two age groups, this is then repeated between the two age groups female populations, but these differences are not seen in any of the other subgroups used here. The organism *C. acnes* was found to be differentially abundant between the age groups in both skin sites, however only in the arm data was it also found to be abundant in the gender subgroups; suggesting it plays a key role or change in the skin ageing process, which may be due to changes in skin lipid and hormonal secretions in older age. Meanwhile co-occurrence network analysis showed a change in the key members of the skin microbiome in ageing, specifically how in the arm the YA have *C. acnes* as a connected member of the community, which is not shown in the OA, then the face networks show that the YA are more likely to have a resilient microbiome than the OA; which is further reflected in the loss of dominance of *C. acnes* in the OA when compared to the YA when viewed using relative abundance. Finally, analysis of the predictive metabolomics shows how the behaviour of the face microbiome changes during age and therefore providing a base for future investigations into ageing skin microbiome research.

Chapter 6. Oral analysis

6.1. Introduction

As the 16S datasets of the arm and face samples have been analysed this leaves only the oral dataset to be characterised to determine any changes or similarities in ageing; which is the main focus of this thesis. As the oral dataset contains samples not derived from the skin and have been sequenced targeting a different variable region performing direct comparisons within the datasets would be inappropriate.

In the oral cavity ageing experiences, a loss of epithelial thickness affecting multiple tissues, along with a reduction in salivary gland output which also undergoes chemical content changes affecting the immune response. Additionally, tooth enamel can become brittle and dental avascularity occurs, coupled with a loss of masticatory muscle functionality can cause problems when eating (Thompson and Chen, 2021; Houge and Ruiz, 2022). These changes are a mixture of intrinsic and extrinsic factors interplaying, in addition to the mitochondrial dysfunction, cellular senescence, impaired immune function contributing to the previously outlined changes there are several extrinsic factors in play caused by dental health and hygiene decisions during life. These start with a diet rich in sugary foods and poor dental hygiene can lead to increasing incidents of dental caries and periodontitis. The standard treatment for caries is the removal of the damaged dental material and replacement with a dental amalgam or similar material, this may eventually lead to teeth being removed and replaced with some form of denture or permanent implant (Selwitz, Ismail and Pitts, 2007). Aside from these, diet will also affect the wear on the enamel surface overtime, along with alcohol and smoking increasing oral cancer risks and other complex dental diseases (Lussi, Schaffner and Jaeggi, 2007; Agnihotri and Gaur, 2014; Priyanka, 2017). Within the microbiome increases in desiccant resistant species can be seen in those who use dentures or other temporarily placed structures, in those missing significant portions of teeth the whole microbiome may shift to resemble that of a young child, while more generally changes in salivary content will affect biofilm formation (Gazdeck *et al.*, 2019; Schwartz *et al.*, 2021).

As outlined in chapter 1 (1.5), and summarised in chapter 5 (5.1), a broad selection of methods can be used to analyse the microbiome and characterise it (Fig 6.1). As the oral dataset is different from the previous skin datasets some minor changes to the protocols were required to accommodate these differences, additionally different metadata groups were relevant to oral health and hygiene.

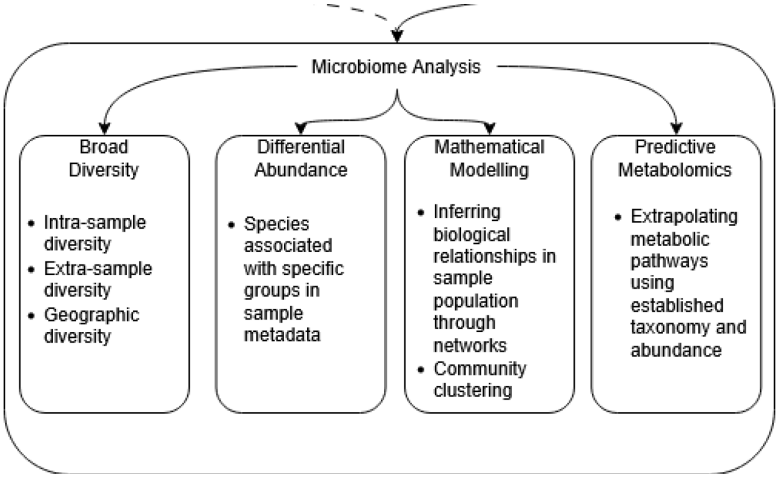


Figure 6.1 Excerpt of Figure 1.5. *Focusing on the microbiome analysis stage of amplicon based 16S microbiome studies.*

6.1.1. Aims

- Provide an overview of the content of the oral microbiome using the previously established dataset.
- Describe and compare the oral microbiome diversities using the age groups and other potentially relevant metadata.
- Identify any notable species associated with the YA and OA oral microbiomes.
- Investigate the community structure of the oral microbiome and any changes during ageing.
- Investigate any changes in the age groups microbiome metabolome in the oral microbiome.

6.2. Methods

6.2.1. Importing dataset into R

The oral dataset previously created in a Qiime2 environment and the participant metadata were imported into R Studio using qiime2R (Jordan E Bisanz, 2018), and converted into a phyloseq object (McMurdie and Holmes, 2013).

6.2.2. Data normalisation

The rarefy_even_depth function was used to normalise the dataset, the oral was normalised to a depth of 100000. During this stage the blank samples were removed from the datasets.

6.2.3. Statistical analysis

Broad ecological diversity metrics using alpha and beta diversity were calculated and statistically compared using the same methods as Chapter 5 (5.2.4).

Differential abundance analysis was carried out using the non-normalised dataset again using the same methods as in Chapter 5 (5.2.4).

To construct co-occurrence networks the non-normalised dataset was collapsed to the species level and then filtered to remove taxa that did not occur in more than 50% of the samples, the construction was carried out in accordance with the methods outlined in Chapter 5 (5.2.4).

Predictive metabolomics was conducted using PICRUSt2 within the Qiime2 environment again using the same methods as in Chapter 5 (5.2.4).

6.3. Results

6.3.1. Top species

After the dataset was finalised, there were 4089 unique features (Table 4.12), these translated to 619 unique taxonomies at the species level. Investigating the top ten species shows that *Streptococcus salivarius* is the most abundant species across the dataset, followed by *Rothia mucilaginosa*, *Haemophilus parainfluenzae*, *Veillonella dispar*, *Prevotella melaninogenica, Campylobacter concisus, Granulicatella adiacens, Streptococcus parasanguinis_clade_411, Veillonella atypica,* and *Fusobacterium periodonticum*. When plotted by relative abundance across all samples, these ten species account for less than half the content in the majority of samples (Fig. 6.2). Although *S. salivarius* was the most abundant species, it is variable in its content within the samples as some have a lower abundance than others or in the case of OA-30 none. Similarly in OA-24 *C. concisus* accounts for nearly 50% of the species present with the sample.

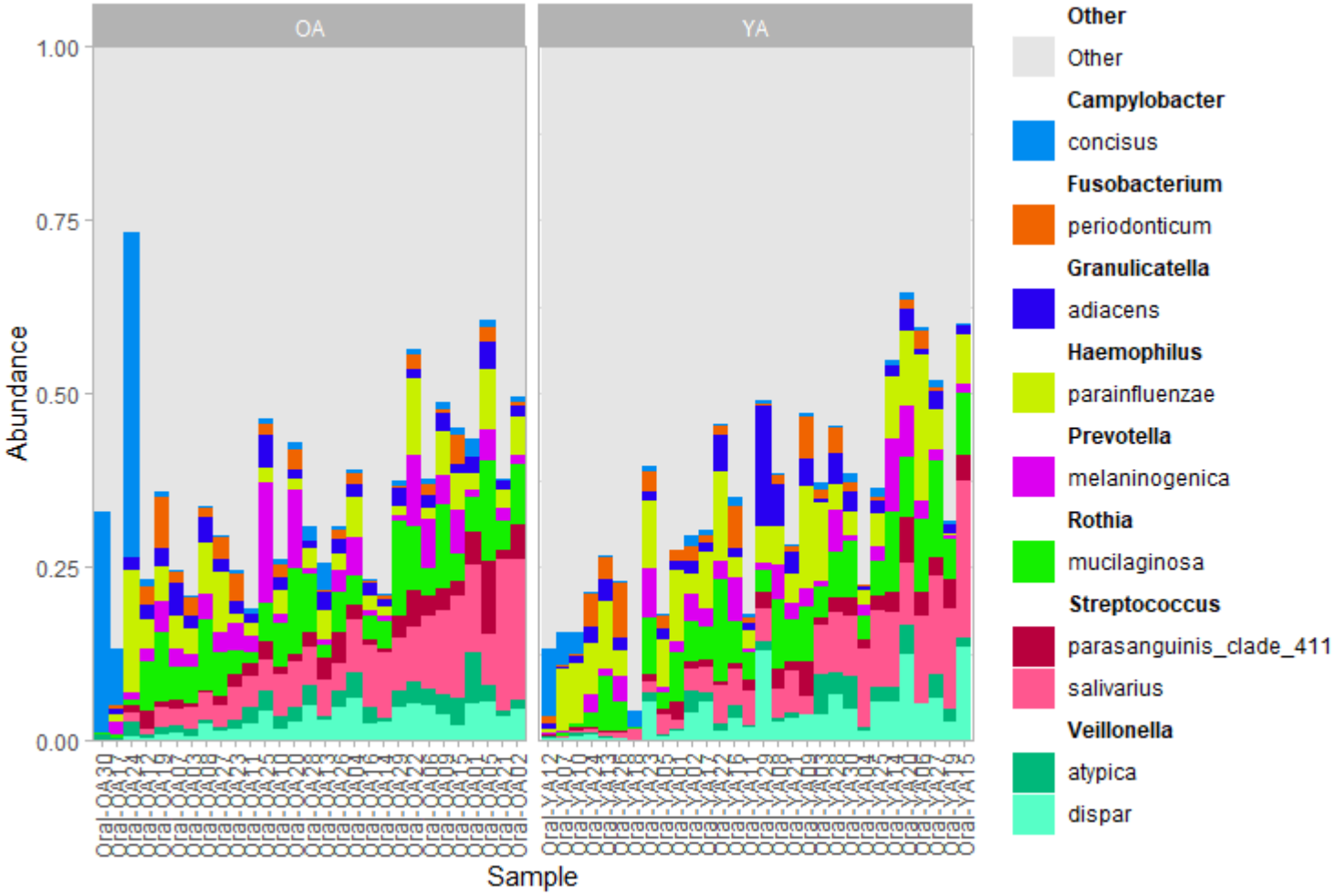


Figure 6.2 Graphs showing the percentage abundance of the species within the oral samples split by age group. *The top 10 species are highlighted, the samples are arranged in ascending order of the genus with the largest proportion in all samples.*

6.3.2. Microbial diversity

After normalising the datasets through alpha rarefaction, the broad diversity metrics for alpha and beta diversity were then calculated.

6.3.2.1. Age Group

Clustering the alpha diversity metrics by the age groups showed no significant differences across the four alpha diversity metrics used here (Fig. 6.3). In all these metrics there is a heavy overlap in both age groups diversity scores, while the YA are seen to have a lower average Shannon and Simpson diversity, and Pielou evenness than the OA this is not significantly so. Therefore, a more detailed investigation of the diversity scores and age was warranted, to investigate these overlaps.

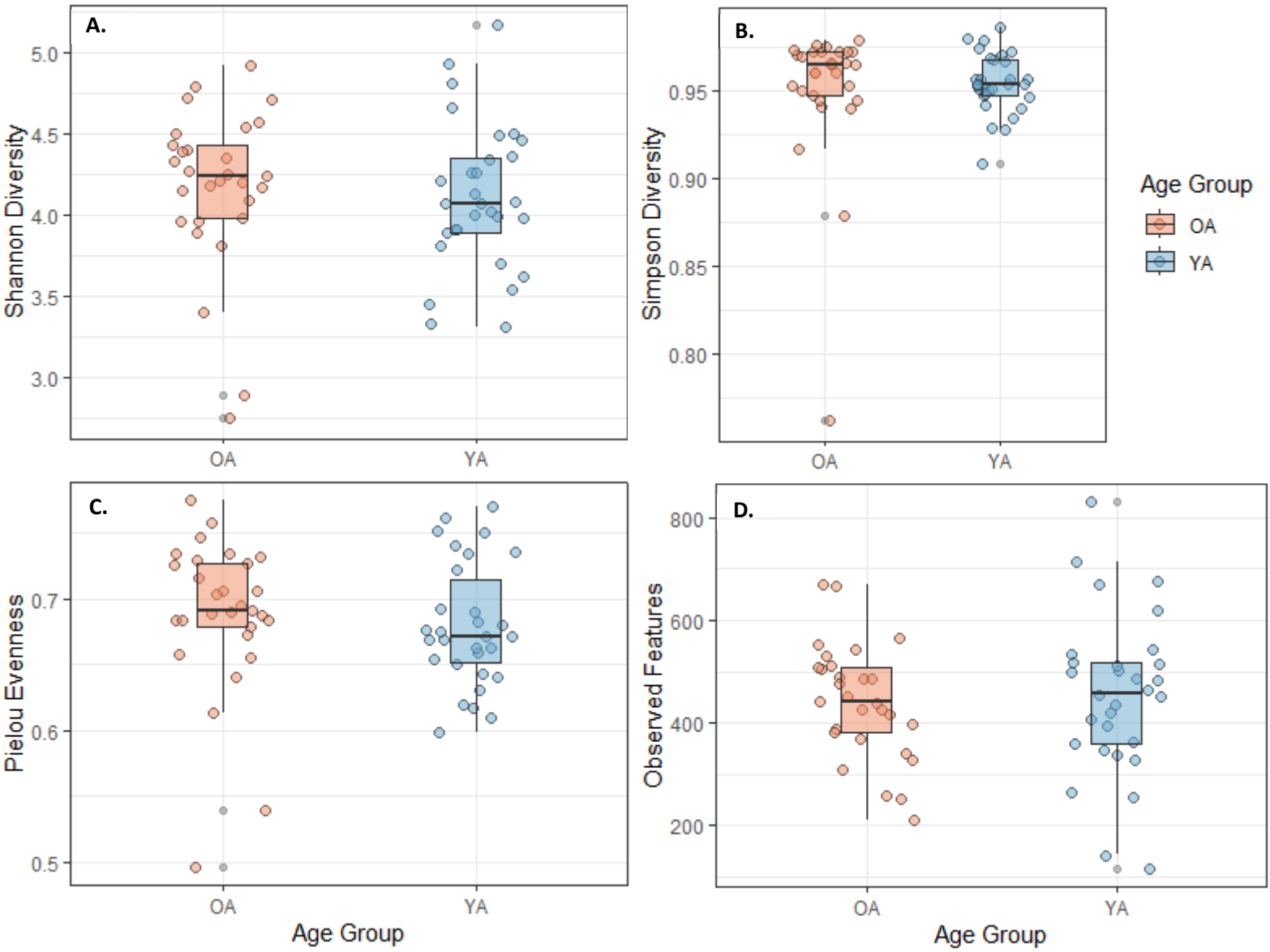


Figure 6.3 Alpha diversity metrics graphs for the oral samples, split by age groups. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

When the Shannon diversity scores are plotted against the participant ages the YA are shown to have a consistent diversity whereas the OA have a slight increase with age (Fig. 6.4) with the regression line for the YA having an R value of 0.059 and the OA of 0.28. However, the p value for both these lines is not below 0.05, which is demonstrated by the wide range of diversity scores seen across both age groups. Exploring this variation concludes the alpha diversity age group analysis, further ecological analysis can be conducted using beta diversity.

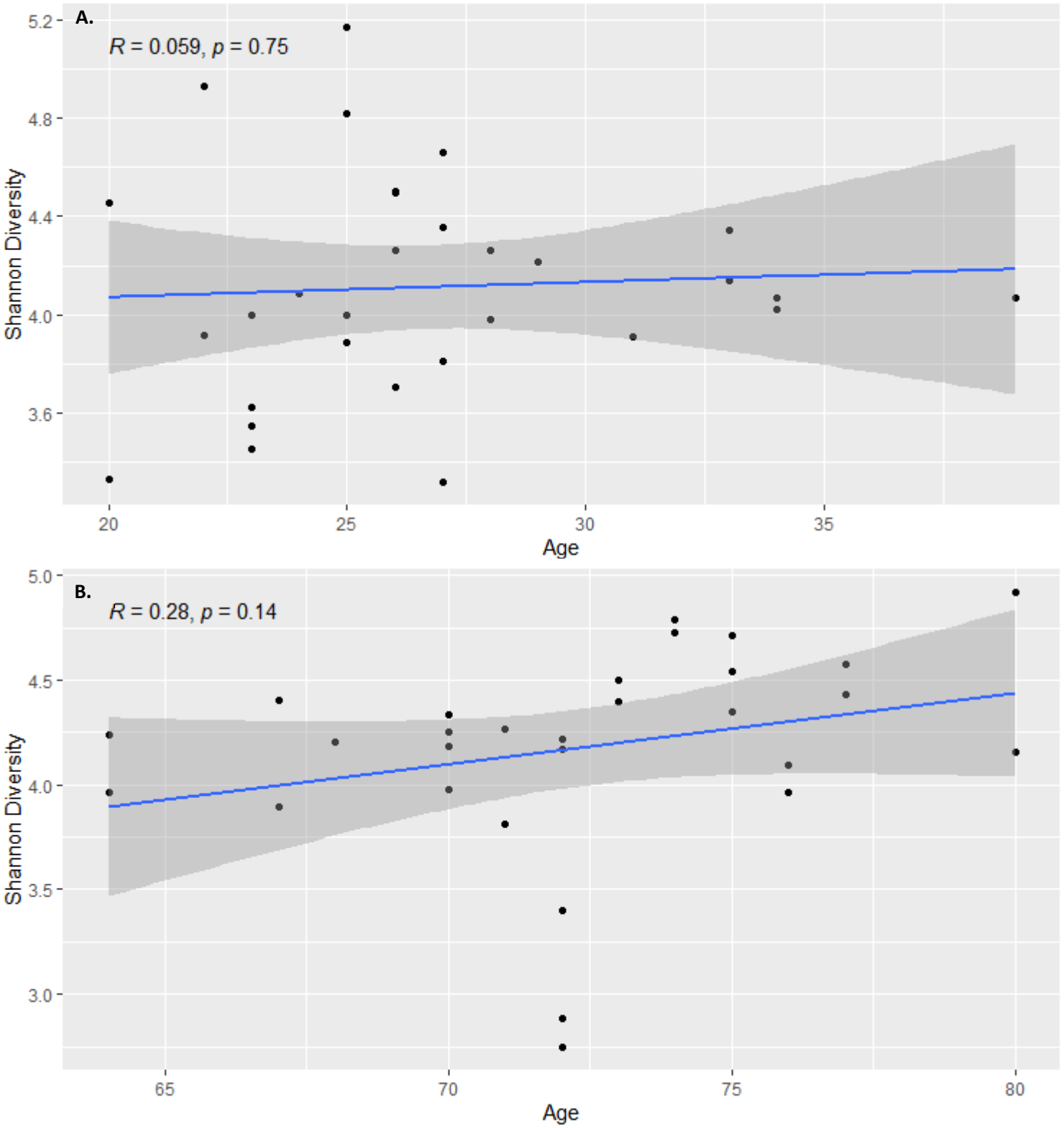


Figure 6.4 Shannon diversity graphs for the oral samples plotted by sample age, including regression line. *Showing an increase in diversity correlated with age. A. YA age group. B. OA age group.*

The Bray Curtis dissimilarity, Jaccard index, and weighted UniFrac beta diversity metrics showed no significant differences when clustered by age group, a significant difference with a p-value of 0.0014 is present in the unweighted UniFrac metric (Fig. 6.5, and Fig. 9.18). This is shown by the heavy overlap in the clustering ellipses in the Bray Curtis, Jaccard index, and weighted UniFrac ordination plots, but distinct clusters are seen in the unweighted UniFrac plot. Therefore, the only significant difference observed between the age groups in broad ecological terms is in the presence and absence of shared species between the groups.

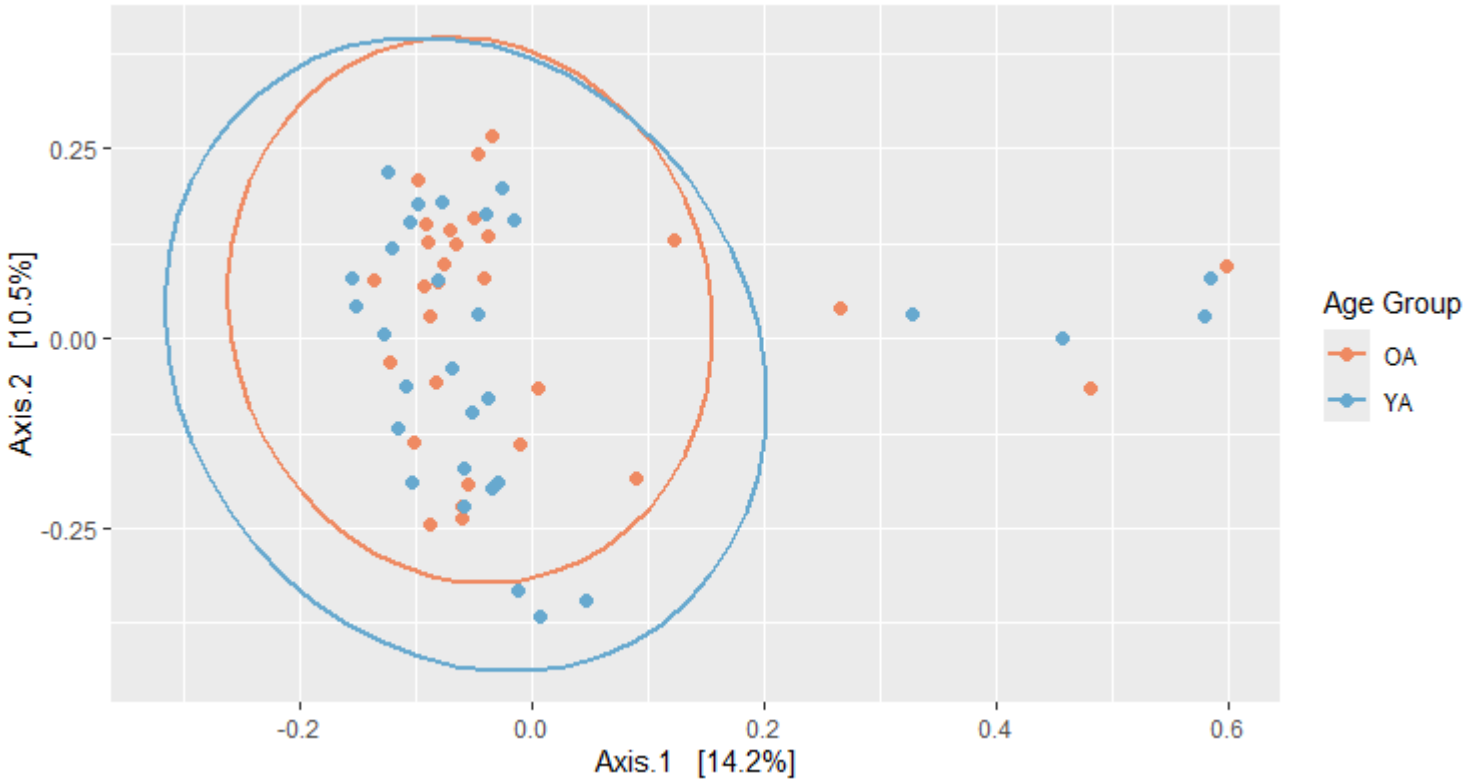


Figure 6.5 Bray Curtis dissimilarity measure of beta diversity as an ordination plot of oral samples. *Where no significant difference was observed.*

6.3.2.2. Gender

The gender of the participants splits them off into large groups that may be relevant to the oral ageing process. When used to compare the alpha diversities no significant differences are seen, however, YA-Females are consistently lower in Shannon and Simpson diversity, and Pielou evenness than the OA-Female (Fig. 6.6), whereas the male groups overlap in these metrics. All genders heavily overlap in observed features. As these gender groups cannot be reasonably broken down more within the scope of this study, it is appropriate to move to the beta diversity analysis.

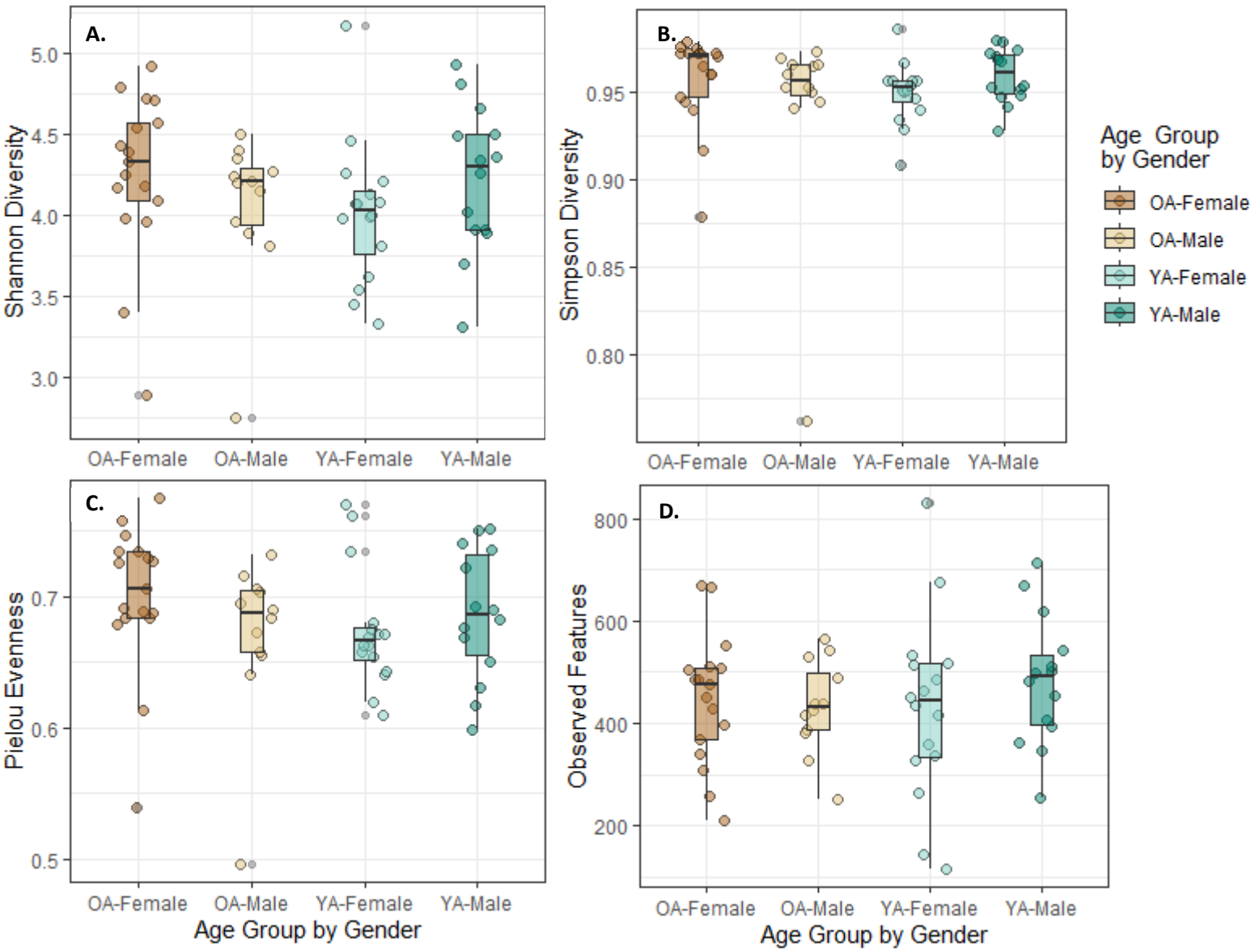


Figure 6.6 Alpha diversity metrics graphs for the oral samples, split by age group gender. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

The beta diversities of the oral samples when clustered by gender showed no significantly different groups in Bray Curtis dissimilarity, Jaccard index, or weighted UniFrac (Fig. 6.7, and Fig. 9.19). However, a significantly different interaction occurred between the OA-Females and YA-Males in the unweighted UniFrac metric with a p-value of 0.012 (Table 6.1), this can be seen as two distinct clusters of these groups in the ordination plot. Again, like the age groupings the only significant difference observed between these gender age groupings is in the presence and absence of shared species between the groups.

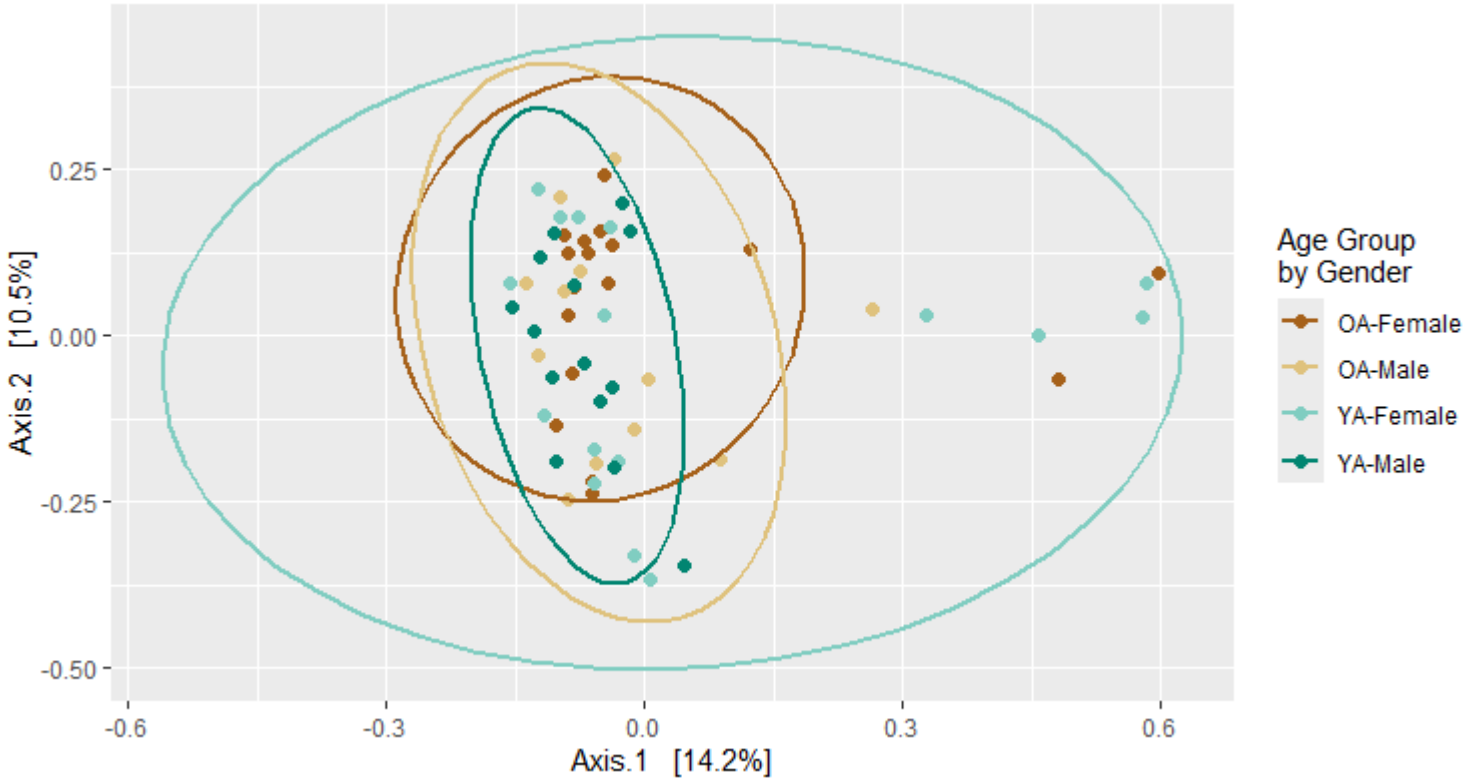


Figure 6.7 Bray Curtis dissimilarity measure of beta diversity as an ordination plot on oral samples. *Split by age group and gender, where no significant differences were observed.*

Table 6.1 Beta diversity p-values for oral sample interactions between the age groups sub-divided into genders. *P-values below 0.05 are highlighted in green.*

	Beta diversity measure			
Interaction	Bray Curtis dissimilarity	Jaccard index	Unweighted UniFrac	Weighted UniFrac
OA-Female vs OA-Male	1.000	1.000	1.000	1.000
OA-Female vs YA-Female	1.000	1.000	0.588	1.000
OA-Female vs YA-Male	0.606	0.714	0.012	1.000
OA-Male vs YA-Female	0.984	1.000	0.906	1.000
OA-Male vs YA-Male	1.000	1.000	0.072	1.000
YA-Female vs YA-Male	1.000	1.000	1.000	1.000

6.3.2.3. Frailty Index

Similarly to the gender groups the FI used to score participants' health can be used to look at their microbiome diversity information, as it may be relevant to the oral ageing process. Clustering for the alpha diversities by the FI scores shows no significant differences between any of the groups (Fig. 6.8). When looked at in detail most of the groups have overlaps in these alpha diversity scores, although a tight clustering is seen in the OA-0 and OA-1 groups. Again, these groups cannot be further broken down for additional alpha diversity investigations, it is therefore appropriate to move to the beta diversity analysis.

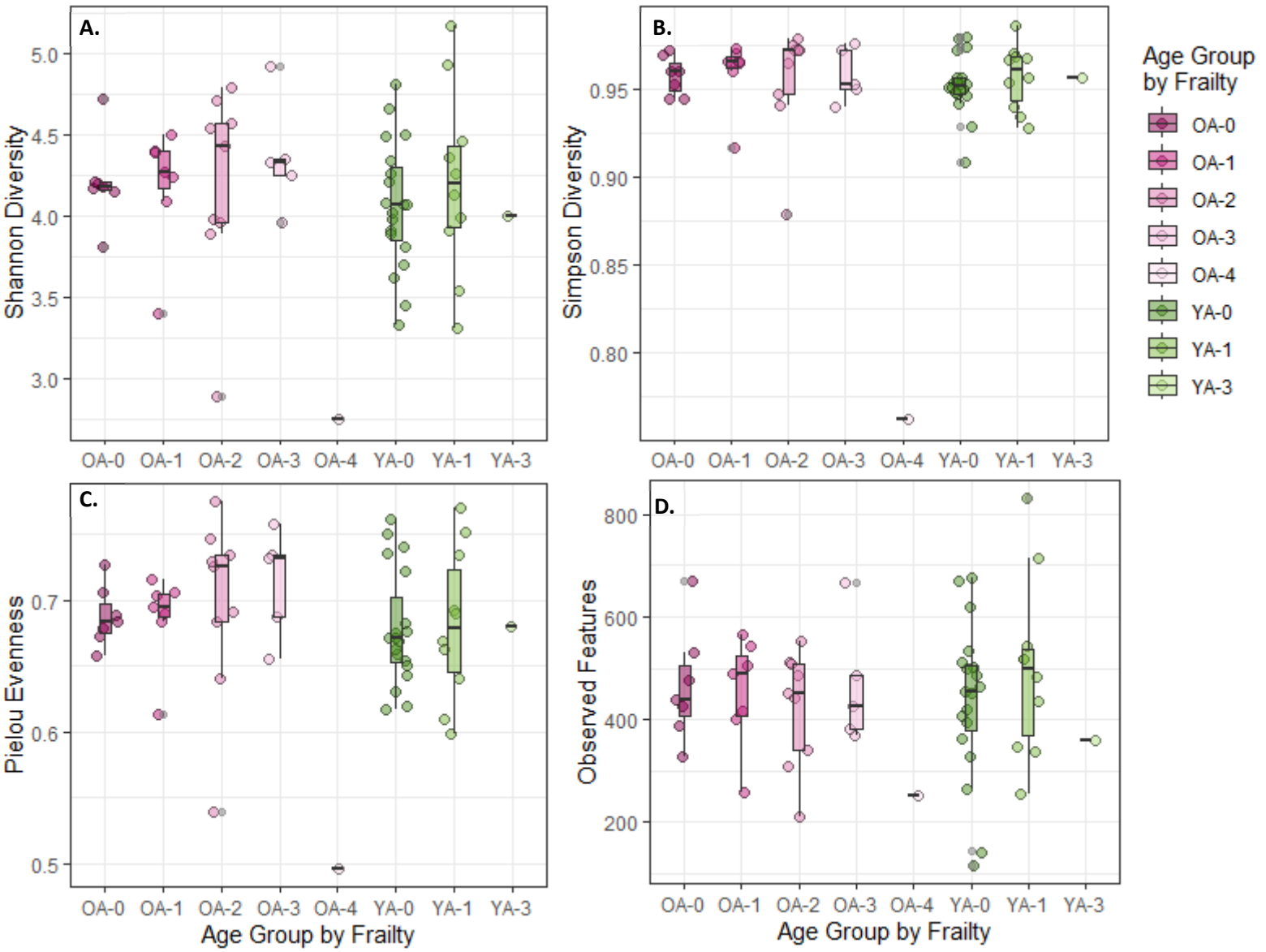


Figure 6.8 Alpha diversity metrics graphs for the oral samples, split by age group and frailty index. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

When the FI scores were used cluster the beta diversity scores no significant differences are seen in any metric or interaction (Fig. 6.9, Fig. 9.20, and Table 6.2). This is demonstrated as a heavy overlap of each clusters ellipses in the ordination plots. Overall, this showed no significant differences in any microbial ecology diversity metric when the oral microbiome diversity data is analysed using the FI system established in chapter 3.

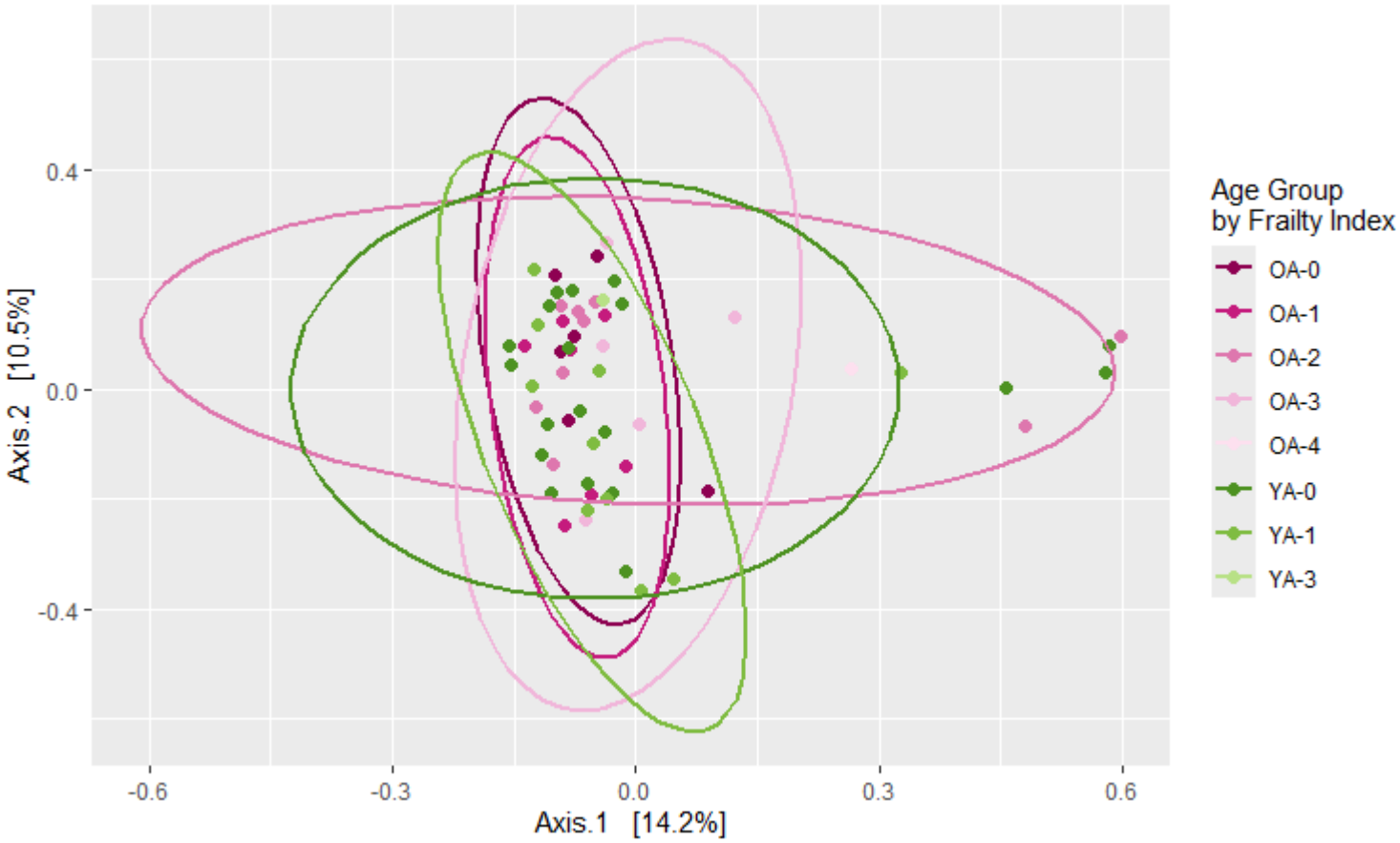


Figure 6.9 Bray Curtis dissimilarity measure of beta diversity as an ordination plot on oral samples. *Split by age group and Frailty Index score, where no significant differences were observed.*

Table 6.2 Beta diversity p-values for oral sample interactions between the age groups sub-divided into FI scores. *P-values below 0.05 are highlighted in green.*

	Beta diversity measure			
Interaction	Bray Curtis dissimilarity	Jaccard index	Unweighted UniFrac	Weighted UniFrac
OA-0 vs OA-1	1.000	1.000	1.000	1.000
OA-0 vs OA-2	1.000	1.000	1.000	1.000
OA-0 vs OA-3	1.000	1.000	1.000	1.000
OA-0 vs OA-4	1.000	1.000	1.000	1.000
OA-0 vs YA-1	1.000	1.000	1.000	1.000
OA-0 vs YA-0	1.000	1.000	1.000	1.000
OA-0 vs YA-3	1.000	1.000	1.000	1.000
OA-1 vs OA-2	1.000	1.000	1.000	1.000
OA-1 vs OA-3	1.000	1.000	1.000	1.000
OA-1 vs OA-4	1.000	1.000	1.000	1.000

OA-1 vs YA-1	1.000	1.000	1.000	1.000
OA-1 vs YA-0	1.000	1.000	1.000	1.000
OA-1 vs YA-3	1.000	1.000	1.000	1.000
OA-2 vs OA-3	1.000	1.000	1.000	1.000
OA-2 vs OA-4	1.000	1.000	1.000	1.000
OA-2 vs YA-1	1.000	1.000	1.000	1.000
OA-2 vs YA-0	1.000	1.000	1.000	1.000
OA-2 vs YA-3	1.000	1.000	1.000	1.000
OA-3 vs OA-4	1.000	1.000	1.000	1.000
OA-3 vs YA-1	1.000	1.000	1.000	1.000
OA-3 vs YA-0	1.000	1.000	1.000	1.000
OA-3 vs YA-3	1.000	1.000	1.000	1.000
OA-4 vs YA-1	1.000	1.000	1.000	1.000
OA-4 vs YA-0	1.000	1.000	1.000	1.000
OA-4 vs YA-3	Invalid	Invalid	Invalid	Invalid
YA-1 vs YA-0	1.000	1.000	1.000	1.000
YA-1 vs YA-3	1.000	1.000	1.000	1.000
YA-0 vs YA-3	1.000	1.000	1.000	1.000

6.3.2.4. Additional metadata groups

Previously in chapter 3 the additional metadata gathered from the participants was reviewed and analysed, identifying potential groups of interest or relevance to the microbiome in this study either through creating significantly different clusters or of specific sample relevance.

Again, the super area output code which groups the participants by their postcodes into convenient clusters of demographics (Fig. 3.3). When used to cluster the alpha diversities metrics no significant differences are seen between the age groups or between any super area output code within any age group. There were consistent overlaps in the diversities across the age group and supergroups, although the OA 3B and 3D cover the largest range in Pielou evenness, Shannon and Simpson diversity (Fig. 6.10). However, further investigation

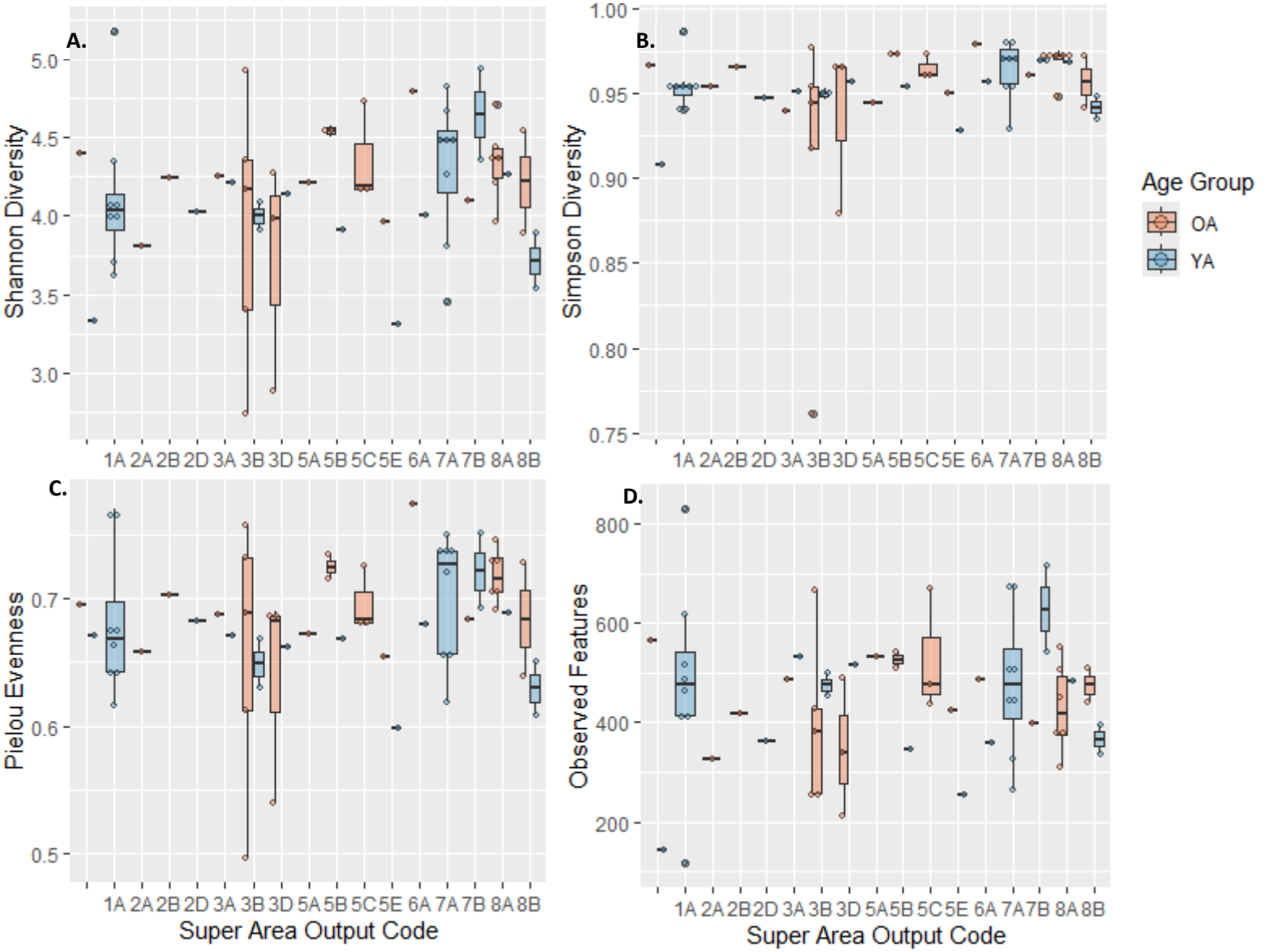


Figure 6.10 Alpha diversity metrics graphs for the oral samples, split by age group and Super area output code. *Showing no significant differences. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

of these groups and their effects on alpha diversity are beyond the scope of this study, therefore the next stage of analysing these groups is to use beta diversity.

When used on the beta diversity metrics again no significant interactions are seen between the age groups across the super area output codes (Fig. 6.11, and Fig 9.21). There are only enough participants to enable ellipses to be generated for the 1A, 3B, 7A, and 8A groups, these ellipses see heavy overlap in the ordination plots. Overall, this showed no significant differences in any microbial ecology diversity metric when the oral microbiome diversity data is analysed using the super area output code established in chapter 3.

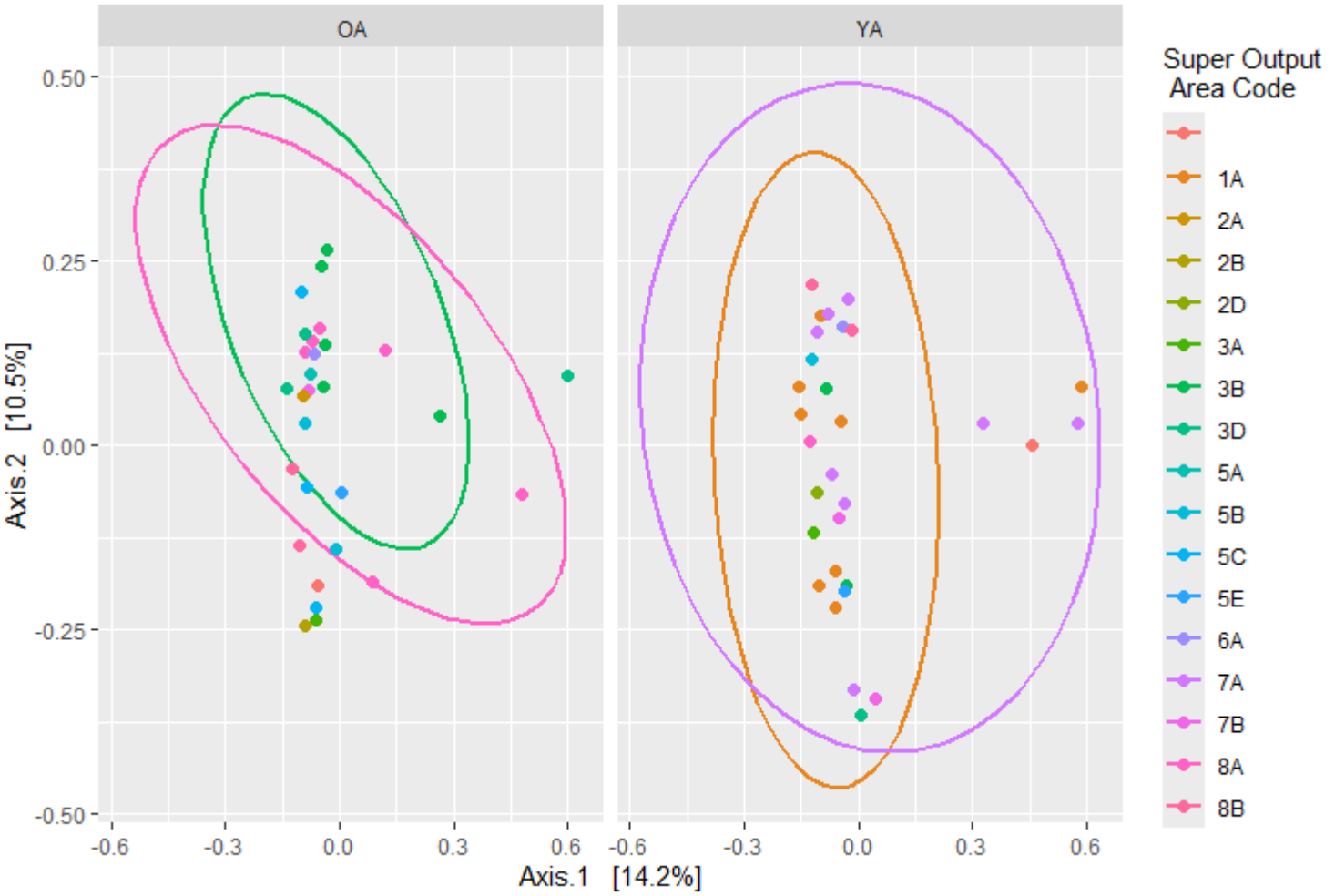


Figure 6.11 Bray Curtis dissimilarity measure of beta diversity as an ordination plot on oral samples. *. Split by age group and super area output code, where no significant differences.*

Another sociodemographic tool that can be used to cluster the participants is the workplace zone code, this indicates the types of industry that operate within the area of the participants addresses. The participants showed a significant difference between the age groups when clustered with using this metric (Fig. 3.3), additionally pollutant exposure has shown to affect the microbiome; therefore, it is pertinent to explore any contribution in this study. However, when used to cluster the alpha diversity metrics no significant differences are seen, again a consistent overlap within and across the age groups and the workplace zone codes is seen (Fig. 6.12). Here there is a large range covered by YA groups B5 and C1, along with large ranges seen in OA C1 and C4 in Pielou evenness, Shannon and Simpson diversity. However, further investigation of these groups and their

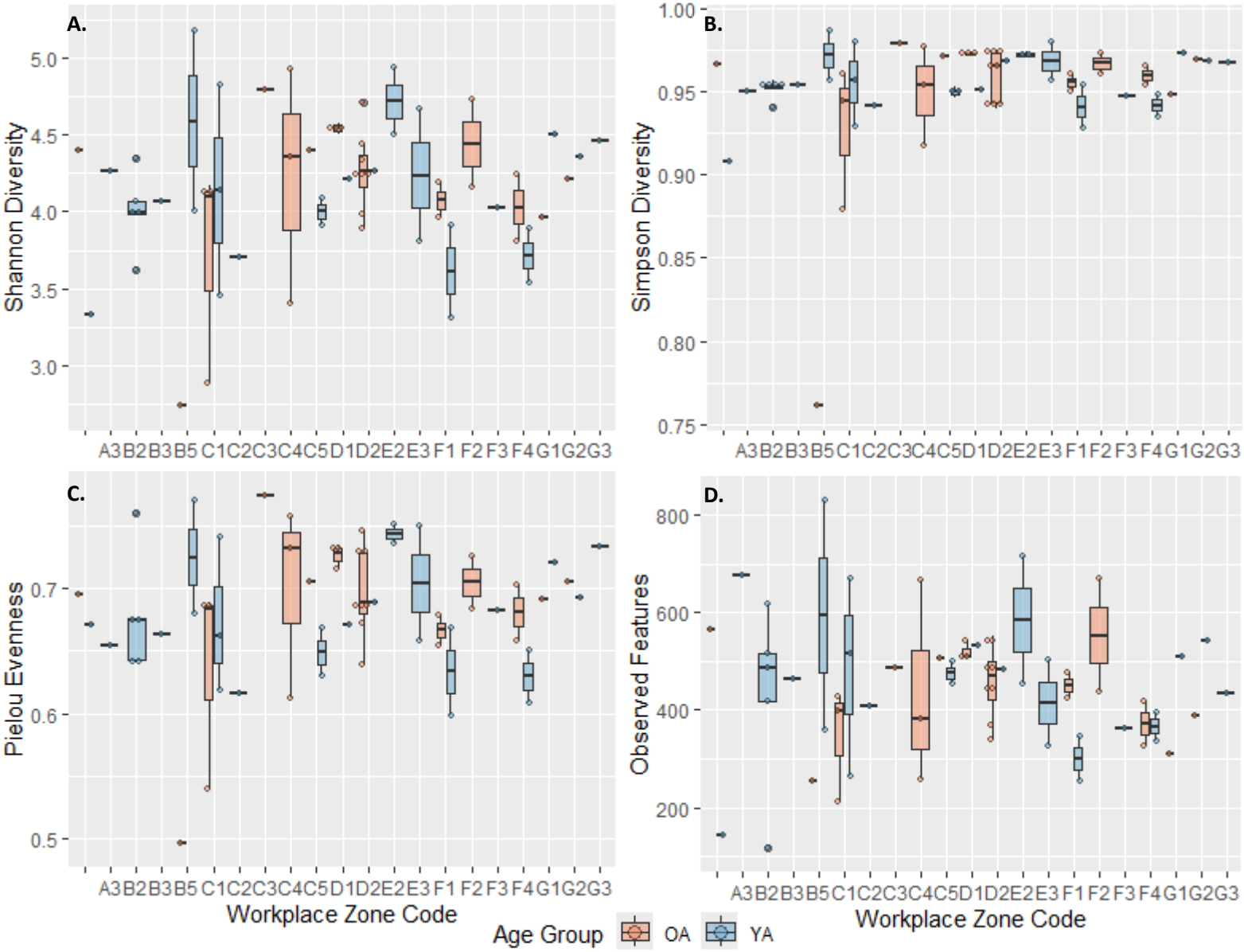


Figure 6.12 Alpha diversity metrics graphs for the oral samples, split by age group and workplace zone code. *A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

effects on alpha diversity are beyond the scope of this study, therefore the next stage of analysing these groups is to use beta diversity.

These workplace zone groups were then used to cluster the beta diversity metrics. There were no significant differences seen in these groups, where clustering ellipses on the ordination plots show overlaps in the B2 and B2 groups across all metrics (Fig. 6.13, and Fig. 9.22). Overall, this showed no significant differences in any microbial ecology diversity metric when the oral microbiome diversity data is analysed using the workplace zone code established in chapter 3.

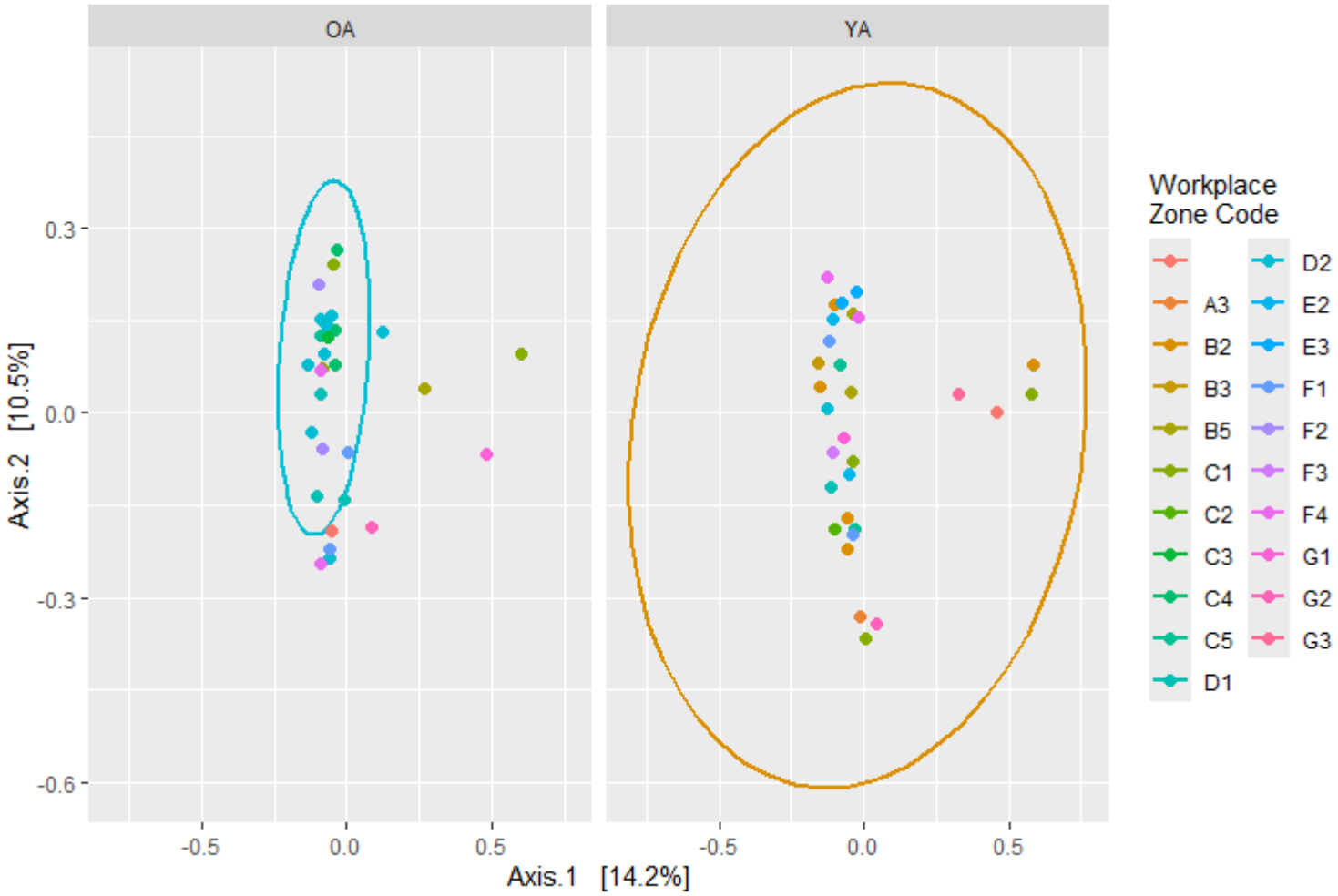


Figure 6.13 Bray Curtis dissimilarity measure of beta diversity as an ordination plot on oral samples. *Split by age group and workplace zone code.*

The AUDIT-C score is a measure relating to participant alcohol consumption, which can affect the microbiome along with having significantly different scores between the YA and OA groups (Fig. 3.6); with the OA scoring higher. This can then be further subdivided by gender where the YA-Females make up the majority of the YA low AUDIT-C scores. Using these groups on the alpha diversities showed no significant differences between any groups (Fig. 6.14). Although the OA-Males that score two were consistently similar to the OA-Females in at score three, both are lower than the other groups in Pielou evenness, Shannon and Simpson diversity. Those that score zero were more tightly clustered in their diversity scores than the other groups where some do not overlap and cover wide ranges. While possibly relevant to the study, further investigation of these

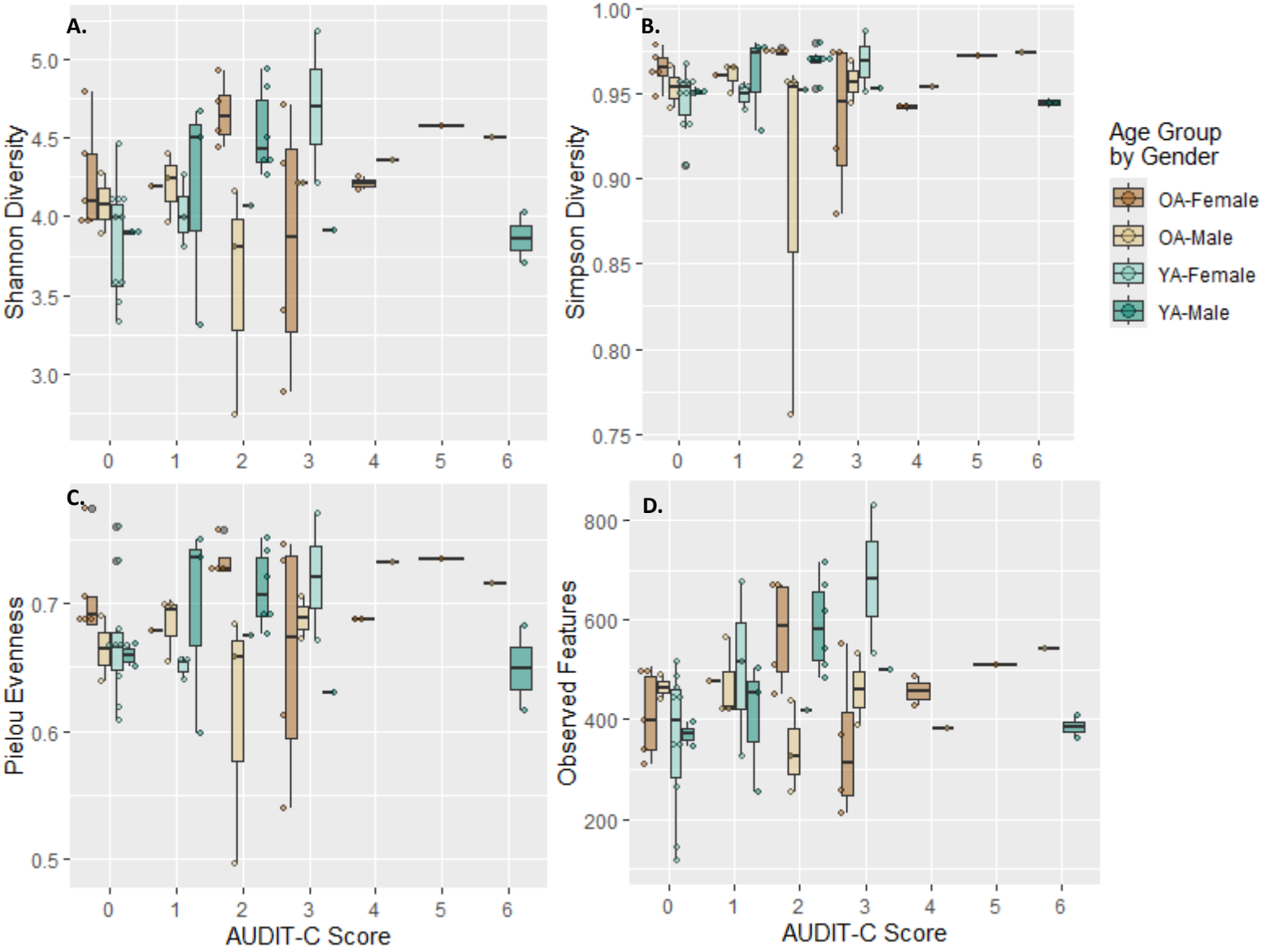


Figure 6.14 Alpha diversity metrics graphs for the oral samples, split by age group and AUDIT-C Score. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

groups and their effects on alpha diversity are beyond the scope of this study, therefore the next stage of analysing these groups is to use beta diversity.

Using these AUDIT groups on the beta diversity metrics again shows no significantly different interactions across these groups (Fig. 6.15, and Fig. 9.23). Again, there were clustered ellipses generated in the OA-Female and YA groups but not the OA-Males, these ellipses showed overlapping across most groups. Overall, this shows no significant differences in any microbial ecology diversity metric when the oral microbiome diversity data is analysed using the AUDIT-C scoring system established in chapter 3.

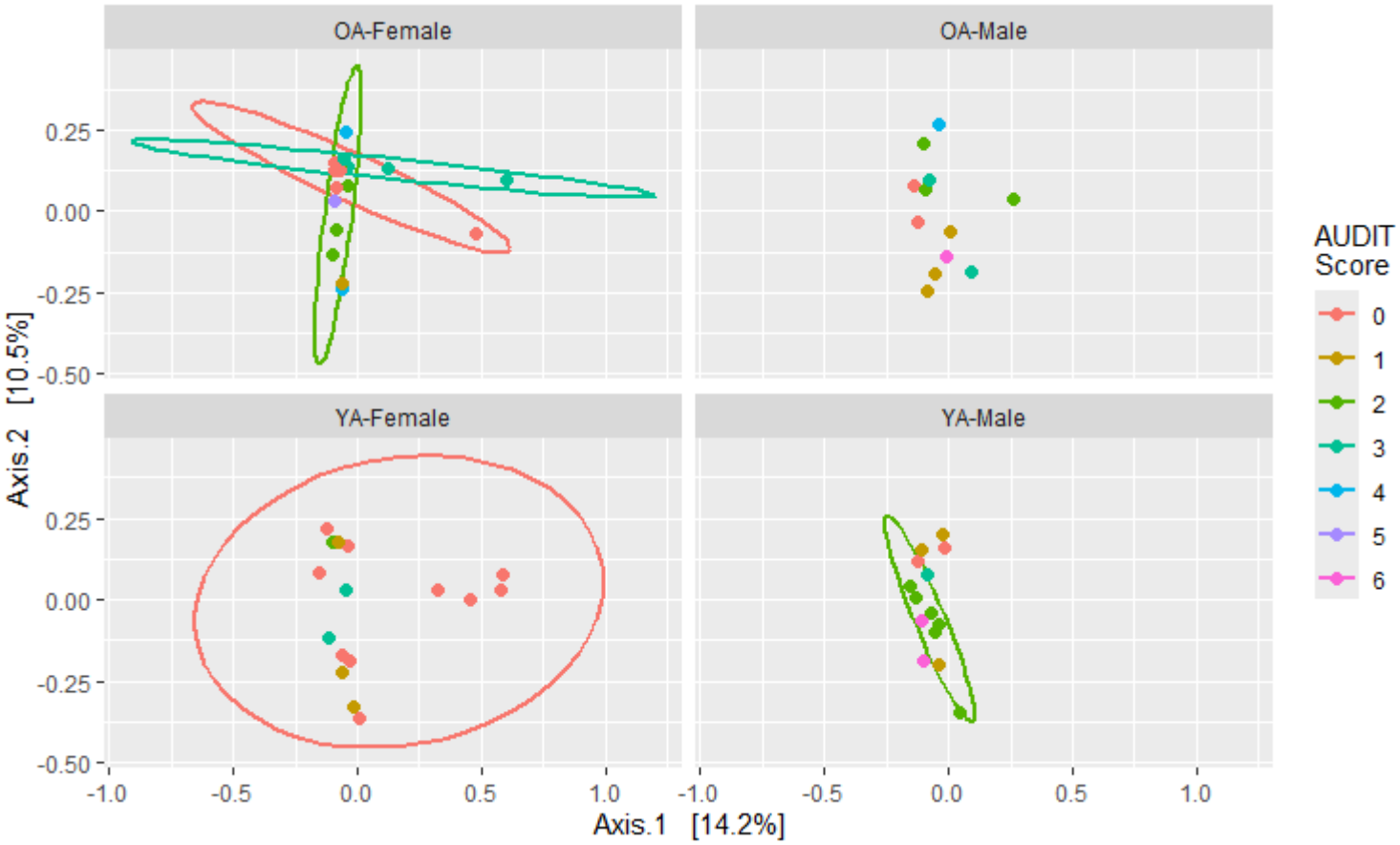


Figure 6.15 Alpha diversity metrics graphs for the oral samples, split by age group and AUDIT-C Score. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features.*

Finally with the metadata two sets of groups help to represent the participant overall dental health and investment in their own dental care, these are if the participant is missing any of their non-wisdom adult teeth and if they visit a dental hygienist regularly, respectively. Having adult teeth other than the wisdom teeth removed suggests either an increase in oral frailty or propensity to dental diseases, while a dental hygienist cleaning the teeth represents suggests that participants are aware of the value of dental health and are committed to doing the best to ensure it. Overall, both of these could have an impact on the oral microbiome.

Across the alpha diversity metrics, no significant differences were seen in these two groups, in both groups there was a consistent large overlap in the ranges of diversities covered (Fig. 6.16). When used on the beta diversity metrics no significant differences were seen in Bray Curtis dissimilarity, Jaccard index, or weighted UniFrac across both metadata groups (Fig. 6.17, Fig. 9.24, Fig. 9.25, and Table 6.3), however, a significant difference was observed in an interaction in both groups using the unweighted UniFrac metric. Specifically, in the missing teeth group there was a significant difference between OA-Yes against YA-No, and in the hygienist group between the OA-Yes against YA-No, both with a p-value of 0.012, while both groups share some participants, they also contained different participants to each other. Therefore, the only significant difference observed between these metrics used to cluster the age groups in broad ecological terms is in the presence and absence of shared species between the OA-Yes against YA-No participants.

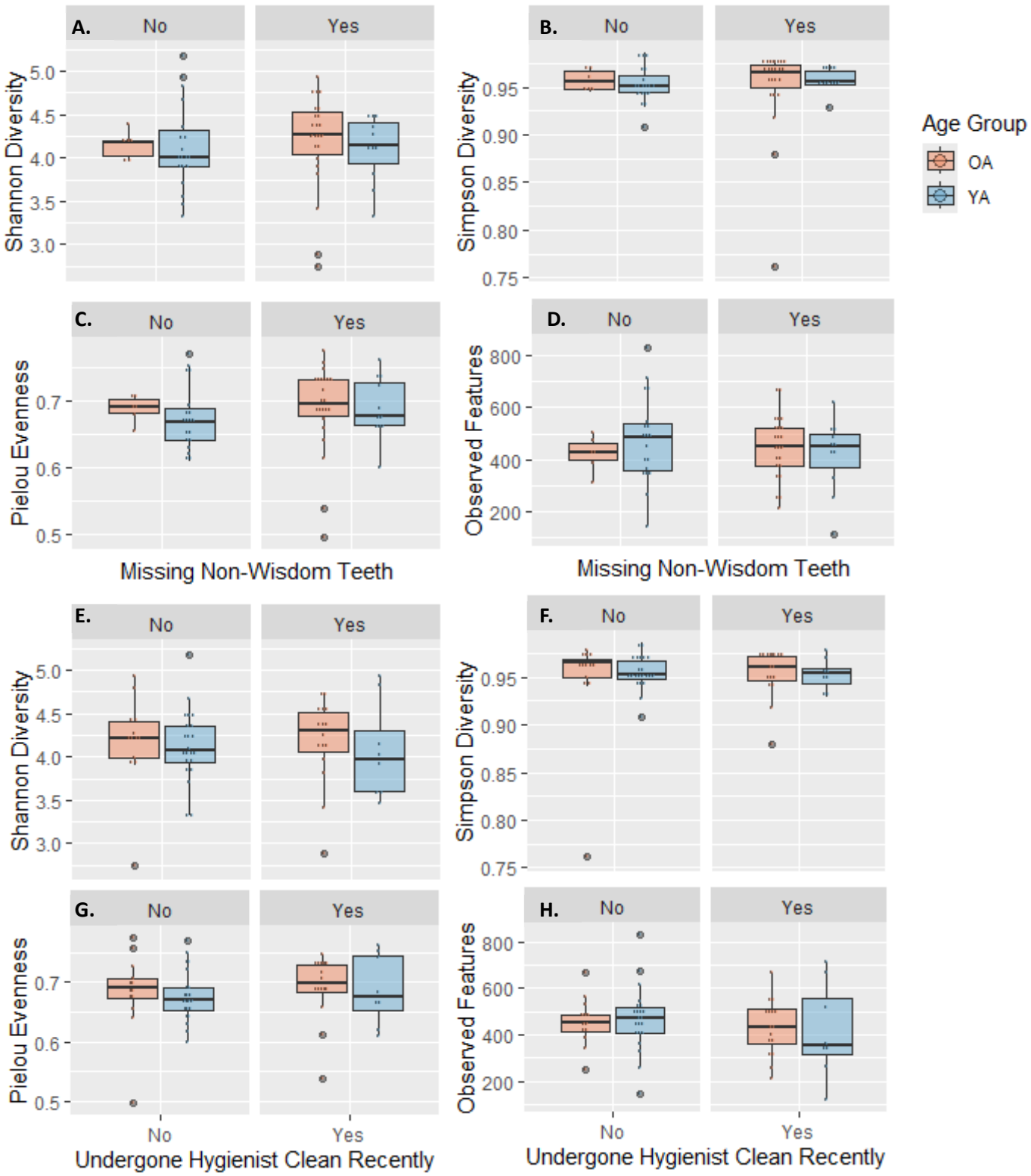


Figure 6.16 Alpha diversity metrics graphs for the oral samples, split by age group, if the participants are missing any non-wisdom adult teeth or if they had visited a dental hygienist within the 6 months prior to sampling. *Where no significant differences were observed. A. Shannon Diversity. B. Simpson Diversity. C. Pielou Evenness. D. Observed Features. E. Shannon Diversity. F. Simpson Diversity. G. Pielou Evenness. H. Observed Features.*

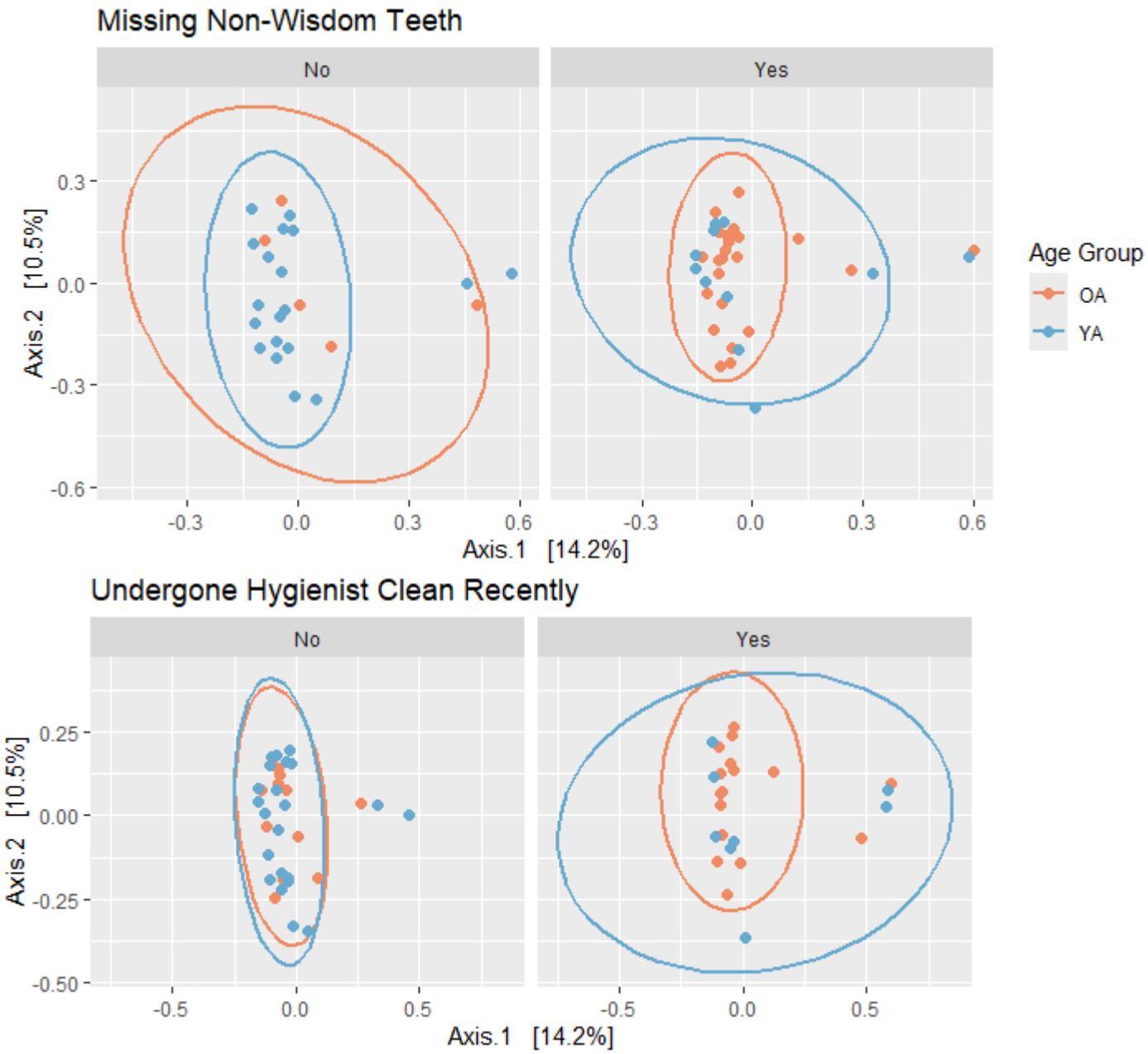


Figure 6.17 Bray Curtis dissimilarity measure of beta diversity as an ordination plot the face samples, split by age group, if the participants are missing any non-wisdom adult teeth or if they had visited a dental hygienist within the 6 months prior to sampling. *Where no significant differences were observed. A. Non-wisdom adult teeth. B. Dental hygienist within the 6 months prior to sampling.*

Table 6.3 Beta diversity p-values for oral sample interactions between the age groups sub-divided into Missing non-wisdom teeth, and Undergone hygienist clean recently. *P-values below 0.05 are highlighted in green.*

		Beta diversity measure		
Interaction	Bray Curtis dissimilarity	Jaccard index	Unweighted UniFrac	Weighted UniFrac
Missing non-wisdom teeth				
OA-Yes vs OA-No	0.432	0.324	1.000	1.000

OA-Yes vs YA-No	0.672	0.432	0.012	1.000
OA-Yes vs YA-Yes	1.000	1.000	1.000	1.000
OA-No vs YA-No	0.408	0.300	1.000	1.000
OA-No vs YA-Yes	1.000	1.000	1.000	1.000
YA-No vs YA-Yes	1.000	1.000	1.000	1.000
Undergone hygienist clean recently				
OA-Yes vs OA-No	1.000	1.000	0.498	1.000
OA-Yes vs YA-No	0.396	0.354	0.012	1.000
OA-Yes vs YA-Yes	1.000	1.000	0.342	1.000
OA-No vs YA-No	1.000	1.000	0.444	1.000
OA-No vs YA-Yes	1.000	1.000	0.060	1.000
YA-No vs YA-Yes	1.000	1.000	1.000	1.000

Across all the metadata groups used to analyse the alpha and beta diversities significantly different interactions were only observed in the unweighted UniFrac metric (Table 6.4), additionally these differences were only seen in four of the eight groups explored here. Overall, this single metric being the only significant difference observed in these groups demonstrates that presence and absence of species is a potentially key factor to this dataset, therefore differential abundant analysis is an appropriate method to try and elucidate these differences.

Table 6.4 Summary of beta diversity significant interactions for the oral samples. ***Detailing the diversity metric and p-value where appropriate.***

Grouping	Significant interaction	Diversity metric	p-value
Age Group	OA vs YA	Unweighted UniFrac	0.0014
Gender	OA-Female vs YA-Male	Unweighted UniFrac	0.012
FI	No significant differences seen in any interaction		
Super Area Output Code			
Workplace Zone Code			
AUDIT-C score			
Missing non-wisdom teeth	OA-Yes vs YA-No	Unweighted UniFrac	0.012
Undergone hygienist clean recently	OA-Yes vs YA-No	Unweighted UniFrac	0.012

6.3.3. Differential Abundance

To identify specific organisms associated with the participant groups differential abundance methods were employed, these were performed on the raw un-normalised datasets.

6.3.3.1. Effects of filtering

The datasets did require some further filtering to aid the differential abundance methods and to ensure that only frequently occurring/core organisms were included along with those that had been identified to an appropriate taxonomic rank. When filtering those taxa that occurred less than 10 times in a sample a reduction of approximately 140 taxa compared to the raw, this changes the number of species identified through the Deseq method to 10 in each age group (Table 6.5). Increasing the filtering to 50 reduced the total taxa by approximately 1500 from the raw, this only reduced the number of identified species to nine and seven in the OA and YA groups respectively. Further increasing the filtering to 100 removed 2300 taxa from the raw dataset, while reducing the identified species to seven in the OA group but increased them to 11 in the YA group. Finally using a filtering threshold of 200 reduced the total taxa to nearly 1000, while this did not reduce the identified species in the OA group from the previous filter it did reduce the YA to nine. After this comparison the filter removing taxa that were present less than 100 times in in the samples was used for the next stages of analysis as a suitable compromise due to lacking an appropriate statistically unbiased method to aid the comparison.

Table 6.5 Demonstrating the effects of various per sample filtering thresholds on the oral samples. ***Detailing the total number of taxa present and number of taxa identified within each age group using the Deseq method.***

Filter	Total taxa	Number of identified Taxa	
		OA	YA
Raw	4089	19	12
10	3932	10	10
50	2478	9	7
100	1766	7	11
200	1174	7	9

6.3.3.2. Age Groups

After filtering the raw oral dataset to remove per sample taxa that occurred less than 100 times and subsequently analysed using five different differential abundance methods, 20 individual taxa were identified (Table 6.6), only *Absconditabacteria_(SR1)_[G-1] bacterium_HMT_345* and *Cloacibacterium sp._HMT_206* were identified by all five methods used. Only one of the identified taxa were not successfully classified fully to the species level, this was *Selenomonas spp.* While each differential abundance method produced its own data on how enriched each species is within a group relative abundance plots were constructed to act as a neutral method to investigate these results.

Table 6.6 Taxa identified as differentially abundant in the oral dataset. ***Only species identified in more than three of the five abundance tests carried out on the oral dataset, when grouped based on age group.***

Species	Abundant Group	Methods Identified by				
		Ancombc	Deseq	Edger	Lefse	Limma-Voom
Absconditabacteria_(SR1)_[G-1] bacterium_HMT_345	YA	✓	✓	✓	✓	✓
Aggregatibacter aphrophilus	OA	✓			✓	✓
Alloprevotella sp._HMT_473	YA	✓	✓		✓	✓
Capnocytophaga granulosa	OA	✓			✓	✓
Cloacibacterium sp._HMT_206	YA	✓	✓	✓	✓	✓
Fusobacterium sp._HMT_203	OA	✓			✓	✓
Haemophilus sputorum	YA	✓			✓	✓
Lachnoanaerobaculum orale	OA	✓			✓	✓
Mollicutes_[G-1] bacterium_HMT_504	YA		✓	✓	✓	✓
Neisseria elongata	OA	✓			✓	✓
Neisseria oralis	OA	✓			✓	✓
Porphyromonas uenonis	YA		✓	✓	✓	✓
Prevotella dentalis	OA		✓	✓	✓	✓
Prevotella histicola	YA	✓			✓	✓
Prevotella maculosa	OA		✓	✓	✓	✓
Prevotella oulorum	OA	✓			✓	✓
Prevotella salivae	OA	✓			✓	✓
Prevotella sp._HMT_317	OA	✓			✓	✓
Selenomonas spp.	OA	✓			✓	✓
Veillonella sp._HMT_780	YA		✓	✓	✓	✓

When the 12 taxa that are abundant in the OA group were plotted using their relative abundance within the sample's a variation is seen in the proportion of abundance across these taxa (Fig. 6.18). Most of the taxa have abundances around 1% or less of the total sample. Of the abundant species *Prevotella dentalis* and *Prevotella maculosa* are only present within the OA group, whereas *Aggregatibacter aphrophilus, Capnocytophaga granulosa, Fusobacterium sp._HMT_203, Neisseria oralis, Prevotella oulorum, Prevotella sp._HMT_317*, and *Selenomonas spp*. are present in both groups but at a minimal level in the YA group. In contrast *Lachnoanaerobaculum orale, Neisseria elongata*, and *Prevotella salivae* are present in both sample groups with relatively similar abundances but a higher abundance in the OA group.

There were eight taxa identified as abundant in the YA group these show a similar variation in abundance when plotted using relative abundance (Fig. 6.19). Of these *Cloacibacterium sp._HMT_206, Mollicutes_[G-1] bacterium_HMT_504*, and *Porphyromonas uenonis* are only present within the YA group, while *Absconditabacteria_(SR1)_[G-1] bacterium_HMT_345, Alloprevotella sp._HMT_473, Prevotella histicola*, and *Veillonella sp._HMT_780* are present in both groups but in low levels in the OA group. Only *Haemophilus*

sputorum was seen in similar levels in both age groups. Overall *Prevotella* species are the most identified using these differential abundance methods. Across all these species the differences observed are genuinely within the dataset and are therefore valid for consideration in their association with the ageing oral microbiome.

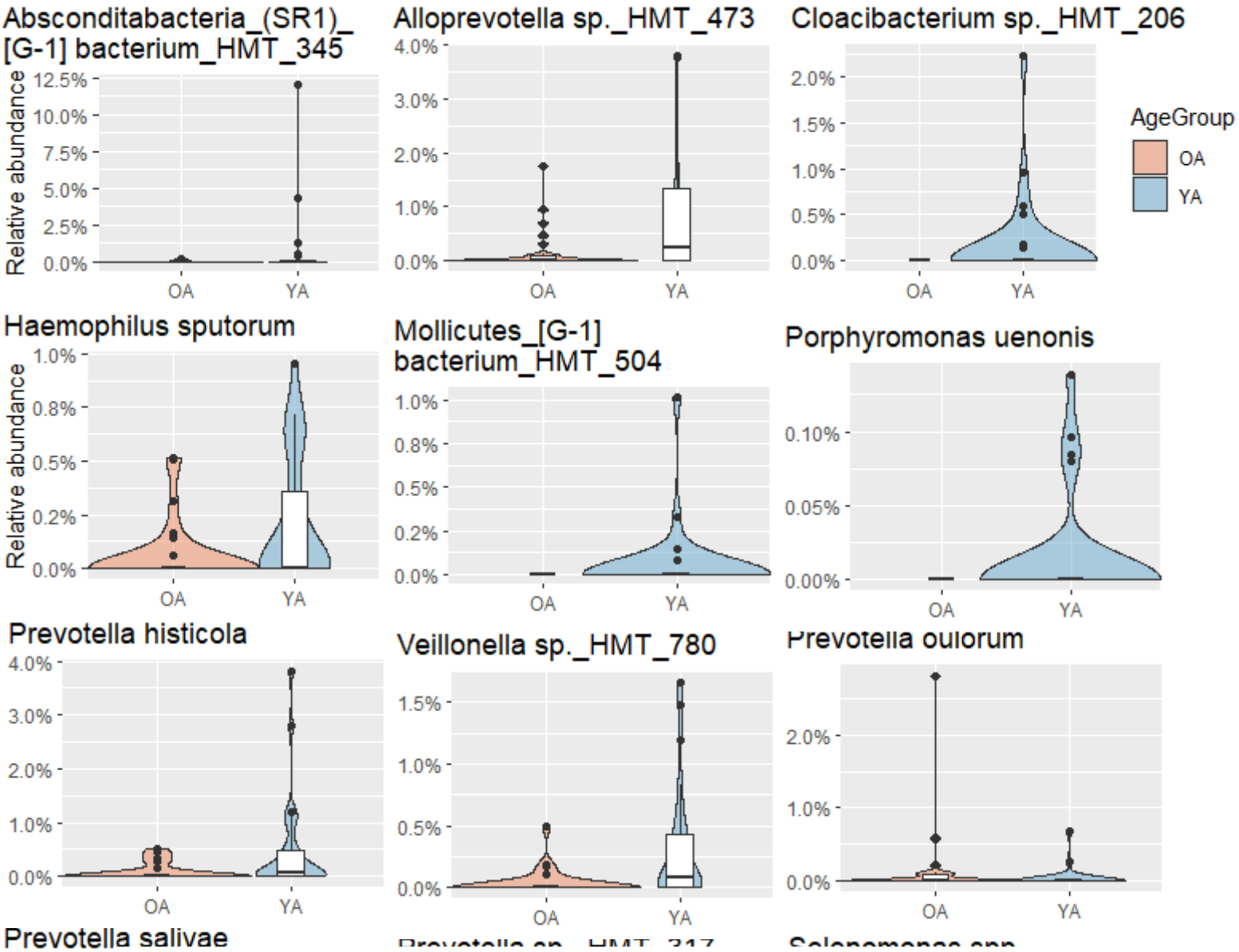


Figure 6.19 Relative abundance plots of the 8 taxa in the oral dataset identified through differential abundance as being abundant in the YA age group.

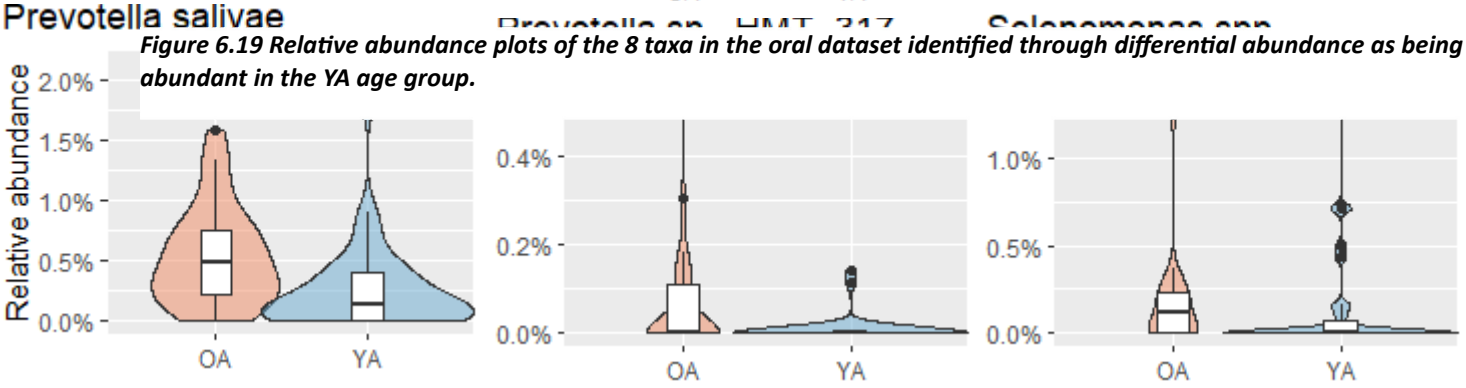


Figure 6.18 Relative abundance plots of the 12 taxa in the oral dataset identified through differential abundance as being abundant in the OA age group.

6.3.3.3. Gender

The differential abundance methods can also be applied to the age groups broken down by gender, this can then be used to follow the taxa identified in the previous age group differential abundance analysis. Both the Ancombc and Lefse methods failed to identify any taxa as differentially abundant, but all methods that did identify taxa did agree on the abundant group in each case (Table 6.7). Of the original 20 taxa only 10 were still identified in the gender groups. Of the remaining taxa *C. granulosa*, *L. orale, N. oralis*, *P. maculosa*, *P. salivae* and *Selenomonas spp.* are identified as abundant within the OA-Female group, *Absconditabacteria_(SR1)_[G-1] bacterium_HMT_345*, and *Mollicutes_[G-1] bacterium_HMT_504* are abundant in the YA-Female group, only *Veillonella sp._HMT_780* is abundant with the YA-Males and no taxa is abundant in the OA-Males. This reduction in the successful identification of these species in the gender groupings limits their ability to be used to interpret the data.

Table 6.7 Taxa previously identified as differentially abundant in the oral dataset grouped based on age group and gender. Only those species identified in more than three of the five abundance tests carried out on the oral dataset, when grouped based on age group.

Species	Abundant Group	Methods Identified by				
		Ancombc	Deseq	Edger	Lefse	Limma-Voom
Absconditabacteria_(SR1)_[G-1] bacterium_HMT_345	YA-Female		✓	✓		✓
Aggregatibacter aphrophilus	None					
Alloprevotella sp._HMT_473	None					
Capnocytophaga granulosa	OA-Female					✓
Cloacibacterium sp._HMT_206	YA-Male			✓		✓
Fusobacterium sp._HMT_203	None					
Haemophilus sputorum	None					
Lachnoanaerobaculum orale	OA-Female					✓
Mollicutes_[G-1] bacterium_HMT_504	YA-Female			✓		
Neisseria elongata	None					
Neisseria oralis	OA-Female					✓
Porphyromonas uenonis	None					
Prevotella dentalis	None					
Prevotella histicola	None					

Prevotella maculosa	OA-Female			✓		✓
Prevotella oulorum	None					
Prevotella salivae	OA-Female					✓
Prevotella sp._HMT_317	None					
Selenomonas spp.	OA-Female					✓
Veillonella sp._HMT_780	YA-Male					✓

The only species to meet the requirement of being recognised by three or more methods is *Absconditabacteria_(SR1)_[G-1] bacterium_HMT_345*, this is abundant in the YA-Female group. When plotted using relative abundance this species is shown to be present in all groups aside from the OA-Males (Fig. 6.20), while the YA-Male and OA-Female demonstrated an abundance of less than 1% of this species. Further

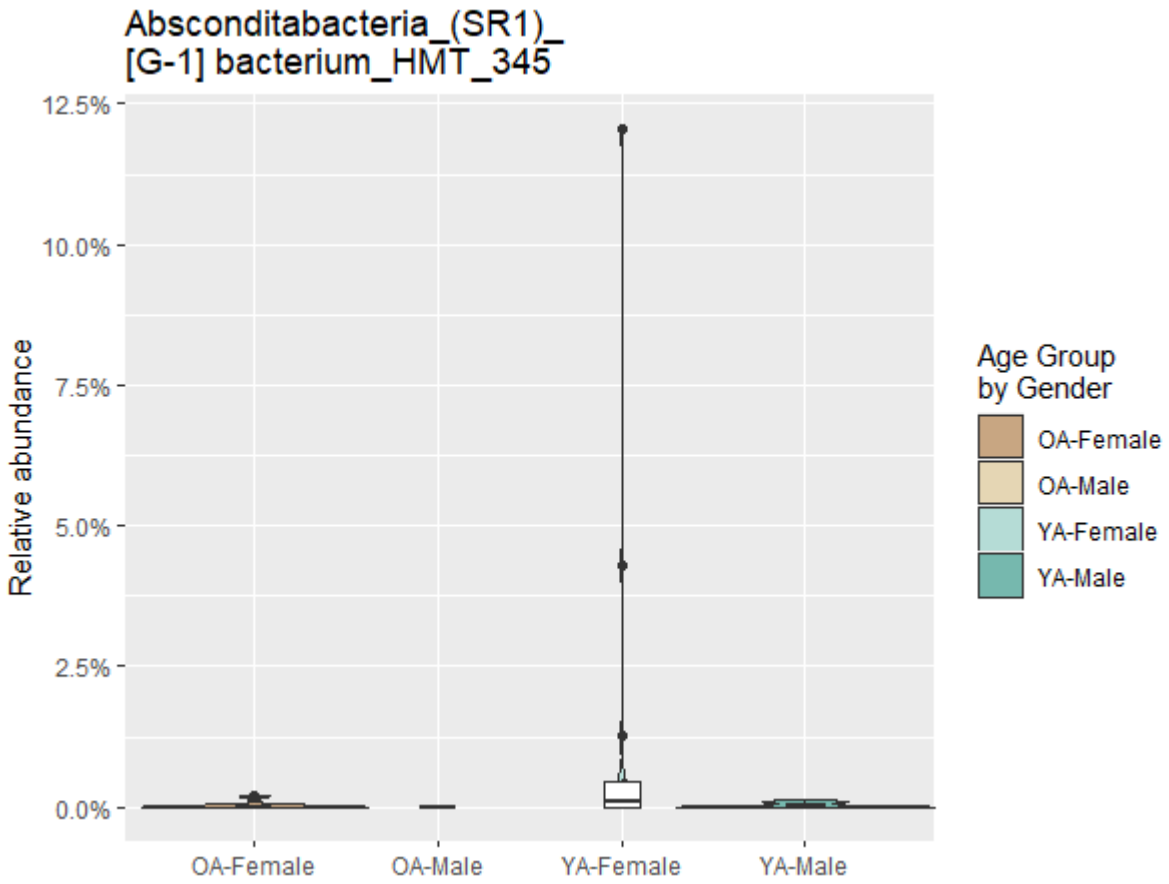


Figure 6.20 Relative abundance plot of the taxa in the oral dataset identified through differential abundance as being abundant in the age groups by gender. ***That were previously identified as abundant in the age groups alone.***

breakdowns using the gender groupings as a basis are not within the scope of this study and therefore it is more appropriate to move to a different metadata grouping.

6.3.3.4. Frailty

Finally, the differential abundance methods can also be applied to the age groups broken down by frailty index score. Here neither the Ancombc nor Deseq methods identified any taxa as differentially abundant (Table 6.8). Of the 20 taxa originally identified in the age groups only 11 were subsequently identified as differentially abundant when clustered by FI score, one of which, *Absconditabacteria_(SR1)_[G-1] bacterium_HMT_345*, was identified as abundant within two separate YA FI groups. Other than this *C. granulosa*, and *L. orale* are identified as abundant in the OA-0 group, *H. sputorum, P. dentalis, P. maculosa,* and *P. sp._HMT_317* were associated with the OA-3 group. Interestingly *H. sputorum* which was previously identified as abundant in the YA group is now associated with the OA-3 group. One species is associated with the OA-2 group, this is *P. salivae,* similarly *V. sp._HMT_780* is only associated clearly with YA-0. While the YA-1 has *C. sp._HMT_206* and *Mollicutes_[G-1] bacterium_HMT_504* identified as abundant with it. This increase in successful identification of these species compared with the previous results may enable greater interpretation of their impact on the dataset.

Table 6.8 Taxa previously identified as differentially abundant in the oral dataset grouped based on age group and FI. Only those species identified in more than three of the five abundance tests carried out on the oral dataset, when grouped based on age group.

Species	Abundant Group	Methods Identified by				
		Ancombc	Deseq	Edger	Lefse	Limma-Voom
Absconditabacteria_(SR1)_[G-1] bacterium_HMT_345	YA-0			✓		
	YA-1					✓
Aggregatibacter aphrophilus	None					
Alloprevotella sp._HMT_473	None					
Capnocytophaga granulosa	OA-0				✓	✓
Cloacibacterium sp._HMT_206	YA-1			✓		✓
Fusobacterium sp._HMT_203	None					
Haemophilus sputorum	OA-3					✓
Lachnoanaerobaculum orale	OA-0					✓
Mollicutes_[G-1] bacterium_HMT_504	YA-1			✓		
Neisseria elongata	None					
Neisseria oralis	None					
Porphyromonas uenonis	None					
Prevotella dentalis	OA-3			✓	✓	✓
Prevotella histicola	None					
Prevotella maculosa	OA-3			✓	✓	✓
Prevotella oulorum	None					
Prevotella salivae	OA-2					✓
Prevotella sp._HMT_317	OA-3			✓	✓	
Selenomonas spp.	None					
Veillonella sp._HMT_780	YA-0			✓	✓	✓

When the 3 species identified as abundant by 3 or more of the methods were plotted by their relative abundance a high abundance in *P. dentalis* and *P. maculosa* is seen in the OA-3 group (Fig. 6.21), whereas the *V. sp._HMT_780* is clearly dominant in the YA-0 group, but also having a strong presence in the YA-1 group. Across these species identified as abundant in the age groups the ability to break down their abundance into different metadata groupings enables an understanding of the significant difference in taxa seen in the ecological diversity metrics previously explored.

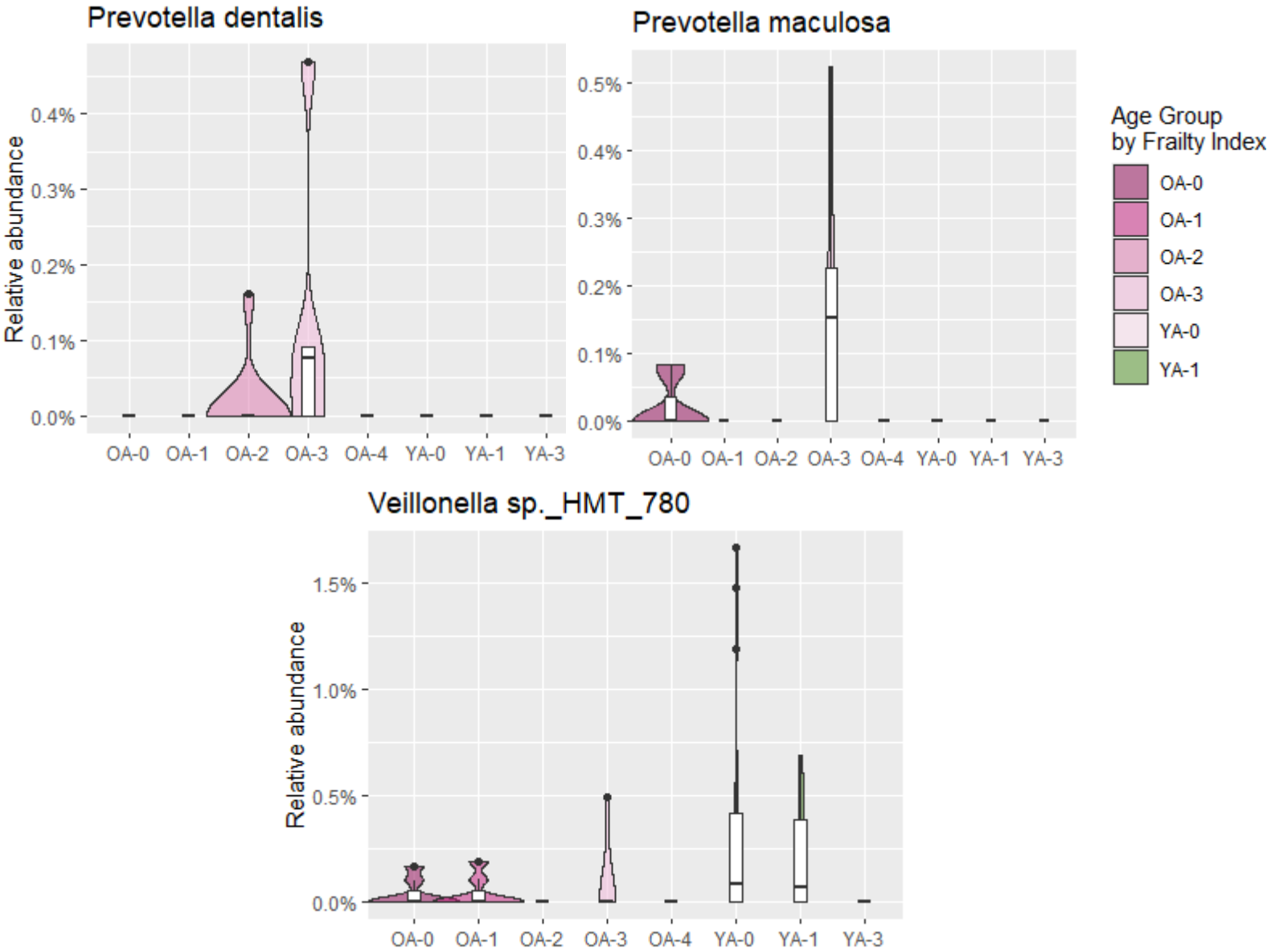


Figure 6.21 Relative abundance plots of the 3 taxa in the oral dataset identified through differential abundance as being abundant in the age groups by FI score. *That were previously identified as abundant in the age groups alone.*

6.3.4. Network analysis

Co-occurrence networks were constructed to infer and analyse potential ecological relationships within the age groups microbiomes. The datasets were filtered to remove low abundance taxa and ensure only a shared core taxa were considered for this computationally intensive analysis, additionally the networks were only compared at the age group level. After filtering a total of 208 taxa remained in the oral dataset for network construction.

These networks show 158 and 207 connected nodes for the OA and YA groups respectively, 10 of the taxa identified in the differential abundance analysis are found in the OA network, while 13 are in the YA network (Fig 6.22). When viewed in more detail we see the LCC in the OA group is 128 taxa and 205 in the YA, however the OA networks and LCC had a higher positive estimated correlation of around 74-76% whereas the YA have around 66% (Table 6.9). Two of the species identified as differentially abundant were also identified through the statistical analysis as hub taxa these are *P. maculosa* which is a hub in the OA network and

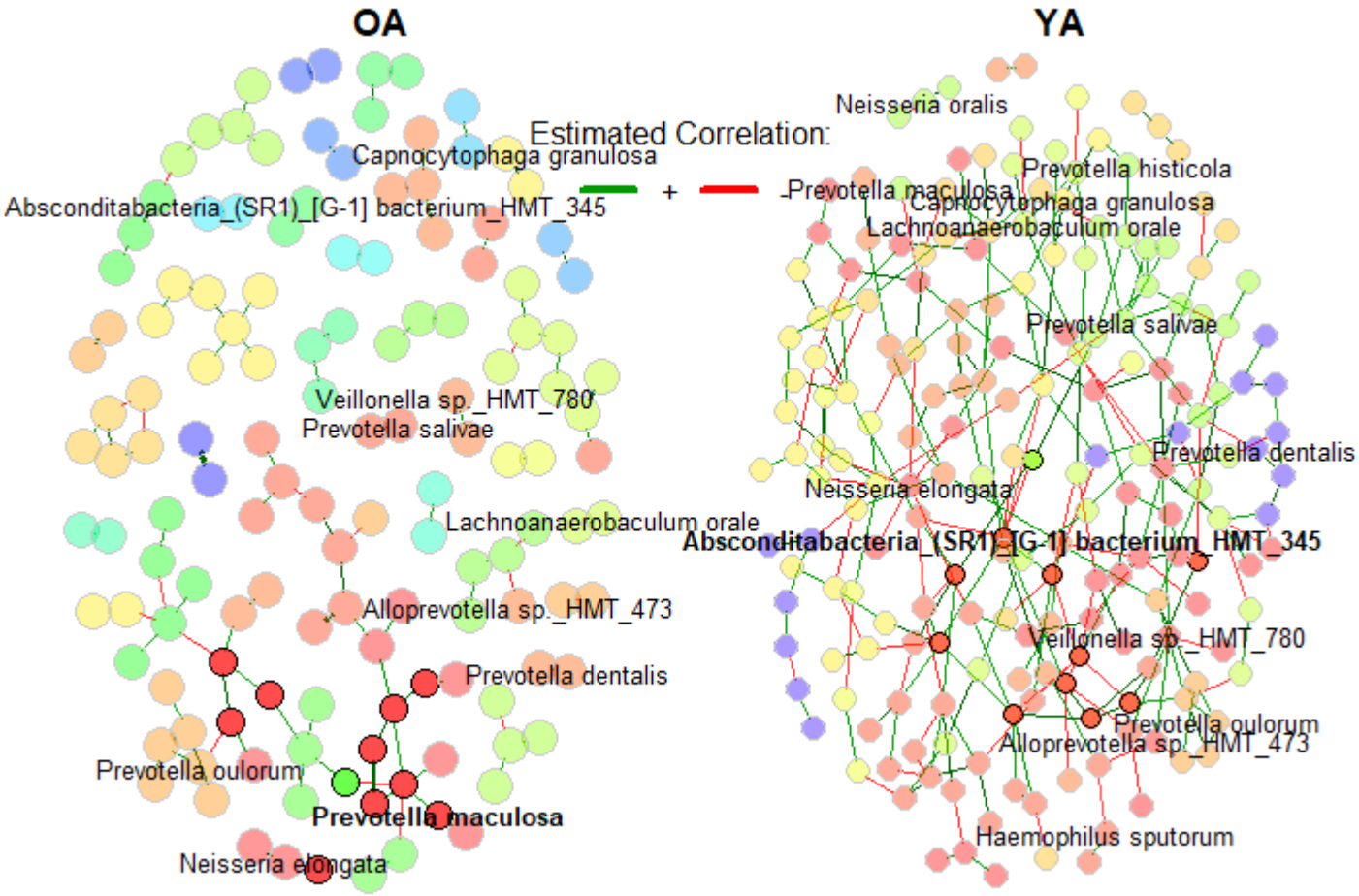


Figure 6.22 Co-Occurrence networks for the oral dataset comparing the OA and YA age groups. *The networks were constructed using SPIEC-EASI statistical method in the Netcomi R package, each node represents a unique taxa (either at the species level or there originally assigned level where speciation unavailable), within the dataset, nodes are coloured based on clustering, unconnected nodes have been removed. Edges are representative of association with nodes, split into estimated positive and negative association by the colours green and red. Nodes representing taxa identified as differentially abundant are labelled where present.*

Absconditabacteria_(SR1)_[G-1] bacterium_HMT_345 which is a hub in the YA network. In total 11 taxa were identified as hub taxa in these networks, interestingly both networks share *C. acnes* and *Selenomonas noxia* as hub organisms while they also share *Veillonella* species these are only shared at the genera level. A detailed statistical comparison of the networks reveals significant differences across the whole network in terms of edge density and natural connectivity with p-values of 0.03 and 0.03 respectively but not any other metric. While the significant differences are observed when a Jaccard index is used on the centrality measures degree, betweenness, and hub taxa, with p-values of 0.004, 0.008, and 0.02 respectively; this shows significant dissimilarity rather than significant similarity between these metrics. Overall there are differences between the age groups co-occurrence networks, this may also contribute to the other differences seen in previous sections.

Table 6.9 A summary of the descriptive information and statistical summary for the oral co-occurrence network. ***P-values below 0.05 are highlighted in green.***

Descriptive information		OA					YA		
Components	Size	128	5	3	2	1	205	2	1
	Number	1	1	1	11	50	1	1	1
Network overview	Number of components	64					3		
		Whole network			LCC		Whole network		LCC
	Clustering coefficient	0.09949			0.10288		0.09940		0.09940
	Modularity	0.78776			0.75493		0.57734		0.57601
	Positive edge percentage	76.02339			74.02597		66.42512		66.58596
	Edge density	0.00794			0.01895		0.01923		0.01975
	Natural connectivity	0.00534			0.00906		0.00589		0.00599
	Relative LCC size	NA			0.61538		NA		0.98558
	Vertex connectivity				1.000				1.000
	Edge connectivity				1.000				1.000
	Average dissimilarity				0.99416				0.99396
	Average path length				4.82311				3.03387
Hub taxa		*Bradyrhizobiaceae 2748*					*Absconditabacteria_(SR1)_[G-1] bacterium_HMT_345*		
		Cutibacterium acnes					*Campylobacter concisus*		
		Fusobacterium periodonticum					*Campylobacter rectus*		
		Mogibacterium 2603					*Cutibacterium acnes*		

	Peptostreptococcus stomatis	*Gemella 3300*
	Prevotella maculosa	*Saccharibacteria_(TM7)_[G-1] bacterium_HMT_346*
	Selenomonas 3705	*Saccharibacteria_(TM7)_[G-3] bacterium_HMT_351*
	Selenomonas noxia	*Selenomonas noxia*
	Stomatobaculum sp._HMT_097	*Staphylococcus capitis*
	Veillonellaceae_[G-1] bacterium_HMT_150	*Streptococcus parasanguinis_clade_411*
	Veillonellaceae_[G-1] bacterium_HMT_155	*Veillonella atypica*

Statistical summary		Absolute difference		p-value	
Network overview		Whole network	LCC	Whole network	LCC
	Clustering coefficient	0.000	0.003	0.999001	0.909091
	Modularity	0.210	0.179	0.142857	0.090909
	Positive edge percentage	9.598	7.440	0.127872	0.200799
	Edge density	0.011	0.001	0.026973	0.618382
	Natural connectivity	0.001	0.003	0.028971	0.188811
	Relative LCC size	NA	0.370	NA	0.170829
	Vertex connectivity		0.000		1.000
	Edge connectivity		0.000		1.000
	Average dissimilarity		0.000		0.673327
	Average path length		1.789		0.153846

Jaccard index	Measure	Jacc	P(<=Jacc)	P(>=Jacc)
	degree	0.190	0.003577	0.998442
	betweenness centr.	0.209	0.008210	0.995969
	closeness centr.	0.300	0.307256	0.771836
	eigenvec. centr.	0.268	0.127837	0.916213
	hub taxa	0.100	0.017593	0.996692

6.3.5. Predictive metabolomics

The raw datasets were used in the Qiime2 Picrust2 plugin to create the predicted metabolomic pathway data. This was normalised through conversion into relative abundance data before broad analysis using beta diversity. In this oral dataset there was no significant difference between the age groups using either Bray Curtis or Jaccard index, this sees overlapping in the statistically generated clusters in the ordination plots of these metrics (Fig. 6.23 and Fig. 9.26). These results do not prevent further analysis using differential abundance methods from being valid, as in the main dataset analysis.

These data were also analysed using the same differential abundance methods as used previously (6.3.3.2). These identified a total of 22 different pathways identified as abundant in three or more of the various methods, however only five of these were identified by all five methods (Table 9.6). When these five abundant pathways were plotted using relative abundance it demonstrates that they constitute a relatively small proportion of the overall number of pathways as they all are present in all samples at levels lower than 0.3% (Fig. 6.24), these are abundant between the OA and YA groups. The superpathway of arginine and polyamine biosynthesis, superpathway of polyamine biosynthesis, and GDP-D-*glycero*-α-D-*manno*-heptose biosynthesis were all abundant in the OA group, as shown by the increased distribution of the pathways when plotted, while the fatty acid β-oxidation and glutaryl-CoA degradation pathways that were abundant in the YA groups invert this trend.

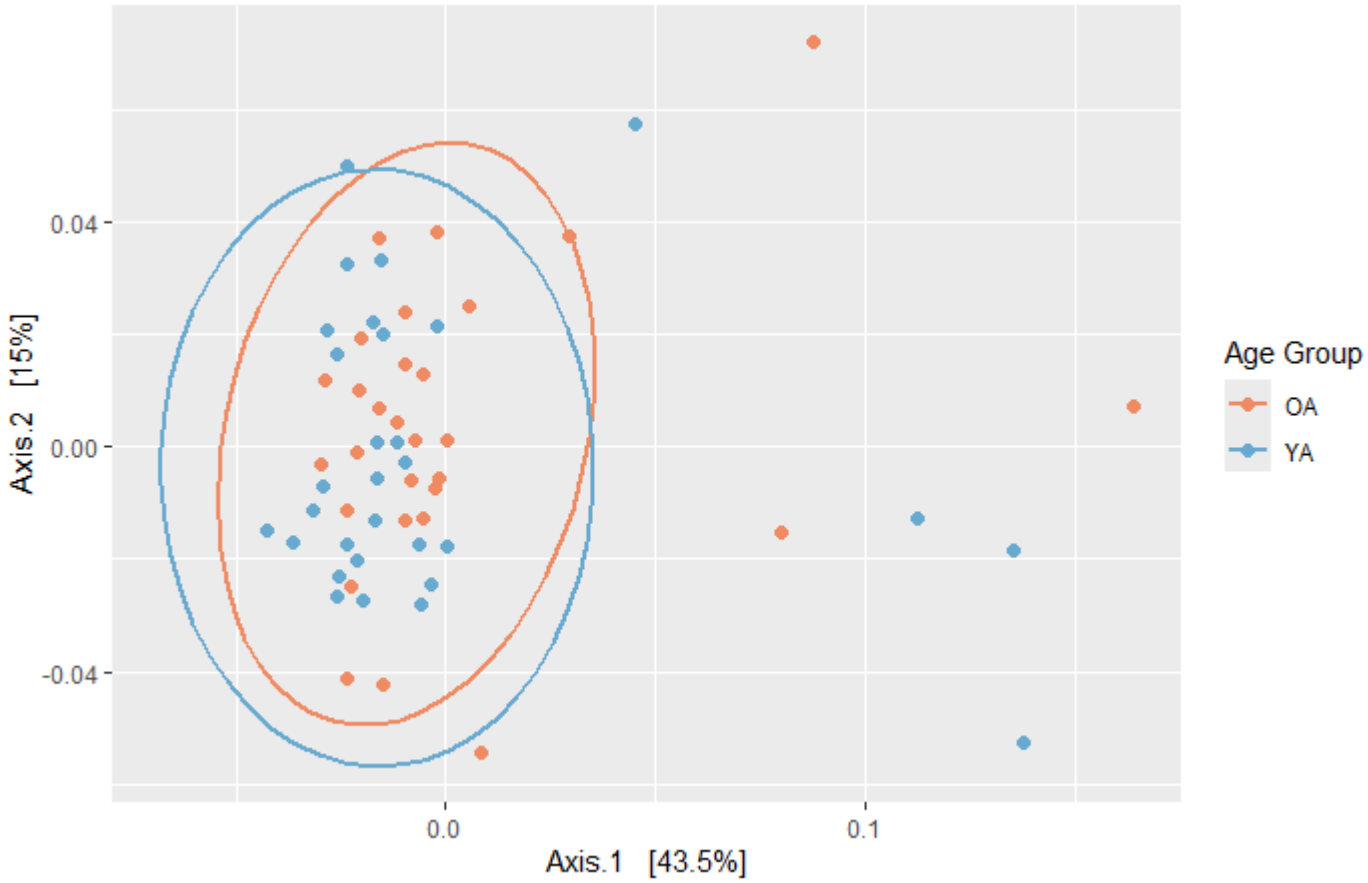


Figure 6.23 Bray Curtis dissimilarity measure of beta diversity as an ordination plot of the predicted metabolomic pathways of the oral dataset grouped by age group. *Where no significant difference was observed.*

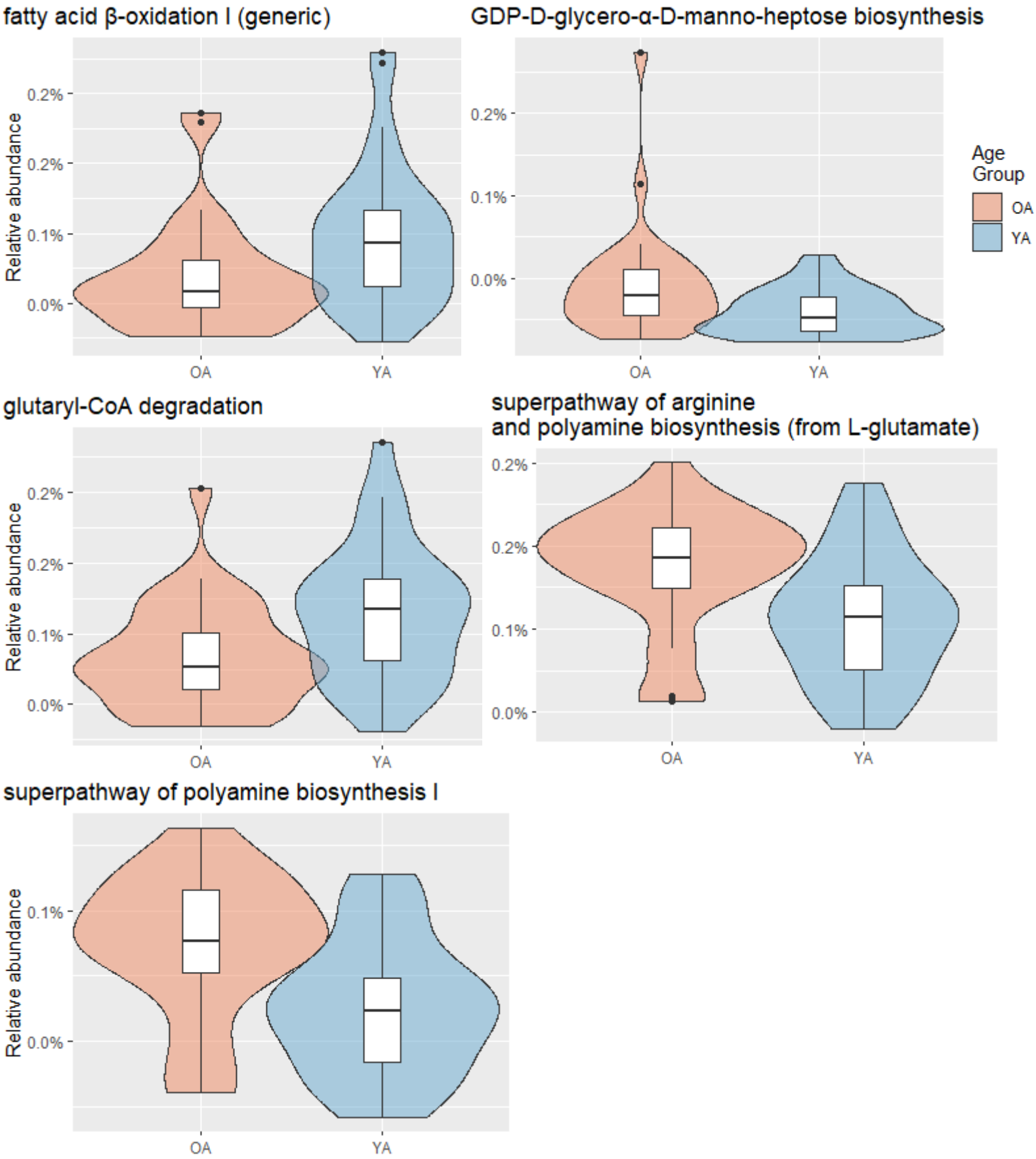


Figure 6.24 Relative abundance plots of the 5 metabolic pathways identified as abundant in the oral dataset.

6.4. Discussion

The focus of this chapter has been the characterisation of the oral microbiome using the dataset previously finalised in chapter 4 with a focus on the age groups the samples were collected from. As this is one of the main themes of this thesis it required a considered and comprehensive analysis covering microbial ecology, mathematical modelling, and post taxonomic toolkits to enable detailed characterisation and comparison. Additionally previous analysis of the participant metadata conducted in chapter 3 identified other groups within the participant panel using some of these to breakdown and compare the data to better understand any differences or similarities seen between the age groups.

The top 10 species identified across all the samples are all associated with the human oral cavity microbiome. However, the very low levels of these species present in most of the samples demonstrates the diverse content of the oral microbiome (6.3.2). In the OA samples *S. salivarius* is seen to have a more consistent dominant presence than the YA, whereas the YA have more *Veillonella* and *H. parainfluenzae* suggesting a taxonomic shift covered by multiple species in age. The dominance of *Streptococci* and *Veillonellae* in the oral cavity have been well established with various roles ascribed to them (Zhou *et al.*, 2021; McLean *et al.*, 2022). Sample OA24 has nearly three quarters of its content taken up by these top 10 species, this low content along with the other low samples at the left side of the graphs may be explained by the low concentration of DNA present in the extract sent for sequencing, this however did not return low numbers of sequence reads as all samples returned over 100000 reads. Without more samples from these participants, it is impossible to know if the differences seen in these low DNA samples compared to the other samples are due to this poor DNA yield, where basis in extraction method and sequence recovery have been established (Abusleme *et al.*, 2014), or are representative of their normal oral microbiome content; therefore the samples are still valid for their inclusion into the study. Overall, this broad overview does demonstrate some differences between the age groups oral microbiomes, and therefore it warrants further statistical analysis to understand these differences better.

The alpha diversity measures showed no significant differences between the age groups, showing continuous overlap between them (Fig. 6.3), which is supported when the Shannon diversity is plotted by participant age (Fig. 6.4), showing a broad spread across the various ages. Given the high overlaps shown it is doubtful that any differences in pipeline processing or additional sampling would show different results. As alpha diversity is an unlabelled metric it does not enable a comparison of the taxonomic differences that were being shown previously, but this does suggest that in terms of total bacterial content and broad species presence any oral microbiome changes are more nuanced than can be detected through these metrics. This is supported when the beta diversity metrics are considered where the only significant difference between the age groups is observed in the unweighted UniFrac metric (Fig. 9.18), which is a measure of presence and absence of the taxa across the samples, whereas the other metrics also factor abundance into the weighting. Therefore, it is possible

that some species, particularly the lower presence members are displaced during ageing but no overall differences in abundance occur in the oral microbiome. Additionally other research which looked at multiple age groups showed that oral microbiome diversity is in a U-shaped curve over adulthood (Kazarina *et al.*, 2023), this demonstrates that the lack of any difference in these metrics observed here is normal and a result of the study design.

The gender-based groups (6.3.3.2), similarly do not show any significant differences at the alpha diversity level. Interestingly however, a trend in the female groups can be seen (Fig. 6.6), where the YA-Females are lower than the OA-Females, there may be a factor due to these participants gender such as hormonal changes or other habit differences that could be studied in further depth either demonstrating or disproving any differences between females and age. Again, the only significant difference seen in beta diversity is in the unweighted UniFrac metric (Table 6.2), here it is only seen in the OA-Female compared to YA-Male. This is interesting as previously gendered differences in this thesis have been observed but mainly focused within the same gender. Again, this difference being focused on the unweighted UniFrac suggests a difference between low level species between this group. In other non-UK based studies differences in the oral microbiome associated with gender have been reported (Minty *et al.*, 2021; Liu *et al.*, 2022; Pinto *et al.*, 2024), additionally however these studies have focused on oral disease states or to investigate specific participant factors such as food consumption. Therefore, while no differences are observed here in these broad diversity metrics is may be that there are differences associated with age in a UK population but various other factors such as participant hormonal levels, feeding regime, or oral disease state which have not been controlled for in this study prevent them from being accurately elucidated.

Clustering the diversities based on participant FI score showed no significant differences in either alpha or beta diversity metrics (6.3.3.3), the alpha diversity graphs only showed overlaps with no specific trends. This is interesting as a previous study (DeClercq *et al.*, 2024), has associated frailty with changes in the oral microbiome diversity in ageing. It may be that there are limitations within this study that prevent these differences from becoming established such as the FI scoring being limited in the measurements assessed and the small number of participants. Similarly there may be a more direct link between oral frailty and the oral microbiome a factor which was not studied here but has been linked in another recent study (Yang *et al.*, 2025). Overall, this study has not been designed to investigate frailty and the oral microbiome, further work would need to be carried out to consider if these factors are linked in a UK population.

Looking at the alpha and beta diversities through socioeconomic groupings showed no significant differences in the oral microbiome diversities (6.3.3.4), while some trends may be observed in the alpha diversity graphs no solid conclusions can be drawn from them. This is likely due a limitation in this study of its small number of participants, if a future study recruited more participants from a variety of socioeconomic groups it

may be possible to observe differences similar to those already established in the gut microbiome and socioeconomic status (Bowyer *et al.*, 2019; Kwak *et al.*, 2024). Similarly using the AUDIT scoring system to break down the alpha and beta diversities showed no significant differences but did demonstrate some trends such as a tight clustering in the no alcohol participants (Fig. 6.14). This lack of differences again may be due to the lack of adequate participants recruited based on their alcohol consumption, as other studies have been able to link alcohol consumption and effects on the oral microbiome (Fan *et al.*, 2018; Ward *et al.*, 2023; Maley *et al.*, 2024).

The other metadata groups investigated here included those who were missing non-wisdom teeth and those who had undergone hygienist cleans recently, both of which reported a significant difference only in unweighted UniFrac and only between OA-Yes vs YA-No groups. As already established, this suggests a difference between low level species between these groups. Regarding the missing teeth group, it is possible that this difference is caused by removal of the teeth causing a change in the environments within the oral cavity allowing different species to become established, some studies have already implicated the loss of teeth with changes in the oral microbiome and even possible effects on health (Takeshita *et al.*, 2016; Nishimoto *et al.*, 2023). Similarly, the difference observed in the hygienist cleans group is possibly due to the yes group having made a commitment to an effective dental hygiene regime that the no group have not, therefore effecting the inhabitation from specific species, however without more information about the participants oral hygiene and health such conclusions are highly speculative. It is possible that there is an overlap in the participants present in each metrics groups which is why the difference is there, as such these factors require further study to investigate if there is any basis in them and how or if they are linked to ageing rather than just being an artifact of this studies population.

Through differential abundance testing some 20 species were identified as being associated with the age groups (6.3.4.2). Of those identified *A. sp_HMT_473*, *C. sp._HMT_206, F. sp._HMT_203*, *M. bacterium_HMT_504*, *P. sp._HMT_317*, and *V. sp._HMT_780* are all known only from genomic studies limiting the ability to interpret their presence in these results (Mark Welch *et al.*, 2016; Welch *et al.*, 2016; Andrew Voorhis *et al.*, 2023), however the genus information provided does allow for some hypothesising to be conducted. Additionally, *Absconditabacteria_(SR1)_[G-1] bacterium_HMT_345* has only been given candidate status as a genus and is yet to be fully recognised outright preventing any full understanding of its presence here (Oren and Göker, 2023). The *Cloacibacterium* genus has species that are associated with environmental water sampling (Allen *et al.*, 2006; Chun *et al.*, 2017), suggesting that it is not a relevant species for this study. Aside from these two organisms the rest of the species (*A. aphrophilus, C. granulosa, H. sputorum, L. orale, N. elongata, N. oralis, P. uenonis, P. dentalis, P. histicola P. maculosa P. oulorum, and P. salivae*) and their associated genera found through differential abundance are all associated with the human microbiome in some capacity (Dworkin *et al.*, 2006; Nørskov-Lauritsen and Kilian, 2006; Rubin, 2008; Hedberg *et al.*, 2012, 2012; Guilloux *et al.*, 2021; Tett *et al.*, 2021; Giacomini *et al.*, 2024, 2025), these are also sometimes associated with infections (Brenner *et*

al., 1989; Acuña-Amador and Barloy-Hubler, 2020). This causes difficulty in understanding the reasons behind their role in oral microbiome ageing in this thesis, as research will be skewed towards their infectious potential and treatment, coupled with a disparity in the literature towards certain organisms (Callaway, 2025) and knowledge of species interactions being limited (Caufield *et al.*, 2017; Zhang *et al.*, 2020; Palmer and Foster, 2022) beyond broad terms, indirect inferences, and functional predictions; whereas the skin has a handful of species dominating the microbiome meaning the results in chapter 5 identified organisms that are implicated in numerous diseases and are well studied with established interactions in this environment.

From what is known about those identified in the OA group (Fig. 6.18), are mostly associated with periodontal infections (Idate *et al.*, 2020; Sharma *et al.*, 2022), however in addition to this *A. aphrophilus* is associated with infective endocarditis (Chambers *et al.*, 2013), *C. granulosa* has been implicated in diabetes mellitus (Ciantar *et al.*, 2005), and *L. orale* has been associated with oral cancer and bacteraemia (Ida *et al.*, 2022; Yang *et al.*, 2022). Otherwise the *Neisseria* species are associated with host nitrate metabolism and may be beneficial to some members of the OA group, while only occasionally being linked to disease sates (Wong and Janda, 1992; Kahler, 2021). Comparatively the YA organisms identified are less associated with periodontal infections, however *H. sputorum* has been linked with infective endocarditis and has begun to appear in genital infections (Costello *et al.*, 2022; Franco-Acosta *et al.*, 2022), *P. uenonis* is understudied in the oral cavity but has been shown to be disruptive in the vaginal microbiome (Acuña-Amador and Barloy-Hubler, 2020; Lithgow *et al.*, 2022), and *P. histicola* while possibly contributing towards inflammation has been shown to reduce arthritis symptoms in mice models (Marietta *et al.*, 2016). Overall, these differences in the differentially abundant organisms suggest that the OA are at greater risk of oral diseases compared to the YA population who carry organisms less related to oral disease including some potentially transient species. The increase in the *Neisseria* species is interesting when considered through the lens of its nitrate reduction role and compared to the change in *Veillonella* compared to the YA (both in the context of the differentially identified species and in the broad overview of taxonomy), as *Veillonella* is also associated with a nitrate reduction role, suggesting that during ageing the changes in the oral microbiome population cause other species to take up community roles from ones that have a reduced presence, of course this may involve more species than those discussed here.

Frustratingly despite the gender age groups showing a significant difference in taxa at the beta diversity level only the discarded *Absconditabacteria* was identified through three or more methods (Table 6.8). This demonstrates that the gender of the participants has a limited effect on the differentially abundant organisms associated with ageing, suggesting that the difference observed in beta diversity was due to other factors and thus effects other species. Comparatively using the FI scoring did identify three species from the age groups as differentially abundant in the FI groups, these are two *Prevotella* species and a *Veillonella* (Fig. 6.21). As *Prevotella* species are linked to inflammation and the abundance is associated with the OA-3 group it suggests

that these participants may be experiencing higher levels of oral inflammation than the other OA groups, however this is a speculative conclusion limited by the applicability of the FI. Although inflammation is an established factor in older adults oral cavities (Bäck *et al.*, 2007), this has also been linked to *Prevotella* in other studies (Ebersole *et al.*, 2021; Könönen *et al.*, 2022). Overall, this difference between differentially abundant organism return from the gender and FI groups suggests that in the oral microbiome health/frailty is a more important factor than gender in ageing.

Moving away from specific species roles in the microbiome and into the overall community structure required the construction of co-occurrence networks for each age group. The YA groups network has a larger LCC which is consists of most of the taxa used to construct the networks, this results in a single large network which is not significantly larger than the OA LCC and whole network (6.3.5); while the YA network does contain more negatively estimated correlated connections this is not significantly so. A significant dissimilarity between the networks was seen when reviewed using a Jaccard index, however this is limited to three out of the five areas measured. One of these areas was the hub taxa which sees two shared species and other shared genera, suggesting that despite being different in other areas some key organisms are retained regardless of age. However, these hub taxa do contain organisms that are not fully classified to the genus or species level or are listed as candidate organisms limiting the conclusions that can be drawn about them. When looking at the species identified through differential abundance and their position in the network more species are present in the YA network than the OA network, including two organisms as hub taxa. This also sees these species disconnected and less likely to co-occur, while the YA sees these linked, in the OA this includes the supportive organisms of *N. elongata* and *V. sp_HMT_780* whereas two of the *Prevotella* species remain linked and are therefore more likely to occur and interact with the microbial community. Overall, these oral networks suggest that the YA have a more connected and established community which has more balanced members than the OA, which may make the OA oral microbiomes more susceptible to colonisation of atypical species and pose an increased threat to oral health, which has been established by other more detailed work into older adults and oral health in Japan and China (Asakawa *et al.*, 2018; Jiang *et al.*, 2019).

Using the post taxonomic predictions of metabolomic pathways allows for the investigation of how the microbiomes may be functioning in the host and identify any differences between the age groups, this is important to demonstrate if the identified population differences are impacting ageing or if the changes are the result of the host age. Here while no beta diversity differences were seen in the pathway abundances, differential abundance analysis identified five pathways between both age groups (6.3.6). Of these five pathways the two associated with the YA, fatty acid β-oxidation I and glutaryl-CoA degradation, are linked to microbial metabolism and growth (Koch *et al.*, 1993; Campbell, Morgan-Kiss and Cronan, 2003; Schiaffi, Barras and Bouveret, 2024), although fatty acids have been considered a biomarker for oral cancers (Praveen *et al.*, 2024), the increased abundance of two metabolic pathways may just be from the diverse diets of the YA population providing different

molecules for breakdown or because of a more integrated community which share processing of biomolecules. The OA abundant pathways include two similar functions the superpathway of arginine and polyamine biosynthesis (from L-glutamate) and superpathway of polyamine biosynthesis I are both involved in polyamine synthesis, increases in polyamines and arginine in the oral cavity have been associated with cancer and infections in older adult populations (Denis *et al.*, 2023; Chu, Chan and Chu, 2024; Bessa *et al.*, 2025). Finally the GDP-D-glycero-α-D-manno-heptose biosynthesis pathway is involved with the production of a key part of the lipopolysaccharide (LPS) system (Holst *et al.*, 1991; Melaugh *et al.*, 1992; Kneidinger *et al.*, 2001), which is naturally associated with increased pathogenicity (Wilson *et al.*, 2002). Looked at together these pathways show that the OA population has a microbiome more prone to creating molecules implicated in oral diseases and increased pathogenic factors whereas the YA are more inclined to have an increased bacterial metabolism, these findings support the previous results from the co-occurrence networks and differential abundant analysis.

During the construction of this chapter investigating the oral microbiome in ageing some limitations have become apparent. The initial culture-based screening even though it has successfully identified some fungal species present in the oral cavity their data was not possible to be included in the 16S dataset preventing their influence on the oral microbiome from being investigated, this could be resolved using whole genome sequencing in future studies, additionally the organisms isolated in this cultural analysis still exist in cold storage and could be recovered for further detailed genetic or cultural study in their own analysis. The nature of the samples taken themselves does limit this study by preventing comparison of the microbiomes at different sites in the oral cavity, this is key as the biogeography of the oral cavity supports different bacterial niches which would enable better conclusions about the changes in the oral microbiome in age by allowing for the locations of these changes to be clarified (Proctor *et al.*, 2020; Giordano-Kelhoffer *et al.*, 2022). Broadly the whole analysis is limited by only having one sample set from each participant, this prevents the creation of averaged microbiomes for the participants which would allow for the confident elimination of transient contaminant species that could be creating noise in the diversity statistics and in the differential abundance results, similarly dynamic shifts in the oral microbiome over time intervals have been observed suggesting long periods between samplings ensure a core microbiome is seen (Jiang *et al.*, 2015). While the classifier used here has more confidence than the skin classifier as it has been created using a curated database, there are still limitations for example the majority of *Neisseria* species in the dataset are only classified to the genus level limiting the conclusions possible about this genera, this may possibly be improved by tweaking the classifier, or the dataset, or obtaining a higher resolution on the sequences by using longer reads.

Over the course of this chapter, it has become apparent that the health of the participants is a key role in the oral microbiome, however, only a rough guide of their health was collected using self-reporting, in the future a more detailed health status should be gathered ideally including an oral health and frailty check carried out by a dentist. Similarly, more participants across all age ranges would enable the investigation into the changes

in the oral microbiome over ageing totally rather than just at the specific age ranges used here. Additionally while the various additional metadata breakdowns explored here did not show any differences some trends were shown and in other studies these metrics have shown differences between the various groups and their microbiomes (Fan *et al.*, 2018; Bowyer *et al.*, 2019), therefore this study is limited in its ability to analyse how these may relate to the ageing oral microbiome, in future studies the number of participants recruited should be increased or the age groups alone should not be the sole factor for recruitment. Finally, some of the difficulties experienced in this chapter have revolved around decisions to do with the selection of data filtering and statistical methodologies, in the future this could be made easier by seeking advice from more experienced bioinformaticians or statisticians as required.

6.5. Conclusion

Overall, this chapter analysed the oral datasets created previously in terms of diversity, mathematical modelling, and predictive metabolism, where possible this was also broken down not just in terms of the age groups but in terms of participant gender, FI, and other potentially relevant metadata subgroups. Through this it demonstrates differences in the oral microbiome during ageing is highly nuanced as there are limited significant differences in the broad microbial diversity measures. Instead, the driving differences occur at the edges of the microbiome, where differential abundance identified that an increase in various *Prevotella* species is associated with the increasing age and frailty, suggesting that they are more vulnerable to oral infections and inflammation. This is reinforced through the co-occurrence networks which showed that the YA are more likely to have a resilient microbiome than the OA, along with beneficial species identified through differential abundance in the OA group being less well connected than in the YA. Finally, analysis of the predictive metabolomics adds to this picture by showing an increase in various pathways associated with pathogenic behaviours in the OA group, therefore this provides a basis for future studies to focus on these differences and vulnerabilities in the oral cavity in ageing.

Chapter 7. General discussion

The understanding of the role played by the oral and skin microbiome in ageing is only just starting to be developed. Previous studies into the microbiome in ageing have focused on the role played by the gut microbiome (Santoro *et al.*, 2020), while a valuable endeavour that contributes valid findings to the field, the human microbiome exists in more than just this one niche and it is therefore necessary to investigate the other niches for any role they may play in ageing. Similarly, studies into the ageing microbiome are usually conducted in North American or East Asian countries (Hung, Kempen and Vries, 2010; Reich *et al.*, 2020; Abdill, Adamowicz and Blekhman, 2022), while studies in these countries provide valid results, their findings may not be applicable to other countries due to differences in standards of healthcare, socioeconomic situations, and other factors which can modify the ageing experience or human microbiome. Therefore, performing this study into the skin and oral microbiome in a UK based population with age being the main participant factor will help close this knowledge gap.

Initially a sampling protocol had to be established, as there are various options for sampling the skin and oral microbiome (Fig. 1.2). This protocol needed to be practical to perform and effective in its recovery of microbes. Using viable count methods chapter 2, identified that oral rinsing had a significantly better recovery of bacteria than oral swabbing, and while no significant differences were observed between the skin swabbing and scraping methods trailed, swabbing was deemed more practical. Without this work the understanding of the effects of sampling as a limitation in the thesis would not be as thorough and the organisation of the sampling on the participants would have been less well thought out. Overall, this enabled the recruitment for the main study of this thesis to commence and allowed for professional practice to be developed before interacting with members of the public.

Having recruited the participants and obtained the samples from them it was also necessary to establish what was different between the populations beyond simple age, as any differences may have an influence on the microbiome, additionally this ensured that the age groups were sufficiently internally similar to enable valid conclusions in the main analysis. In chapter 3, the age groups were found to have a significant difference in terms of frailty (Fig. 3.5), with the OA group being frailer than the YA, additionally, this also found a significant increase in alcohol consumption among the OA group (Fig. 3.6). Further established differences included socioeconomic factors, along with some specific skin and dental health metrics. The discovery of these sub-groups enabled some of them to be used (where relevant/feasible), in the analysis of the microbiome. Additionally, such an in-depth understanding of population data aids in understanding the study limitations and identifies areas which require consideration by future studies.

While other 16S microbiome studies have been carried out at Aston University previously none had sequenced these specific variable regions on skin and oral sample types, therefore the assembly of the 16S datasets for analysis required deep consideration. During the variations in pipeline trialled in chapter 4, the skin sequences were shown to have problems with Phred quality compared to the oral, which lead to the removal of one sample from the arm dataset. Again, these in-depth pipeline variations aided in developing an understanding of the study limitations while successfully creating optimal datasets for the sample types.

These finalised skin datasets were then analysed in terms of microbial ecology, mathematical modelling, and predictive metabolomics using the previously generated metadata analysis to breakdown and guide the interpretations where appropriate. Here significant differences were observed in beta diversity mainly between the age groups and the YA-Female vs OA-Female group (Table 5.6), in both the arm and face samples. Through differential abundance the species *C. acnes* was identified to undergo significant decreases in the arm and face samples during ageing, this was also shown to be significantly abundant in the arm samples of YA-Females (5.3.3.2). Statistically constructed co-occurrence networks showed significant dissimilarity between the age groups in both samples, in the face network the YA showed a high level of complexity indicating resilience to disruption by other species (5.3.4). Finally predictive metabolomics showed changes to the face microbiome behaviour in age (5.3.5). These results establish that the differences in the skin microbiome during age revolve around a dominant species, specifically in the arm microbiome; gender also plays a large role in some of these changes.

Comparatively in the oral dataset analysis using the same methods and similar groupings showed limited significantly different beta diversity results (Table 6.5), these were only shown in the unweighted UniFrac metric (which compares the presence of taxa without considering abundance) suggesting a stable core microbiome. Through differential abundance the OA group was shown to have an increase in various *Prevotella* species compared to the YA, this was not shown to significantly be affected by gender, instead significant abundance was seen when clustered by frailty index score in the OA group (6.3.3). Using the co-occurrence networks limited significant dissimilarity was shown, but the YA still showed a high level of complexity indicating resilience to disruption by other species (6.3.4). Finally predictive metabolomics showed an increase in pathogenic associated pathways in the OA group (6.3.5). These establish that the differences in the oral microbiome during age revolve around small changes in low abundance species and that these are driven by health along with age.

Taking these findings from the skin and oral microbiomes together, these results demonstrate that within the same participant different changes in the ageing process may be driving various changes to the microbiome. As the hormonal changes possibly responsible for the gendered shift in skin microbiome are not having the same effect in the oral cavity, instead other changes such as immune system presence may be allowing the increase in pathogenic species seen. Similarly, the face and oral co-occurrence networks in the YA are more resilient than in

the arm, suggesting that even for this group care should be taken to not disrupt the delicate balance in the arm microbiome. This could not have been seen without the gathering and analysis of participant metadata along with having three sampling locations from each participant, this demonstrates the value of a multi-faceted study in investigating the microbiome.

Despite these valuable insights some limitations have persisted throughout the study. While a suitable/acceptable sampling method was identified in the initial phases of this study, the skin sampling method has provided problems with low DNA yield and poor sequence quality. This has led to the removal of some participants from the skin microbiome analysis and may have impacted the ability to accurately classify the species present, although other recent research has suggested that swabbing is the best method currently available and that intra-participant variation is a key factor (Alyami *et al.*, 2025; Balacco *et al.*, 2025), but other research finds variation in the microbiome in deeper layers of skin tissue (Pardo *et al.*, 2024), showing that this is a complex topic with impacts beyond the scope of this thesis. Therefore, further studies could be done into the sampling techniques for the skin microbiome, at the very least more factors can now be considered that were not considered at the time of this study commencing due to the research being released after sampling had taken place. This study has been limited by its scale, while it is sufficiently powered, other studies into the microbiome into ageing have used more participants, as this enables a greater analysis by having more ages and participants with a variety of backgrounds that may affect the microbiome. Additionally, only one set of samples per participant does affect the ability to factor in intra-participant variation in the microbiome while also correcting for low quality sequences, a future study could be done comparing the variability over time in ageing populations, which could provide additional support to the interpretation of the co-occurrence networks. The overall understanding of these results is also impacted by a lack of clinical assessment of the participants skin and oral health, for example the increased inflammation associated organisms seen in the OA group could simply be due to enough of them having some mild gum disease or similar (Martínez-García and Hernández-Lemus, 2021). Although this would not necessarily detract from the conclusion that OA are more at risk of oral disease by their microbiome, it may be possible to use these as a biomarker for oral ageing health, which is why determining the oral health of participants is important. Aside from participant factors the sequencing method used also has influenced the study, only sequencing a select variable region of the 16S gene limits the classification depth along with also limiting the study focus to bacteria. Sequencing longer regions can allow for deeper classification including to the strain level, while whole genome sequencing would open the study up to fungi and other microbes, recent studies and reviews have called for multiple sequencing methods to be employed in microbiome studies to allow for the various methods to complement each other (Rubiola *et al.*, 2022; Filardo, Di Pietro and Sessa, 2024). Beyond this some of the statistical analysis of the microbiome datasets used here are considered controversial by some (Goberna and Verdú, 2022; Nearing *et al.*, 2022; Jung, 2025). While more could be done in future studies to address these concerns the techniques used are still considered

valid and are improving over time, so such concerns do little to dent the validity of this study but need to be considered for future work. Overall, these limitations demonstrate an understanding of the multifaceted nature of this and similar studies into the human microbiome, and while attempts to control these have been taken where possible here their impact is understood to affect other studies of this nature.

Various aspects of this study can contribute to future studies into the skin and oral microbiome in ageing, principally this is a main outcome of this study as its scale and scope makes it effectively a pilot study on a UK population. From the metadata gathered in this study significant differences and trends between the age groups have been identified beyond simply age, these have included aspects of health, habits, and hygiene. In future studies these factors may either be recruited for to enable distinct populations within the study, or otherwise controlled to aid the interpretation of the results in the microbiome. While the results from the skin analysis have identified several key species and metabolic pathways that warrant further investigation along with identifying gender as a driving factor, these may either become the focus of a future study or become an aspect that is compared between the studies. Similarly, the oral aspect has identified different species that are not dominant members of the microbiome that change during ageing along with metabolic pathways that could be further investigated along with identifying whether health is the driving factor, again these may be the focus of a future study or an aspect for comparison. Regardless of what a future study utilises from these findings hopefully it will include a wider range of participants and ages to enable a deeper understanding of the skin and oral microbiomes in ageing to be developed. Overall, this study has broadly found similarities to other studies into the skin and oral microbiome in ageing in other countries including the USA, China, Japan, and Korea, but has also found possible aspects which are distinctive to a UK population, including differences in species abundance and microbiome metabolome.

In conclusion, this thesis has performed a characterisation of the skin and oral microbiome in ageing in a UK based population, demonstrating differences in these microbiomes and that these differences are driven by different factors. It identified an optimal skin and oral sampling method suitable for implementation at Aston University or other non-clinical settings. Analysis of the participant metadata showed differences in age of a broadly healthy UK population, while identifying other differences that warrant further investigation. The skin characterisation demonstrated that a reduction in the key species *C. acnes* occurs in both the arm and face during ageing, while also demonstrating how the microbiome becomes vulnerable to disruption by other species in age, significantly in the skin microbiomes striking differences were observed in females in age whereas the male populations tended to be similar. Comparatively the oral characterisation demonstrated an increase in in various *Prevotella* species associated with increasing age and an increase in the pathogenic potential of the microbiome metabolome, here however the main driver appears to be overall health and frailty rather than gender. Together this demonstrates how different aspects of ageing can affect different locations in the body differently and therefore, shows how a holistic view of the ageing process in humans and the microbiome must be considered

to enable the development of monitoring systems or treatments to help manage the progress of ageing, where specific species could be used as biomarkers or become targets for treatments to alleviate the ageing process. Finally, this not only adds knowledge about the ageing skin and oral microbiome in a UK population to the already known changes in other countries populations, but its status as a pilot study means that it has identified other areas that future studies can use to help guide their design and implementation. The principal conclusion from this thesis is that there are differences in the skin and oral microbiome in an ageing UK population, and that these differences are not the same in each area and are also driven by different factors.

Chapter 8. References

Aas, J.A. *et al.* (2005) 'Defining the Normal Bacterial Flora of the Oral Cavity', *Journal of Clinical Microbiology*, 43(11), p. 5721. Available at: https://doi.org/10.1128/JCM.43.11.5721-5732.2005.

Abat, C., Garcia, V. and Rolain, J.-M. (2016) '*Facklamia hominis* scapula abscess, Marseille, France', *New Microbes and New Infections*, 9, pp. 13–14. Available at: https://doi.org/10.1016/j.nmni.2015.11.003.

Abdill, R.J., Adamowicz, E.M. and Blekhman, R. (2022) 'Public human microbiome data are dominated by highly developed countries', *PLoS Biology*, 20(2), p. e3001536. Available at: https://doi.org/10.1371/journal.pbio.3001536.

Abusleme, L. *et al.* (2014) 'Influence of DNA extraction on oral microbial profiles obtained via 16S rRNA gene sequencing', *Journal of Oral Microbiology*, 6(1), p. 23990. Available at: https://doi.org/10.3402/jom.v6.23990.

Acuña-Amador, L. and Barloy-Hubler, F. (2020) '*Porphyromonas* spp. have an extensive host range in ill and healthy individuals and an unexpected environmental distribution: A systematic review and meta-analysis', *Anaerobe*, 66, p. 102280. Available at: https://doi.org/10.1016/j.anaerobe.2020.102280.

Adachi, M. *et al.* (2007) 'Professional oral health care by dental hygienists reduced respiratory infections in elderly persons requiring nursing care', *International Journal of Dental Hygiene*, 5(2), pp. 69–74. Available at: https://doi.org/10.1111/j.1601-5037.2007.00233.x.

Adami, G.R. (ed.) (2021) *The Oral Microbiome*. New York, NY: Springer US (Methods in Molecular Biology). Available at: https://doi.org/10.1007/978-1-0716-1518-8.

Agnihotri, R. and Gaur, S. (2014) 'Implications of tobacco smoking on the oral health of older adults', *Geriatrics & Gerontology International*, 14(3), pp. 526–540. Available at: https://doi.org/10.1111/GGI.12285.

Ahlers, J.M.D. *et al.* (2022) 'Single-Cell RNA Profiling of Human Skin Reveals Age-Related Loss of Dermal Sheath Cells and Their Contribution to a Juvenile Phenotype', *Frontiers in Genetics*, 12(January), pp. 1–10. Available at: https://doi.org/10.3389/fgene.2021.797747.

Akers, L. and Gordon, J.S. (2018) 'Using Facebook for Large-Scale Online Randomized Clinical Trial Recruitment: Effective Advertising Strategies', *Journal of Medical Internet Research*, 20(11), p. e9372. Available at: https://doi.org/10.2196/jmir.9372.

Alkema, W. *et al.* (2021) 'Charting host-microbe co-metabolism in skin aging and application to metagenomics data', *PLOS ONE*, 16(11), p. e0258960. Available at: https://doi.org/10.1371/JOURNAL.PONE.0258960.

Allen, T.D. *et al.* (2006) 'Cloacibacterium normanense gen. nov., sp. nov., a novel bacterium in the family Flavobacteriaceae isolated from municipal wastewater', *International Journal of Systematic and Evolutionary Microbiology*, 56(Pt 6), pp. 1311–1316. Available at: https://doi.org/10.1099/ijs.0.64218-0.

Allender, S., Hutchinson, L. and Foster, C. (2008) 'Life-change events and participation in physical activity: a systematic review', *Health Promotion International*, 23(2), pp. 160–172. Available at: https://doi.org/10.1093/heapro/dan012.

Almeida-Santos, A. *et al.* (2021) 'Characterization of the Oral Microbiome of Medicated Type-2 Diabetes Patients', *Frontiers in Microbiology*, 12, p. 56. Available at: https://doi.org/10.3389/FMICB.2021.610370/BIBTEX.

Alyami, R.Y. *et al.* (2025) 'Methods for the analysis of skin microbiomes: a comparison of sampling processes and 16S rRNA hypervariable regions.', *Access Microbiology* [Preprint]. Available at: https://doi.org/10.1099/acmi.0.001037.v1.

Amieva, M.R. and El-Omar, E.M. (2008) 'Host-Bacterial Interactions in Helicobacter pylori Infection', *Gastroenterology*, 134(1), pp. 306–323. Available at: https://doi.org/10.1053/J.GASTRO.2007.11.009.

Aminu, S. *et al.* (2024) 'Exploring microbial worlds: a review of whole genome sequencing and its application in characterizing the microbial communities', *Critical Reviews in Microbiology*, 50(5), pp. 805–829. Available at: https://doi.org/10.1080/1040841X.2023.2282447.

Andrew Voorhis *et al.* (2023) *HOMD :: Human Oral Microbiome Database*. Available at: https://www.homd.org/ (Accessed: 5 March 2025).

Asakawa, M. *et al.* (2018) 'Tongue Microbiota and Oral Health Status in Community-Dwelling Elderly Adults', *mSphere*, 3(4), p. 10.1128/msphere.00332-18. Available at: https://doi.org/10.1128/msphere.00332-18.

Bäck, M. *et al.* (2007) 'The Oral Cavity and Age: A Site of Chronic Inflammation?', *PLOS ONE*, 2(12), p. e1351. Available at: https://doi.org/10.1371/journal.pone.0001351.

Bai, X. (2014) 'Images of Ageing in Society: A Literature Review', *Journal of Population Ageing*, 7(3), pp. 231–253. Available at: https://doi.org/10.1007/s12062-014-9103-x.

Baker, L.B. (2019) 'Physiology of sweat gland function: The roles of sweating and sweat composition in human health', *Temperature*, 6(3), pp. 211–259. Available at: https://doi.org/10.1080/23328940.2019.1632145.

Balacco, D.L. *et al.* (2025) 'Optimisation of cutaneous microbiota sampling methodology', *Frontiers in Microbiomes*, 4, p. 1559981. Available at: https://doi.org/10.3389/frmbi.2025.1559981.

Bana, B. and Cabreiro, F. (2019) 'The Microbiome and Aging', *Annual Review of Genetics*, 53(1), pp. 239–261. Available at: https://doi.org/10.1146/annurev-genet-112618-043650.

Barer, B.M. (1994) 'Men and Women Aging Differently', *The International Journal of Aging and Human Development*, 38(1), pp. 29–40. Available at: https://doi.org/10.2190/MCH0-VTER-DL6G-LECR.

Bars-Cortina, D. *et al.* (2024) 'Comparison between 16S rRNA and shotgun sequencing in colorectal cancer, advanced colorectal lesions, and healthy human gut microbiota', *BMC Genomics*, 25(1), pp. 1–20. Available at: https://doi.org/10.1186/s12864-024-10621-7.

Behera, S.S. *et al.* (2020) 'Traditionally fermented pickles: How the microbial diversity associated with their nutritional and health benefits?', *Journal of Functional Foods*, 70, p. 103971. Available at: https://doi.org/10.1016/j.jff.2020.103971.

Behr, L.C. *et al.* (2023) '60 years of healthy aging: On definitions, biomarkers, scores and challenges', *Ageing Research Reviews*, 88, p. 101934. Available at: https://doi.org/10.1016/j.arr.2023.101934.

Bélanger, M., Townsend, N. and Foster, C. (2011) 'Age-related differences in physical activity profiles of English adults', *Preventive Medicine*, 52(3), pp. 247–249. Available at: https://doi.org/10.1016/j.ypmed.2011.02.008.

Belotti, L. and Frazão, P. (2022) 'Effectiveness of water fluoridation in an upper-middle-income country: A systematic review and meta-analysis', *International Journal of Paediatric Dentistry*, 32(4), pp. 503–513. Available at: https://doi.org/10.1111/ipd.12928.

Benito-Cruz, B. *et al.* (2016) 'Oral *Candida* isolates and fluconazole susceptibility patterns in older Mexican women', *Archives of Gerontology and Geriatrics*, 65, pp. 204–210. Available at: https://doi.org/10.1016/j.archger.2016.04.001.

Berkowitz, D.M. *et al.* (1974) 'Adhesive Tape: Potential Source of Nosocomial Bacteria', *Applied Microbiology*, 28(4), p. 651. Available at: https://doi.org/10.1128/AM.28.4.651-654.1974.

Berthoud, R., Blekesaune, M. and Hancock, R. (2009) 'Ageing, income and living standards: evidence from the British Household Panel Survey', *Ageing & Society*, 29(7), pp. 1105–1122. Available at: https://doi.org/10.1017/S0144686X09008605.

Bessa, L.J. *et al.* (2025) 'Linking peri-implantitis to microbiome changes in affected implants, healthy implants, and saliva: a cross-sectional pilot study', *Frontiers in Cellular and Infection Microbiology*, 15, p. 1543100. Available at: https://doi.org/10.3389/fcimb.2025.1543100.

Best, A., Lieberman, D.E. and Kamilar, J.M. (2019) 'Diversity and evolution of human eccrine sweat gland density', *Journal of Thermal Biology*, 84, pp. 331–338. Available at: https://doi.org/10.1016/j.jtherbio.2019.07.024.

Biesterbos, J.W.H. *et al.* (2013) 'Usage patterns of personal care products: Important factors for exposure assessment', *Food and Chemical Toxicology*, 55, pp. 8–17. Available at: https://doi.org/10.1016/j.fct.2012.11.014.

Bjerre, R.D. *et al.* (2019) 'Effects of sampling strategy and DNA extraction on human skin microbiome investigations', *Scientific Reports*, 9(1), pp. 1–11. Available at: https://doi.org/10.1038/s41598-019-53599-z.

Bockmühl, D.P. (2017) 'Laundry hygiene-how to get more than clean', *Journal of Applied Microbiology*, 122(5), pp. 1124–1133. Available at: https://doi.org/10.1111/jam.13402.

Bokulich, N.A. *et al.* (2016) 'mockrobiota: a Public Resource for Microbiome Bioinformatics Benchmarking', *mSystems*, 1(5), pp. e00062-16. Available at: https://doi.org/10.1128/mSystems.00062-16.

Bokulich, N.A. *et al.* (2018) 'Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin', *Microbiome*, 6(1), pp. 1–17. Available at: https://doi.org/10.1186/s40168-018-0470-z.

Bolyen, E. *et al.* (2019) 'Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2', *Nature Biotechnology*, 37(8), pp. 852–857. Available at: https://doi.org/10.1038/s41587-019-0209-9.

Boulesnane, Y. *et al.* (2020) 'Impact of sampling and DNA extraction methods on skin microbiota assessment', *Journal of Microbiological Methods*, 171(December 2019), p. 105880. Available at: https://doi.org/10.1016/j.mimet.2020.105880.

Bouslimani, A. *et al.* (2019) 'The impact of skin care products on skin chemistry and microbiome dynamics', *BMC Biology*, 17(1), p. 47. Available at: https://doi.org/10.1186/s12915-019-0660-6.

Bowyer, R.C.E. *et al.* (2019) 'Socioeconomic Status and the Gut Microbiome: A TwinsUK Cohort Study', *Microorganisms*, 7(1), p. 17. Available at: https://doi.org/10.3390/microorganisms7010017.

Brandwein, M. *et al.* (2021) 'Biogeographical Landscape of the Human Face Skin Microbiome Viewed in High Definition', *Acta Dermato-Venereologica*, 101(11), pp. adv00603–adv00603. Available at: https://doi.org/10.2340/00015555-3929.

Brenner, D.J. *et al.* (1989) 'Capnocytophaga canimorsus sp. nov. (formerly CDC group DF-2), a cause of septicemia following dog bite, and C. cynodegmi sp. nov., a cause of localized wound infection following dog bite', *Journal of Clinical Microbiology*, 27(2), pp. 231–235. Available at: https://doi.org/10.1128/jcm.27.2.231-235.1989.

Brito, S., Baek, M. and Bin, B.-H. (2024) 'Skin Structure, Physiology, and Pathology in Topical and Transdermal Drug Delivery', *Pharmaceutics*, 16(11), p. 1403. Available at: https://doi.org/10.3390/pharmaceutics16111403.

Brookes, Z.L.S. *et al.* (2021) 'Effects of chlorhexidine mouthwash on the oral microbiome', *Journal of Dentistry*, 113(July). Available at: https://doi.org/10.1016/j.jdent.2021.103768.

Brooks, S.G. *et al.* (2025) 'The Skin Acid Mantle: An Update on Skin pH', *Journal of Investigative Dermatology*, 145(3), pp. 509–521. Available at: https://doi.org/10.1016/j.jid.2024.07.009.

Brown, M., Thomas, B. and Blake, K. (2018) 'Oral Health and Ageing: A Literature Review'. Available at: https://doi.org/10.7727/wimj.2018.164.

Burke-Garcia, A. and Mathew, S. (2017) 'Leveraging social and digital media for participant recruitment: A review of methods from the Bayley Short Form Formative Study', *Journal of Clinical and Translational Science*, 1(3), pp. 205–207. Available at: https://doi.org/10.1017/cts.2017.9.

Byrd, A.L., Belkaid, Y. and Segre, J.A. (2018) 'The human skin microbiome', *Nature Reviews Microbiology*, 16(3), pp. 143–155. Available at: https://doi.org/10.1038/nrmicro.2017.157.

Callahan, B.J. *et al.* (2016) 'DADA2: High resolution sample inference from Illumina amplicon data', *Nature methods*, 13(7), pp. 581–583. Available at: https://doi.org/10.1038/nmeth.3869.

Callaway, E. (2025) 'These are the 20 most-studied bacteria — the majority have been ignored', *Nature*, 637(8047), pp. 770–771. Available at: https://doi.org/10.1038/d41586-025-00038-x.

Callewaert, C. *et al.* (2014) 'Deodorants and antiperspirants affect the axillary bacterial community', *Archives of Dermatological Research*, 306(8), pp. 701–710. Available at: https://doi.org/10.1007/S00403-014-1487-1/TABLES/1.

Campbell, J.W., Morgan-Kiss, R.M. and Cronan, J.E. (2003) 'A new Escherichia coli metabolic competency: growth on fatty acids by a novel anaerobic beta-oxidation pathway', *Molecular Microbiology*, 47(3), pp. 793–805. Available at: https://doi.org/10.1046/j.1365-2958.2003.03341.x.

Campbell, M.J. and Walters, S.J. (2014) *How to Design, Analyse and Report Cluster Randomised Trials in Medicine and Health Related Research*. Newark, UNITED KINGDOM: John Wiley & Sons, Incorporated. Available at: http://ebookcentral.proquest.com/lib/aston/detail.action?docID=1662762 (Accessed: 1 September 2023).

Cao, Q. *et al.* (2021) 'Effects of Rare Microbiome Taxa Filtering on Statistical Analysis', *Frontiers in Microbiology*, 11. Available at: https://doi.org/10.3389/fmicb.2020.607325.

Cao, Y. *et al.* (2022) 'microbiomeMarker: an R/Bioconductor package for microbiome marker identification and visualization', *Bioinformatics*. Edited by T. Marschall, 38(16), pp. 4027–4029. Available at: https://doi.org/10.1093/bioinformatics/btac438.

Cappellato, M., Baruzzo, G. and Camillo, B.D. (2022) 'Investigating differential abundance methods in microbiome data: A benchmark study', *PLOS Computational Biology*, 18(9), p. e1010467. Available at: https://doi.org/10.1371/journal.pcbi.1010467.

Carville, K. *et al.* (2014) 'The effectiveness of a twice-daily skin-moisturising regimen for reducing the incidence of skin tears', *International Wound Journal*, 11(4), pp. 446–453. Available at: https://doi.org/10.1111/iwj.12326.

Caselli, E. *et al.* (2020) 'Defining the oral microbiome by whole-genome sequencing and resistome analysis: The complexity of the healthy picture', *BMC Microbiology*, 20(1), pp. 1–19. Available at: https://doi.org/10.1186/s12866-020-01801-y.

Castelino, M. *et al.* (2017) 'Optimisation of methods for bacterial skin microbiome investigation: primer selection and comparison of the 454 versus MiSeq platform', *BMC Microbiology*, 17(1), pp. 1–12. Available at: https://doi.org/10.1186/s12866-017-0927-4.

Catney, G., Wright, R. and Ellis, M. (2021) 'The evolution and stability of multi-ethnic residential neighbourhoods in England', *Transactions (Institute of British Geographers : 1965)*, 46(2), pp. 330–346. Available at: https://doi.org/10.1111/tran.12416.

Cattaneo, C. *et al.* (2019) 'New insights into the relationship between taste perception and oral microbiota composition', *Scientific Reports 2019 9:1*, 9(1), pp. 1–8. Available at: https://doi.org/10.1038/s41598-019-40374-3.

Caufield, J.H. *et al.* (2017) 'Bacterial protein meta-interactomes predict cross-species interactions and protein function', *BMC Bioinformatics*, 18(1), p. 171. Available at: https://doi.org/10.1186/s12859-017-1585-0.

Chambers, S.T. *et al.* (2013) 'HACEK Infective Endocarditis: Characteristics and Outcomes from a Large, Multi-National Cohort', *PLOS ONE*, 8(5), p. e63181. Available at: https://doi.org/10.1371/journal.pone.0063181.

Chen, G. *et al.* (2021) 'Gut–Brain–Skin Axis in Psoriasis: A Review', *Dermatology and Therapy*, 11(1), pp. 25–38. Available at: https://doi.org/10.1007/S13555-020-00466-9/TABLES/2.

Chen, J. *et al.* (2020) 'The gallbladder and vermiform appendix influence the assemblage of intestinal microorganisms', *https://doi.org/10.2217/fmb-2019-0325*, 15(7), pp. 541–555. Available at: https://doi.org/10.2217/FMB-2019-0325.

Chen, W. *et al.* (2010) 'Gender aspects in skin diseases', *Journal of the European Academy of Dermatology and Venereology*, 24(12), pp. 1378–1385. Available at: https://doi.org/10.1111/j.1468-3083.2010.03668.x.

Chiller, K., Selkin, B.A. and Murakawa, G.J. (2001) 'Skin Microflora and Bacterial Infections of the Skin', *Journal of Investigative Dermatology Symposium Proceedings*, 6(3), pp. 170–174. Available at: https://doi.org/10.1046/J.0022-202X.2001.00043.X.

Cho, T. *et al.* (2014) 'Importance of Diversity in the Oral Microbiota including Candida Species Revealed by High-Throughput Technologies', *International Journal of Dentistry*, 2014, pp. 1–5. Available at: https://doi.org/10.1155/2014/454391.

Choi, E.H. (2019) 'Aging of the skin barrier', *Clinics in Dermatology*, 37(4), pp. 336–345. Available at: https://doi.org/10.1016/j.clindermatol.2019.04.009.

Choo, J.M., Leong, L.E. and Rogers, G.B. (2015) 'Sample storage conditions significantly influence faecal microbiome profiles', *Scientific Reports*, 5(1), p. 16350. Available at: https://doi.org/10.1038/srep16350.

Chowdhry, A. *et al.* (2023) 'Exploring the oral microbiome: an updated multidisciplinary oral healthcare perspective', *Discoveries*, 11(2), p. e165. Available at: https://doi.org/10.15190/d.2023.4.

Chrisman, B. *et al.* (2022) 'The human "contaminome": bacterial, viral, and computational contamination in whole genome sequences from 1000 families', *Scientific Reports*, 12(1), p. 9863. Available at: https://doi.org/10.1038/s41598-022-13269-z.

Christensen, J.J. and Ruoff, K.L. (2015) 'Aerococcus, Abiotrophia, and Other Aerobic Catalase-Negative, Gram-Positive Cocci', in *Manual of Clinical Microbiology*. John Wiley & Sons, Ltd, pp. 422–436. Available at: https://doi.org/10.1128/9781555817381.ch24.

Chu, S., Chan, A.K.Y. and Chu, C.H. (2024) 'Polyamines in Dysbiotic Oral Conditions of Older Adults: A Scoping Review', *International Journal of Molecular Sciences*, 25(19), p. 10596. Available at: https://doi.org/10.3390/ijms251910596.

Chun, B.H. *et al.* (2017) 'Cloacibacterium caeni sp. nov., isolated from activated sludge', *International Journal of Systematic and Evolutionary Microbiology*, 67(6), pp. 1688–1692. Available at: https://doi.org/10.1099/ijsem.0.001841.

Church, D.L. *et al.* (2020) 'Performance and Application of 16S rRNA Gene Cycle Sequencing for Routine Identification of Bacteria in the Clinical Microbiology Laboratory', *Clinical Microbiology Reviews*, 33(4), p. 10.1128/cmr.00053-19. Available at: https://doi.org/10.1128/cmr.00053-19.

Ciantar, M. *et al.* (2005) 'Capnocytophaga spp. in Periodontitis Patients Manifesting Diabetes Mellitus', *Journal of Periodontology*, 76(2), pp. 194–203. Available at: https://doi.org/10.1902/jop.2005.76.2.194.

Clark, A. and Mach, N. (2017) 'The crosstalk between the gut microbiota and mitochondria during exercise', *Frontiers in Physiology*, 8(MAY), p. 319. Available at: https://doi.org/10.3389/FPHYS.2017.00319/BIBTEX.

Clark, W.A. and Coulter, R. (2015) 'Who wants to move? The role of neighbourhood change', *Environment and Planning A: Economy and Space*, 47(12), pp. 2683–2709. Available at: https://doi.org/10.1177/0308518X15615367.

Clavel, T. *et al.* (2004) 'Survival of Bacillus cereus spores and vegetative cells in acid media simulating human stomach', *Journal of Applied Microbiology*, 97(1), pp. 214–219. Available at: https://doi.org/10.1111/J.1365-2672.2004.02292.X.

Collins, S.E. (2016) 'Associations Between Socioeconomic Factors and Alcohol Outcomes', *Alcohol Research : Current Reviews*, 38(1), pp. 83–94.

Cooper, A.J. *et al.* (2015) 'The skin microbiome: Associations between altered microbial communities and disease', *Australasian Journal of Dermatology*, 56(4), pp. 268–274. Available at: https://doi.org/10.1111/ajd.12253.

Corvec, S. *et al.* (2019) 'Taxonomy and phylogeny of *Cutibacterium* (formerly *Propionibacterium*) acnes in inflammatory skin diseases', *Annales de Dermatologie et de Vénéréologie*, 146(1), pp. 26–30. Available at: https://doi.org/10.1016/j.annder.2018.11.002.

Costello, V.H. *et al.* (2022) 'Infective endocarditis due to Haemophilus sputorum', *Access Microbiology*, 4(12), p. 000410. Available at: https://doi.org/10.1099/acmi.0.000410.

Crawford, J.M., Taubman, M.A. and Smith, D.J. (1975) 'Minor Salivary Glands as a Major Source of Secretory Immunoglobulin A in the Human Oral Cavity', *Science*, 190(4220), pp. 1206–1209. Available at: https://doi.org/10.1126/science.1198107.

Cundell, A.M. (2018) 'Microbial Ecology of the Human Skin', *Microbial Ecology*, 76(1), pp. 113–120. Available at: https://doi.org/10.1007/S00248-016-0789-6/TABLES/2.

Dąbrowska, A. k. *et al.* (2018) 'The relationship between skin function, barrier properties, and body-dependent factors', *Skin Research and Technology*, 24(2), pp. 165–174. Available at: https://doi.org/10.1111/srt.12424.

Davis, N.M. *et al.* (2018) 'Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data', *Microbiome*, 6(1), pp. 1–14. Available at: https://doi.org/10.1186/s40168-018-0605-2.

Davis-Turak, J. *et al.* (2017) 'Genomics pipelines and data integration: challenges and opportunities in the research setting', *Expert review of molecular diagnostics*, 17(3), pp. 225–237. Available at: https://doi.org/10.1080/14737159.2017.1282822.

De Sousa, V. *et al.* (2020) 'Bacterial diversity and prevalence of antibiotic resistance genes in the oral microbiome', *PLOS ONE*, 15(9), p. e0239664.

Dean, M. *et al.* (2009) 'Factors influencing eating a varied diet in old age', *Public Health Nutrition*, 12(12), pp. 2421–2427. Available at: https://doi.org/10.1017/S1368980009005448.

DeClercq, V. *et al.* (2024) 'Oral microbial signatures associated with age and frailty in Canadian adults', *Scientific Reports*, 14(1), p. 9685. Available at: https://doi.org/10.1038/s41598-024-60409-8.

Defazio, G. *et al.* (2025) 'kMetaShot: a fast and reliable taxonomy classifier for metagenome-assembled genomes', *Briefings in Bioinformatics*, 26(1), p. bbae680. Available at: https://doi.org/10.1093/bib/bbae680.

Dein, S. and Huline-Dickens, S. (1997) 'Cultural aspects of aging and psychopathology', *Aging & Mental Health*, 1(2), pp. 112–120. Available at: https://doi.org/10.1080/13607869757209.

Delaney, C. *et al.* (2019) 'Interkingdom interactions on the denture surface: Implications for oral hygiene', *Biofilm*, 1(April), p. 100002. Available at: https://doi.org/10.1016/j.bioflm.2019.100002.

Delgado-Angulo, E.K., Mangal, M. and Bernabé, E. (2019) 'Socioeconomic inequalities in adult oral health across different ethnic groups in England', *Health and Quality of Life Outcomes*, 17(1), pp. 1–7. Available at: https://doi.org/10.1186/s12955-019-1156-3.

Denis, S. *et al.* (2023) 'Chapter 4 - Metabolomic study of biofilm-forming natural microbiota of oral biofilm', in H. Sarma et al. (eds) *Microbial Biofilms*. Academic Press (Advances in Biotechnology and Bioengineering), pp. 49–60. Available at: https://doi.org/10.1016/B978-0-323-95715-1.00014-5.

Dent, E., Kowal, P. and Hoogendijk, E.O. (2016) 'Frailty measurement in research and clinical practice: A review', *European Journal of Internal Medicine*, 31, pp. 3–10. Available at: https://doi.org/10.1016/j.ejim.2016.03.007.

Deo, P.N. and Deshmukh, R. (2019) 'Oral microbiome: Unveiling the fundamentals', *Journal of Oral and Maxillofacial Pathology : JOMFP*, 23(1), p. 122. Available at: https://doi.org/10.4103/JOMFP.JOMFP_304_18.

Diebel, L.W.M. and Rockwood, K. (2021) 'Determination of Biological Age: Geriatric Assessment vs Biological Biomarkers', *Current Oncology Reports*, 23(9), p. 104. Available at: https://doi.org/10.1007/s11912-021-01097-9.

Dikeocha, I.J. *et al.* (2022) 'Pharmacomicrobiomics: Influence of gut microbiota on drug and xenobiotic metabolism', *The FASEB Journal*, 36(6), p. e22350. Available at: https://doi.org/10.1096/fj.202101986R.

Dimitriu, P.A. *et al.* (2019) 'New Insights into the Intrinsic and Extrinsic Factors That Shape the Human Skin Microbiome', *mBio*. Edited by D.S. Guttman, 10(4). Available at: https://doi.org/10.1128/mBio.00839-19.

Donaldson, G.P., Lee, S.M. and Mazmanian, S.K. (2016) 'Gut biogeography of the bacterial microbiota', *Nature Reviews Microbiology*, 14(1), pp. 20–32. Available at: https://doi.org/10.1038/nrmicro3552.

Douglas, G.M. *et al.* (2020) 'PICRUSt2 for prediction of metagenome functions', *Nature Biotechnology*, 38(6), pp. 685–688. Available at: https://doi.org/10.1038/s41587-020-0548-6.

Duran-Pinedo, A.E. and Frias-Lopez, J. (2015) 'Beyond microbial community composition: functional activities of the oral microbiome in health and disease', *Microbes and Infection*, 17(7), pp. 505–516. Available at: https://doi.org/10.1016/j.micinf.2015.03.014.

Durazzi, F. *et al.* (2021) 'Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota', *Scientific Reports*, 11(1), p. 3030. Available at: https://doi.org/10.1038/s41598-021-82726-y.

Dutzan, N. *et al.* (2016) 'Characterization of the human immune cell network at the gingival barrier', *Mucosal Immunology*, 9(5), pp. 1163–1172. Available at: https://doi.org/10.1038/mi.2015.136.

Dworkin, M. *et al.* (eds) (2006) *The Prokaryotes: Volume 4: Bacteria: Firmicutes, Cyanobacteria*. New York, NY: Springer US. Available at: https://doi.org/10.1007/0-387-30744-3.

Dwyer-Hemmings, L. (2018) '"A Wicked Operation"? Tonsillectomy in Twentieth-Century Britain', *Medical History*, 62(2), pp. 217–241. Available at: https://doi.org/10.1017/mdh.2018.5.

Ebersole, J.L. *et al.* (2021) 'Oral Microbiome and Gingival Gene Expression of Inflammatory Biomolecules With Aging and Periodontitis', *Frontiers in Oral Health*, 2. Available at: https://doi.org/10.3389/froh.2021.725115.

Eisenhofer, R. *et al.* (2019) 'Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations', *Trends in Microbiology*, 27(2), pp. 105–117. Available at: https://doi.org/10.1016/j.tim.2018.11.003.

Elias, P.M. (2017) 'The how, why and clinical importance of stratum corneum acidification', *Experimental Dermatology*, 26(11), pp. 999–1003. Available at: https://doi.org/10.1111/exd.13329.

Ertel, K. *et al.* (2009) 'Gender differences in attitudes and practices toward body skin care', *Journal of the American Academy of Dermatology*, 60(3, Supplement 1), p. AB85. Available at: https://doi.org/10.1016/j.jaad.2008.11.383.

Essche, M. van *et al.* (2013) 'Bacterial Antagonism Against Periodontopathogens', *Journal of Periodontology*, 84(6), pp. 801–811. Available at: https://doi.org/10.1902/JOP.2012.120261.

Ewels, P. *et al.* (2016) 'MultiQC: summarize analysis results for multiple tools and samples in a single report', *Bioinformatics*, 32(19), pp. 3047–3048. Available at: https://doi.org/10.1093/bioinformatics/btw354.

Ewing, B. and Green, P. (1998) 'Base-Calling of Automated Sequencer Traces Using Phred. II. Error Probabilities', *Genome Research*, 8(3), pp. 186–194. Available at: https://doi.org/10.1101/gr.8.3.186.

Fan, X. *et al.* (2018) 'Drinking alcohol is associated with variation in the human oral microbiome in a large study of American adults', *Microbiome*, 6(1), pp. 1–15. Available at: https://doi.org/10.1186/s40168-018-0448-x.

Farage, M.A., Miller, K.W. and Maibach, H.I. (2016) *Textbook of Aging Skin*. Berlin, Heidelberg, GERMANY: Springer Berlin / Heidelberg. Available at: http://ebookcentral.proquest.com/lib/aston/detail.action?docID=6314731 (Accessed: 31 August 2023).

Feng, Y. *et al.* (2018) 'The Associations between Biochemical and Microbiological Variables and Taste Differ in Whole Saliva and in the Film Lining the Tongue', *BioMed Research International*, 2018. Available at: https://doi.org/10.1155/2018/2838052.

Fernandez-Flores, A. and Saeb-Lima, M. (2019) 'Histopathology of Cutaneous Aging', *American Journal of Dermatopathology*, 41(7), pp. 469–479. Available at: https://doi.org/10.1097/DAD.0000000000001260.

Filardo, S., Di Pietro, M. and Sessa, R. (2024) 'Current progresses and challenges for microbiome research in human health: a perspective', *Frontiers in Cellular and Infection Microbiology*, 14. Available at: https://doi.org/10.3389/fcimb.2024.1377012.

Fink, J.K. (2018) *Materials, Chemicals and Methods for Dental Applications*. Newark, UNITED STATES: John Wiley & Sons, Incorporated. Available at: http://ebookcentral.proquest.com/lib/aston/detail.action?docID=5328368 (Accessed: 29 September 2023).

Flowers, S.A., Bhat, S. and Lee, J.C. (2020) 'Potential Implications of Gut Microbiota in Drug Pharmacokinetics and Bioavailability', *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, 40(7), pp. 704–712. Available at: https://doi.org/10.1002/PHAR.2428.

Fluhr, J.W. and Elias, P.M. (2002) 'Stratum corneum pH: Formation and Function of the "Acid Mantle"', *Exogenous Dermatology*, 1(4), pp. 163–175. Available at: https://doi.org/10.1159/000066140.

Fluoride - NHS (2018). Available at: https://www.nhs.uk/conditions/fluoride/ (Accessed: 14 July 2022).

Foley, M. and Akers, H. (2019) 'Does poverty cause dental caries?', *Australian Dental Journal*, 64(1), pp. 96–102. Available at: https://doi.org/10.1111/adj.12666.

Franco, A.C., Aveleira, C. and Cavadas, C. (2022) 'Skin senescence: mechanisms and impact on whole-body aging', *Trends in Molecular Medicine*, 28(2), pp. 97–109. Available at: https://doi.org/10.1016/J.MOLMED.2021.12.003.

Franco-Acosta, A. *et al.* (2022) 'Emergence of genital infections due to *Haemophilus pittmaniae* and *Haemophilus sputorum*', *Infectious Diseases Now*, 52(4), pp. 227–229. Available at: https://doi.org/10.1016/j.idnow.2022.04.003.

Funahara, M. *et al.* (2018) 'An analysis of the factors affecting the number of bacteria in the saliva of elderly adults in need of care', *International Journal of Gerontology*, 12(3), pp. 205–207. Available at: https://doi.org/10.1016/J.IJGE.2018.01.003.

Galloway-Peña, J. and Hanson, B. (2020) 'Tools for Analysis of the Microbiome', *Digestive Diseases and Sciences*, 65(3), pp. 674–685. Available at: https://doi.org/10.1007/s10620-020-06091-y.

Gazdeck, R.K. *et al.* (2019) 'Diversity of the oral microbiome between dentate and edentulous individuals', *Oral Diseases*, 25(3), pp. 911–918. Available at: https://doi.org/10.1111/ODI.13039.

Ghimire, N. *et al.* (2024) 'Comparative genomics based exploration of xenobiotic degradation patterns in Glutamicibacter, Arthrobacter, and Pseudarthrobacter isolated from diverse ecological habitats', *Heliyon*, 10(22). Available at: https://doi.org/10.1016/j.heliyon.2024.e40280.

Giacomini, J.J. *et al.* (2024) 'Spatial ecology of Haemophilus and Aggregatibacter in the human oral cavity', *Microbiology Spectrum*, 12(4), pp. e04017-23. Available at: https://doi.org/10.1128/spectrum.04017-23.

Giacomini, J.J. *et al.* (2025) 'Spatial ecology of the Neisseriaceae family in the human oral cavity', *Microbiology Spectrum*, 13(5), pp. e03275-24. Available at: https://doi.org/10.1128/spectrum.03275-24.

Giordano-Kelhoffer, B. *et al.* (2022) 'Oral Microbiota, Its Equilibrium and Implications in the Pathophysiology of Human Diseases: A Systematic Review', *Biomedicines*, 10(8), p. 1803. Available at: https://doi.org/10.3390/biomedicines10081803.

Glassing, A. *et al.* (2016) 'Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples', *Gut Pathogens*, 8, p. 24. Available at: https://doi.org/10.1186/s13099-016-0103-7.

Goberna, M. and Verdú, M. (2022) 'Cautionary notes on the use of co-occurrence networks in soil ecology', *Soil Biology and Biochemistry*, 166, p. 108534. Available at: https://doi.org/10.1016/j.soilbio.2021.108534.

Goig, G.A. *et al.* (2020) 'Contaminant DNA in bacterial sequencing experiments is a major source of false genetic variability', *BMC Biology*, 18(1), pp. 1–15. Available at: https://doi.org/10.1186/s12915-020-0748-z.

Golob, J.L. *et al.* (2017) 'Evaluating the accuracy of amplicon-based microbiome computational pipelines on simulated human gut microbial communities', *BMC Bioinformatics*, 18(1), p. 283. Available at: https://doi.org/10.1186/s12859-017-1690-0.

Gong, B. *et al.* (2021) 'Association Between Gut Microbiota and Autoimmune Thyroid Disease: A Systematic Review and Meta-Analysis', *Frontiers in Endocrinology*, 12. Available at: https://www.frontiersin.org/articles/10.3389/fendo.2021.774362 (Accessed: 31 August 2023).

Gotschlich, E.C., Colbert, R.A. and Gill, T. (2019) 'Methods in microbiome research: Past, present, and future', *Best Practice & Research Clinical Rheumatology*, 33(6), p. 101498. Available at: https://doi.org/10.1016/j.berh.2020.101498.

Graham, E.H. *et al.* (2022) 'The application of the skin virome for human identification', *Forensic Science International: Genetics*, 57. Available at: https://doi.org/10.1016/J.FSIGEN.2022.102662.

'GraphPad Prism' (2023). San Diego, California USA: GraphPad Software. Available at: www.graphpad.com.

Grice, E.A. *et al.* (2008) 'A diversity profile of the human skin microbiota', *Genome Research*, 18(7), pp. 1043–1050. Available at: https://doi.org/10.1101/gr.075549.107.

Grice, E.A. *et al.* (2009) 'Topographical and Temporal Diversity of the Human Skin Microbiome', *Science*, 324(5931), pp. 1190–1192. Available at: https://doi.org/10.1126/science.1171700.

Guarner, F. and Malagelada, J.-R. (2003) 'Gut flora in health and disease', *The Lancet*, 361(9356), pp. 512–519. Available at: https://doi.org/10.1016/S0140-6736(03)12489-0.

Guilloux, C.-A. *et al.* (2021) '*Porphyromonas*: A neglected potential key genus in human microbiomes', *Anaerobe*, 68, p. 102230. Available at: https://doi.org/10.1016/j.anaerobe.2020.102230.

Guo, B. *et al.* (2022) 'Microbial co-occurrence network topological properties link with reactor parameters and reveal importance of low-abundance genera', *npj Biofilms and Microbiomes*, 8(1), pp. 1–13. Available at: https://doi.org/10.1038/s41522-021-00263-y.

Guo, Y. *et al.* (2022) 'Forming 4-Methylcatechol as the Dominant Bioavailable Metabolite of Intraruminal Rutin Inhibits p-Cresol Production in Dairy Cows', *Metabolites*, 12(1), p. 16. Available at: https://doi.org/10.3390/metabo12010016.

Gupta, V.K., Paul, S. and Dutta, C. (2017) 'Geography, ethnicity or subsistence-specific variations in human microbiome composition and diversity', *Frontiers in Microbiology*, 8(JUN). Available at: https://doi.org/10.3389/fmicb.2017.01162.

Haider, D. *et al.* (2024) 'Mock microbial community meta-analysis using different trimming of amplicon read lengths', *Environmental Microbiology*, 26(1), p. e16566. Available at: https://doi.org/10.1111/1462-2920.16566.

Hamada, M. *et al.* (2010) 'Mobilicoccus pelagius gen. nov., sp. nov. and Piscicoccus intestinalis gen. nov., sp. nov., two new members of the family Dermatophilaceae, and reclassification of Dermatophilus chelonae (Masters et al. 1995) as Austwickia chelonae gen. nov., comb. nov', *The Journal of General and Applied Microbiology*, 56(6), pp. 427–436. Available at: https://doi.org/10.2323/jgam.56.427.

Han, Y.W. (2015) 'Fusobacterium nucleatum: a commensal-turned pathogen', *Current Opinion in Microbiology*, 23, pp. 141–147. Available at: https://doi.org/10.1016/j.mib.2014.11.013.

Hannigan, G.D. *et al.* (2015) 'The Human Skin Double-Stranded DNA Virome: Topographical and Temporal Diversity, Genetic Enrichment, and Dynamic Associations with the Host Microbiome', *mBio*. Edited by J. Heitman, 6(5). Available at: https://doi.org/10.1128/mBio.01578-15.

Hapca, A. *et al.* (2014) 'Effectiveness of newspaper advertising for patient recruitment into a clinical trial', *British Journal of Clinical Pharmacology*, 77(6), pp. 1064–1072. Available at: https://doi.org/10.1111/bcp.12262.

Haraszthy, V.I., Raylae, C.C. and Sreenivasan, P.K. (2019) 'Antimicrobial effects of a stannous fluoride toothpaste in distinct oral microenvironments', *JOURNAL OF THE AMERICAN DENTAL ASSOCIATION*, 150(4), pp. S14–S24. Available at: https://doi.org/10.1016/j.adaj.2019.01.007.

Hausauer, A.K. *et al.* (2011) 'Increases in Melanoma Among Adolescent Girls and Young Women in California: Trends by Socioeconomic Status and UV Radiation Exposure', *Archives of Dermatology*, 147(7), pp. 783–789. Available at: https://doi.org/10.1001/ARCHDERMATOL.2011.44.

Hay, R.J. *et al.* (2014) 'The Global Burden of Skin Disease in 2010: An Analysis of the Prevalence and Impact of Skin Conditions', *Journal of Investigative Dermatology*, 134(6), pp. 1527–1534. Available at: https://doi.org/10.1038/jid.2013.446.

Hayes, W. and Sahu, S. (2020) 'The Human Microbiome: History and Future', *Journal of Pharmacy & Pharmaceutical Sciences*, 23(8), pp. 406–411. Available at: https://doi.org/10.18433/jpps31525.

Hedberg, M.E. *et al.* (2012) 'Lachnoanaerobaculum gen. nov., a new genus in the Lachnospiraceae: characterization of Lachnoanaerobaculum umeaense gen. nov., sp. nov., isolated from the human small intestine, and Lachnoanaerobaculum orale sp. nov., isolated from saliva, and reclassification of Eubacterium

saburreum (Prévot 1966) Holdeman and Moore 1970 as Lachnoanaerobaculum saburreum comb. nov.', *International Journal of Systematic and Evolutionary Microbiology*, 62(Pt 11), pp. 2685–2690. Available at: https://doi.org/10.1099/ijs.0.033613-0.

Hiergeist, A. *et al.* (2015) 'Analyses of Intestinal Microbiota: Culture versus Sequencing', *ILAR Journal*, 56(2), pp. 228–240. Available at: https://doi.org/10.1093/ilar/ilv017.

Hillebrand, G.G. *et al.* (2021) 'Temporal Variation of the Facial Skin Microbiome: A 2-Year Longitudinal Study in Healthy Adults', *Plastic and reconstructive surgery*, 147(1 2), pp. 50S-61S. Available at: https://doi.org/10.1097/PRS.0000000000007621.

Hiltunen, K. and Vehkalahti, M.M. (2023) 'Why and when older people lose their teeth: A study of public healthcare patients aged 60 years and over in 2007-2015', *Gerodontology*, 40(3), pp. 326–333. Available at: https://doi.org/10.1111/ger.12657.

Ho, C.Y. and Dreesen, O. (2021) 'Faces of cellular senescence in skin aging', *Mechanisms of Ageing and Development*, 198, p. 111525. Available at: https://doi.org/10.1016/J.MAD.2021.111525.

Holden, G. *et al.* (1993) 'The Recruitment of Research Participants:: A Review', *Social Work in Health Care*, 19(2), pp. 1–44. Available at: https://doi.org/10.1300/J010v19n02_01.

Holmes, I., Harris, K. and Quince, C. (2012) 'Dirichlet Multinomial Mixtures: Generative Models for Microbial Metagenomics', *PLOS ONE*, 7(2), p. e30126. Available at: https://doi.org/10.1371/journal.pone.0030126.

Holst, O. *et al.* (1991) 'Structural analysis of the heptose/hexose region of the lipopolysaccharide from Escherichia coli K-12 strain W3100', *Carbohydrate Research*, 215(2), pp. 323–335. Available at: https://doi.org/10.1016/0008-6215(91)84031-9.

Hon, K. *et al.* (2015) 'Patient acceptability, efficacy, and skin biophysiology of a cream and cleanser containing lipid complex with shea butter extract versus a ceramide product for eczema', *Hong Kong Medical Journal* [Preprint]. Available at: https://doi.org/10.12809/hkmj144472.

Hong, I. *et al.* (2020) 'Clinical and Microbiological Efficacy of Pyrophosphate Containing Toothpaste: A Double-Blinded Placebo-Controlled Randomized Clinical Trial', *MICROORGANISMS*, 8(11), p. 1806. Available at: https://doi.org/10.3390/microorganisms8111806.

Houge, C.-M. and Ruiz, J.G. (2022) *Oral Health and Aging*. Edited by C.-M. Hogue and J.G. Ruiz. Cham: Springer International Publishing. Available at: https://doi.org/10.1007/978-3-030-85993-0.

House, J.S., Kessler, R.C. and Herzog, A.R. (1990) 'Age, Socioeconomic Status, and Health', *The Milbank Quarterly*, 68(3), pp. 383–411. Available at: https://doi.org/10.2307/3350111.

Hovav, A.H. (2014) 'Dendritic cells of the oral mucosa', *Mucosal Immunology*, 7(1), pp. 27–37. Available at: https://doi.org/10.1038/mi.2013.42.

How, K.Y., Song, K.P. and Chan, K.G. (2016) 'Porphyromonas gingivalis: An Overview of Periodontopathic Pathogen below the Gum Line', *Frontiers in Microbiology*, 7, p. 53. Available at: https://doi.org/10.3389/fmicb.2016.00053.

Hung, L.-W., Kempen, G.I.J.M. and Vries, N.K.D. (2010) 'Cross-cultural comparison between academic and lay views of healthy ageing: a literature review', *Ageing & Society*, 30(8), pp. 1373–1391. Available at: https://doi.org/10.1017/S0144686X10000589.

Hwang, K. and Baik, S.H. (1997) 'Distribution of Hairs and Sweat Glands on the Bodies of Korean Adults: A Morphometric Study', *Cells Tissues Organs*, 158(2), pp. 112–120. Available at: https://doi.org/10.1159/000147920.

'IBM SPSS Statistics' (2021). Armonk, New York: IBM Corp.

Ida, Y. *et al.* (2022) 'First description of Lachnoanaerobaculum orale as a possible cause of human bacteremia', *Anaerobe*, 73, p. 102506. Available at: https://doi.org/10.1016/j.anaerobe.2021.102506.

Idate, U. *et al.* (2020) 'Molecular identification of Capnocytophaga species from the oral cavity of patients with chronic periodontitis and healthy individuals', *Journal of Oral and Maxillofacial Pathology*, 24(2), p. 397. Available at: https://doi.org/10.4103/jomfp.JOMFP_33_20.

Iglesia, S., Kononov, T. and Zahr, A.S. (2022) 'A multi-functional anti-aging moisturizer maintains a diverse and balanced facial skin microbiome', *Journal of Applied Microbiology*, 133(3), pp. 1791–1799. Available at: https://doi.org/10.1111/jam.15663.

Iheozor-Ejiofor, Z. *et al.* (2024) 'Water fluoridation for the prevention of dental caries', *Cochrane Database of Systematic Reviews* [Preprint], (10). Available at: https://doi.org/10.1002/14651858.CD010856.pub3.

Ii, M.S.R. *et al.* (2021) 'RESCRIPt: Reproducible sequence taxonomy reference database management', *PLOS Computational Biology*, 17(11), p. e1009581. Available at: https://doi.org/10.1371/journal.pcbi.1009581.

Ilha, E.C. *et al.* (2016) 'Comparison of real-time PCR assay and plate count for Lactobacillus paracasei enumeration in yoghurt', *Annals of Microbiology*, 66(2), pp. 597–606. Available at: https://doi.org/10.1007/s13213-015-1137-7.

Iqbal, A. *et al.* (2024) 'A Literature Review of Cutibacterium Acnes: From Skin Commensal to Pathogen in Shoulder Surgery', *Cureus*, 16(9). Available at: https://doi.org/10.7759/cureus.69460.

Ishii, Y. *et al.* (2006) 'Identification of biochemically atypical Staphylococcus aureus clinical isolates with three automated identification systems', *Journal of medical microbiology*, 55(Pt 4), pp. 387–392. Available at: https://doi.org/10.1099/JMM.0.46231-0.

Iwauchi, M. *et al.* (2019) 'Relationship between oral and gut microbiota in elderly people', *Immunity, Inflammation and Disease*, 7(3), pp. 229–236. Available at: https://doi.org/10.1002/IID3.266.

Jackson, D.N. and Theiss, A.L. (2020) 'Gut bacteria signaling to mitochondria in intestinal inflammation and cancer', *Gut microbes*, 11(3), pp. 285–304. Available at: https://doi.org/10.1080/19490976.2019.1592421.

Jansson, L. *et al.* (2020) 'Impact of swab material on microbial surface sampling', *Journal of Microbiological Methods*, 176, p. 106006. Available at: https://doi.org/10.1016/j.mimet.2020.106006.

Jari Oksanen *et al.* (2025) 'vegan: Community Ecology Package'. Available at: https://CRAN.R-project.org/package=vegan.

Jensen, M.G. *et al.* (2023) 'Species- and strain-level diversity of Corynebacteria isolated from human facial skin', *BMC Microbiology*, 23(1), pp. 1–14. Available at: https://doi.org/10.1186/s12866-023-03129-9.

Jeske, J.T. and Gallert, C. (2022) 'Microbiome Analysis via OTU and ASV-Based Pipelines—A Comparative Interpretation of Ecological Data in WWTP Systems', *Bioengineering*, 9(4), p. 146. Available at: https://doi.org/10.3390/bioengineering9040146.

Jiang, Q. *et al.* (2019) 'The Oral Microbiome in the Elderly With Dental Caries and Health', *Frontiers in Cellular and Infection Microbiology*, 8. Available at: https://doi.org/10.3389/fcimb.2018.00442.

Jiang, W. *et al.* (2015) 'The Impact of Various Time Intervals on the Supragingival Plaque Dynamic Core Microbiome', *PLOS ONE*, 10(5), p. e0124631. Available at: https://doi.org/10.1371/journal.pone.0124631.

John E. Hall and Arthur C. Guyton (2010) *Guyton and Hall Textbook of Medical Physiology*. Saunders/Elsevier. Available at: https://shop.elsevier.com/books/guyton-and-hall-textbook-of-medical-physiology/hall/978-1-4160-4574-8 (Accessed: 20 May 2025).

Jordan E Bisanz (2018) 'qiime2R: Importing QIIME2 artifacts and associated data into R sessions'. Available at: https://github.com/jbisanz/qiime2R.

Jovel, J. *et al.* (2016) 'Characterization of the Gut Microbiome Using 16S or Shotgun Metagenomics', *Frontiers in Microbiology*, 7. Available at: https://doi.org/10.3389/fmicb.2016.00459.

Ju, F. and Zhang, T. (2015) '16S rRNA gene high-throughput sequencing data mining of microbial diversity and interactions', *Applied Microbiology and Biotechnology*, 99(10), pp. 4119–4129. Available at: https://doi.org/10.1007/s00253-015-6536-y.

Jung, A. *et al.* (2018) 'A Review of Enterococcus cecorum Infection in Poultry', *Avian Diseases*, 62(3), pp. 261–271. Available at: https://doi.org/10.1637/11825-030618-Review.1.

Jung, S. (2025) 'Advances in functional analysis of the microbiome: Integrating metabolic modeling, metabolite prediction, and pathway inference with Next-Generation Sequencing data', *Journal of Microbiology*, 63(1), p. e:2411006. Available at: https://doi.org/10.71150/jm.2411006.

Jung, Y. *et al.* (2024) 'Aging-Induced Changes in Cutibacterium acnes and Their Effects on Skin Elasticity and Wrinkle Formation', *Microorganisms*, 12(11), p. 2179. Available at: https://doi.org/10.3390/microorganisms12112179.

Kahler, C.M. (2021) 'Neisseria species and their complicated relationships with human health', *Microbiology Australia*, 42(2), pp. 79–83. Available at: https://doi.org/10.1071/MA21024.

Kajihara, K.T. and Hynson, N.A. (2024) 'Networks as tools for defining emergent properties of microbiomes and their stability', *Microbiome*, 12(1), pp. 1–13. Available at: https://doi.org/10.1186/s40168-024-01868-z.

Kamarajan, P. *et al.* (2015) 'Nisin ZP, a Bacteriocin and Food Preservative, Inhibits Head and Neck Cancer Tumorigenesis and Prolongs Survival', *PLOS ONE*, 10(7), p. e0131008. Available at: https://doi.org/10.1371/JOURNAL.PONE.0131008.

Kandrack, M.-A., Grant, K.R. and Segall, A. (1991) 'Gender differences in health related behaviour: Some unanswered questions', *Social Science & Medicine*, 32(5), pp. 579–590. Available at: https://doi.org/10.1016/0277-9536(91)90293-L.

Kane, A.E. and Howlett, S.E. (2021) 'Sex differences in frailty: Comparisons between humans and preclinical models', *Mechanisms of Ageing and Development*, 198, p. 111546. Available at: https://doi.org/10.1016/j.mad.2021.111546.

Kateete, D.P. *et al.* (2010) 'Identification of Staphylococcus aureus: DNase and Mannitol salt agar improve the efficiency of the tube coagulase test', *Annals of Clinical Microbiology and Antimicrobials*, 9(1), pp. 1–7. Available at: https://doi.org/10.1186/1476-0711-9-23.

Kay, J.G., Kramer, J.M. and Visser, M.B. (2019) 'Danger signals in oral cavity-related diseases', *Journal of Leukocyte Biology*, 106(1), pp. 193–200. Available at: https://doi.org/10.1002/JLB.4MIR1118-439R.

Kazarina, A. *et al.* (2023) 'Oral microbiome variations related to ageing: possible implications beyond oral health', *Archives of Microbiology*, 205(4), p. 116. Available at: https://doi.org/10.1007/s00203-023-03464-5.

Keaney, T.C. (2016) 'Aging in the Male Face: Intrinsic and Extrinsic Factors', *Dermatologic Surgery*, 42(7), p. 797. Available at: https://doi.org/10.1097/DSS.0000000000000505.

Killinger, B. and Labrie, V. (2019) 'The Appendix in Parkinson's Disease: From Vestigial Remnant to Vital Organ?', *Journal of Parkinson's Disease*, 9(s2), pp. S345–S358. Available at: https://doi.org/10.3233/JPD-191703.

Kim, D.H. and Rockwood, K. (2024) 'Frailty in Older Adults', *The New England journal of medicine*, 391(6), pp. 538–548. Available at: https://doi.org/10.1056/NEJMra2301292.

Kim, G. *et al.* (2021) 'Spermidine-induced recovery of human dermal structure and barrier function by skin microbiome', *Communications Biology*, 4(1), pp. 1–11. Available at: https://doi.org/10.1038/s42003-020-01619-4.

Kim, H.-J. *et al.* (2019) 'Segregation of age-related skin microbiome characteristics by functionality', *Scientific Reports*, 9(1), p. 16748. Available at: https://doi.org/10.1038/s41598-019-53266-3.

Kim, H.J. *et al.* (2022) 'Aged related human skin microbiome and mycobiome in Korean women', *Scientific Reports 2022 12:1*, 12(1), pp. 1–11. Available at: https://doi.org/10.1038/s41598-022-06189-5.

Kim, J.-H. *et al.* (2021) 'Taxonomic profiling of skin microbiome and correlation with clinical skin parameters in healthy Koreans', *Scientific Reports*, 11(1), p. 16269. Available at: https://doi.org/10.1038/s41598-021-95734-9.

Kim, N. *et al.* (2024) 'Genome-resolved metagenomics: a game changer for microbiome medicine', *Experimental & Molecular Medicine*, 56(7), pp. 1501–1512. Available at: https://doi.org/10.1038/s12276-024-01262-7.

Kim, S. and Jazwinski, S.M. (2015) 'Quantitative measures of healthy aging and biological age', *Healthy aging research*, 4, p. 26. Available at: https://doi.org/10.12715/har.2015.4.26.

Kleinegger, C.L. *et al.* (1996) 'Frequency, Intensity, Species, and Strains of Oral Candida Vary as a Function of Host Age', *JOURNAL OF CLINICAL MICROBIOLOGY*, 34(9), pp. 2246–2254.

Kneidinger, B. *et al.* (2001) 'Biosynthesis of nucleotide-activated D-glycero-D-manno-heptose', *The Journal of Biological Chemistry*, 276(24), pp. 20935–20944. Available at: https://doi.org/10.1074/jbc.M100378200.

Koch, J. *et al.* (1993) 'Products of enzymatic reduction of benzoyl-CoA, a key reaction in anaerobic aromatic metabolism', *European Journal of Biochemistry*, 211(3), pp. 649–661. Available at: https://doi.org/10.1111/j.1432-1033.1993.tb17593.x.

Könönen, E. *et al.* (2022) 'Prevotella species as oral residents and infectious agents with potential impact on systemic conditions', *Journal of Oral Microbiology*, 14(1), p. 2079814. Available at: https://doi.org/10.1080/20002297.2022.2079814.

Kool, J. *et al.* (2023) 'Reducing bias in microbiome research: Comparing methods from sample collection to sequencing', *Frontiers in Microbiology*, 14. Available at: https://doi.org/10.3389/fmicb.2023.1094800.

Kraus, L. *et al.* (2020) 'Long waves of consumption or a unique social generation? Exploring recent declines in youth drinking', *Addiction Research & Theory*, 28(3), pp. 183–193. Available at: https://doi.org/10.1080/16066359.2019.1629426.

Kuehne, A. *et al.* (2017) 'An integrative metabolomics and transcriptomics study to identify metabolic alterations in aged skin of humans in vivo', *BMC Genomics*, 18(1), p. 169. Available at: https://doi.org/10.1186/s12864-017-3547-3.

Kutuzova, S. *et al.* (2024) 'Taxometer: Improving taxonomic classification of metagenomics contigs', *Nature Communications*, 15(1), p. 8357. Available at: https://doi.org/10.1038/s41467-024-52771-y.

Kwak, S. *et al.* (2024) 'Sociobiome - Individual and neighborhood socioeconomic status influence the gut microbiome in a multi-ethnic population in the US', *npj Biofilms and Microbiomes*, 10(1), pp. 1–10. Available at: https://doi.org/10.1038/s41522-024-00491-y.

Lagler, H. *et al.* (2021) 'Comparison of non-invasive Staphylococcus aureus sampling methods on lesional skin in patients with atopic dermatitis', *European Journal of Clinical Microbiology and Infectious Diseases* [Preprint], (0123456789). Available at: https://doi.org/10.1007/s10096-021-04365-5.

Laidlaw, K. *et al.* (2010) 'Attitudes to ageing and expectations for filial piety across Chinese and British cultures: A pilot exploratory evaluation', *Aging & Mental Health*, 14(3), pp. 283–292. Available at: https://doi.org/10.1080/13607860903483060.

Lake, A.D. *et al.* (2020) 'Case of Enterococcus cecorum Human Bacteremia, United States', *HCA Healthcare Journal of Medicine*, 1(6), pp. 495–498. Available at: https://doi.org/10.36518/2689-0216.1126.

Lamaudiere, M. and Sergaki, C. (2024) 'Appropriate Biological Standardisation to Bring Confidence to Omics Data: The Example of the Microbiome', *Food Science and Nutrition Cases*, 2024, p. fsncases20240004. Available at: https://doi.org/10.1079/fsncases.2024.0004.

Lang, I.A. *et al.* (2008) 'Neighborhood Deprivation, Individual Socioeconomic Status, and Cognitive Function in Older People: Analyses from the English Longitudinal Study of Ageing', *Journal of the American Geriatrics Society*, 56(2), pp. 191–198. Available at: https://doi.org/10.1111/j.1532-5415.2007.01557.x.

Larson, P.J. *et al.* (2022) 'Associations of the skin, oral and gut microbiome with aging, frailty and infection risk reservoirs in older adults', *Nature Aging*, 2(10), pp. 941–955. Available at: https://doi.org/10.1038/s43587-022-00287-9.

Laube, S. (2004) 'Skin infections and ageing', *Ageing Research Reviews*, 3(1), pp. 69–89. Available at: https://doi.org/10.1016/J.ARR.2003.08.003.

Lee, H. *et al.* (2021) 'Comparative analysis of human facial skin microbiome between topical sites compared to entire face', *Genes & Genomics*, 43(12), pp. 1483–1495. Available at: https://doi.org/10.1007/s13258-021-01180-2.

Lee, H., Hong, Y. and Kim, M. (2021) 'Structural and Functional Changes and Possible Molecular Mechanisms in Aged Skin', *International Journal of Molecular Sciences 2021, Vol. 22, Page 12489*, 22(22), p. 12489. Available at: https://doi.org/10.3390/IJMS222212489.

Leo Lahti and Sudarshan Shetty (2025) 'microbiome R package'. Available at: http://microbiome.github.io.

Leung, M.H.Y. *et al.* (2020) 'Changes of the human skin microbiota upon chronic exposure to polycyclic aromatic hydrocarbon pollutants', *Microbiome*, 8(1), p. 100. Available at: https://doi.org/10.1186/s40168-020-00874-1.

Lewis, S. *et al.* (2021) 'Comparison of 16S and whole genome dog microbiomes using machine learning', *BioData Mining*, 14(1), pp. 1–15. Available at: https://doi.org/10.1186/s13040-021-00270-x.

Li, M. *et al.* (2019) 'The influence of age, gender and race/ethnicity on the composition of the human axillary microbiome', *International Journal of Cosmetic Science*, 41(4), pp. 371–377. Available at: https://doi.org/10.1111/ics.12549.

Li, R. *et al.* (2023) 'Association Between Skin Acid Mantle, Natural Moisturizing Factors, and Antibacterial Activity Against S. aureus in the Stratum Corneum', *Clinical, Cosmetic and Investigational Dermatology*, 16, pp. 1595–1606. Available at: https://doi.org/10.2147/CCID.S409534.

Li, Y., Jiang, B. and Dai, W. (2022) 'A large-scale whole-genome sequencing analysis reveals false positives of bacterial essential genes', *Applied Microbiology and Biotechnology*, 106(1), pp. 341–347. Available at: https://doi.org/10.1007/s00253-021-11702-3.

Lin, S.-Y. *et al.* (2018) 'Sphingomonas colocasiae sp. nov., isolated from taro (Colocasia esculanta)', *International Journal of Systematic and Evolutionary Microbiology*, 68(1), pp. 133–140. Available at: https://doi.org/10.1099/ijsem.0.002471.

Lithgow, K.V. *et al.* (2022) 'Protease activities of vaginal Porphyromonas species disrupt coagulation and extracellular matrix in the cervicovaginal niche', *npj Biofilms and Microbiomes*, 8(1), pp. 1–15. Available at: https://doi.org/10.1038/s41522-022-00270-7.

Liu, H. *et al.* (2023) 'From nitrate to NO: potential effects of nitrate-reducing bacteria on systemic health and disease', *European Journal of Medical Research*, 28, p. 425. Available at: https://doi.org/10.1186/s40001-023-01413-y.

Liu, L. and Chen, J. (2023) 'Advances in Relationship Between Alcohol Consumption and Skin Diseases', *Clinical, Cosmetic and Investigational Dermatology*, 16, pp. 3785–3791. Available at: https://doi.org/10.2147/CCID.S443128.

Liu, X. *et al.* (2022) 'Sex differences in the oral microbiome, host traits, and their causal relationships', *iScience*, 26(1), p. 105839. Available at: https://doi.org/10.1016/j.isci.2022.105839.

Logan, J.M. *et al.* (2000) 'Campylobacter lanienae sp. nov., a new species isolated from workers in an abattoir.', *International Journal of Systematic and Evolutionary Microbiology*, 50(2), pp. 865–872. Available at: https://doi.org/10.1099/00207713-50-2-865.

López-Otín, C. *et al.* (2013) 'The Hallmarks of Aging', *Cell*, 153(6), pp. 1194–1217. Available at: https://doi.org/10.1016/J.CELL.2013.05.039.

Lousada, M.B. *et al.* (2021) 'Exploring the human hair follicle microbiome*', *British Journal of Dermatology*, 184(5), pp. 802–815. Available at: https://doi.org/10.1111/bjd.19461.

Luengo, J.M., García, J.L. and Olivera, E.R. (2001) 'The phenylacetyl-CoA catabolon: a complex catabolic unit with broad biotechnological applications', *Molecular Microbiology*, 39(6), pp. 1434–1442. Available at: https://doi.org/10.1046/j.1365-2958.2001.02344.x.

Luna, P.C. (2020) 'Skin Microbiome as Years Go By', *American Journal of Clinical Dermatology*, 21(s1), pp. 12–17. Available at: https://doi.org/10.1007/s40257-020-00549-5.

Lussi, A., Schaffner, M. and Jaeggi, T. (2007) 'Dental erosion - Diagnosis and prevention in children and adults', *International Dental Journal*, 57(6 SUPPL.), pp. 385–398. Available at: https://doi.org/10.1111/J.1875-595X.2007.TB00166.X.

Ma, X. *et al.* (2019) 'Analysis of error profiles in deep next-generation sequencing data', *Genome Biology*, 20(1), p. 50. Available at: https://doi.org/10.1186/s13059-019-1659-6.

Maitre, Y. *et al.* (2021) 'Evidence and Therapeutic Perspectives in the Relationship between the Oral Microbiome and Alzheimer's Disease: A Systematic Review', *International Journal of Environmental Research and Public Health*, 18(21), p. 11157. Available at: https://doi.org/10.3390/ijerph182111157.

Makrantonaki, E. *et al.* (2013) 'Skin aging', *MMW Fortschritte der Medizin*, 155 Spec N(2), pp. 50–4; quiz 55.

Maley, S.J. *et al.* (2024) 'Alcohol Consumption and the Diversity of the Oral Microbiome in Postmenopausal Women', *The Journal of Nutrition*, 154(1), pp. 202–212. Available at: https://doi.org/10.1016/j.tjnut.2023.10.025.

Malla, M.A. *et al.* (2019) 'Exploring the Human Microbiome: The Potential Future Role of Next-Generation Sequencing in Disease Diagnosis and Treatment', *Frontiers in Immunology*, 9. Available at: https://doi.org/10.3389/fimmu.2018.02868.

Mansfield, K. and Naik, S. (2020) 'Unraveling Immune-Epithelial Interactions in Skin Homeostasis and Injury.', *The Yale journal of biology and medicine*, 93(1), pp. 133–143.

Manus, M.B. *et al.* (2022) 'Comparing different sample collection and storage methods for field-based skin microbiome research', *American Journal of Human Biology*, 34(1), p. e23584. Available at: https://doi.org/10.1002/ajhb.23584.

Marietta, E.V. *et al.* (2016) 'Human Gut-Derived Prevotella histicola Suppresses Inflammatory Arthritis in Humanized Mice', *Arthritis & rheumatology (Hoboken, N.J.)*, 68(12), pp. 2878–2888. Available at: https://doi.org/10.1002/art.39785.

Mark Welch, J.L. *et al.* (2016) 'Biogeography of a human oral microbiome at the micron scale', *Proceedings of the National Academy of Sciences*, 113(6), pp. E791–E800. Available at: https://doi.org/10.1073/pnas.1522149113.

Markey, L. *et al.* (2025) 'Microbiome diversity of low biomass skin sites is captured by metagenomics but not 16S amplicon sequencing'.

Marples, R.R. and Kligman, A.M. (1969) 'Growth of Bacteria Under Adhesive Tapes', *Archives of Dermatology*, 99(1), pp. 107–110. Available at: https://doi.org/10.1001/ARCHDERM.1969.01610190113020.

Marsh, P.D. *et al.* (2016) 'Influence of saliva on the oral microbiota', *Periodontology 2000*, 70(1), pp. 80–92. Available at: https://doi.org/10.1111/PRD.12098.

Martin, C.R. *et al.* (2018) 'The Brain-Gut-Microbiome Axis', *Cellular and Molecular Gastroenterology and Hepatology*, 6(2), pp. 133–148. Available at: https://doi.org/10.1016/j.jcmgh.2018.04.003.

Martin, M. (2011) 'Cutadapt removes adapter sequences from high-throughput sequencing reads', *EMBnet.journal*, 17(1), pp. 10–12. Available at: https://doi.org/10.14806/ej.17.1.200.

Martínez-García, M. and Hernández-Lemus, E. (2021) 'Periodontal Inflammation and Systemic Diseases: An Overview', *Frontiers in Physiology*, 12, p. 709438. Available at: https://doi.org/10.3389/fphys.2021.709438.

McLean, A.R. *et al.* (2022) 'Site-tropism of streptococci in the oral microbiome', *Molecular Oral Microbiology*, 37(6), pp. 229–243. Available at: https://doi.org/10.1111/omi.12387.

McMahan, R.H. *et al.* (2023) 'Changes in gut microbiome correlate with intestinal barrier dysfunction and inflammation following a 3-day ethanol exposure in aged mice', *Alcohol*, 107, pp. 136–143. Available at: https://doi.org/10.1016/j.alcohol.2022.08.011.

McMurdie, P.J. and Holmes, S. (2013) 'phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data', *PLOS ONE*, 8(4), p. e61217. Available at: https://doi.org/10.1371/journal.pone.0061217.

Melaugh, W. *et al.* (1992) 'Partial characterization of the major lipooligosaccharide from a strain of Haemophilus ducreyi, the causative agent of chancroid, a genital ulcer disease', *The Journal of Biological Chemistry*, 267(19), pp. 13434–13439.

Miles, A.A., Misra, S.S. and Irwin, J.O. (1938) 'The estimation of the bactericidal power of the blood', *The Journal of Hygiene*, 38(6), p. 732. Available at: https://doi.org/10.1017/S002217240001158X.

Mim, M.F. *et al.* (2024) 'The dynamic relationship between skin microbiomes and personal care products: A comprehensive review', *Heliyon*, 10(14). Available at: https://doi.org/10.1016/j.heliyon.2024.e34549.

Min, K. *et al.* (2024) 'Quantitative analysis of the effects of brushing, flossing, and mouthrinsing on supragingival and subgingival plaque microbiota: 12-week clinical trial', *BMC Oral Health*, 24(1), p. 575. Available at: https://doi.org/10.1186/s12903-024-04362-y.

Ministry of Housing, Communities & Local Government (2019) *English indices of deprivation 2019*, *GOV.UK*. Available at: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019 (Accessed: 15 September 2023).

Minty, M. *et al.* (2021) 'Gender-associated differences in oral microbiota and salivary biochemical parameters in response to feeding', *Journal of Physiology and Biochemistry*, 77(1), pp. 155–166. Available at: https://doi.org/10.1007/s13105-020-00757-x.

Mitsui, T., Saito, M. and Harasawa, R. (2018) 'Salivary nitrate-nitrite conversion capacity after nitrate ingestion and incidence of Veillonella spp. in elderly individuals', *Journal of Oral Science*, 60(3), pp. 405–410. Available at: https://doi.org/10.2334/JOSNUSD.17-0337.

Mitsutake, S. *et al.* (2023) 'The effects of dental visits on the occurrence of acute hospitalization for systemic diseases among patients aged 75 years or older: A propensity score-matched study', *Archives of Gerontology and Geriatrics*, 107, p. 104876. Available at: https://doi.org/10.1016/j.archger.2022.104876.

Moissl-Eichinger, C. *et al.* (2017) 'Human age and skin physiology shape diversity and abundance of Archaea on skin', *Scientific Reports 2017 7:1*, 7(1), pp. 1–10. Available at: https://doi.org/10.1038/s41598-017-04197-4.

Moitinho-Silva, L. *et al.* (2021) 'Host traits, lifestyle and environment are associated with human skin bacteria', *British Journal of Dermatology*, 185(3), pp. 573–584. Available at: https://doi.org/10.1111/bjd.20072.

Moreno-Arribas, M.V. *et al.* (2020) 'Relationship between Wine Consumption, Diet and Microbiome Modulation in Alzheimer's Disease', *Nutrients*, 12(10), p. 3082. Available at: https://doi.org/10.3390/nu12103082.

Morley, J.E. (2007) 'The Aging Gut: Physiology', *Clinics in Geriatric Medicine*, 23(4), pp. 757–767. Available at: https://doi.org/10.1016/J.CGER.2007.06.002.

Mosaddad, S.A. *et al.* (2019) 'Oral microbial biofilms: an update', *European Journal of Clinical Microbiology & Infectious Diseases*, 38(11), pp. 2005–2019. Available at: https://doi.org/10.1007/s10096-019-03641-9.

Muhamad Rizal, N.S. *et al.* (2020) 'Advantages and Limitations of 16S rRNA Next-Generation Sequencing for Pathogen Identification in the Diagnostic Microbiology Laboratory: Perspectives from a Middle-Income Country', *Diagnostics*, 10(10), p. 816. Available at: https://doi.org/10.3390/diagnostics10100816.

Myers, T. *et al.* (2024) 'A multi-study analysis enables identification of potential microbial features associated with skin aging signs', *Frontiers in Aging*, 4. Available at: https://doi.org/10.3389/fragi.2023.1304705.

Nagase, S. *et al.* (2020) 'Distinct Skin Microbiome and Skin Physiological Functions Between Bedridden Older Patients and Healthy People: A Single-Center Study in Japan', *Frontiers in Medicine*, 7(April), pp. 1–12. Available at: https://doi.org/10.3389/fmed.2020.00101.

Nakatsuji, T. *et al.* (2013) 'The microbiome extends to subepidermal compartments of normal skin', *Nature Communications 2013 4:1*, 4(1), pp. 1–8. Available at: https://doi.org/10.1038/ncomms2441.

Nearing, J.T. *et al.* (2020) 'Assessing the variation within the oral microbiome of healthy adults', *bioRxiv*, (May), pp. 1–20. Available at: https://doi.org/10.1101/2020.05.07.083634.

Nearing, J.T. *et al.* (2022) 'Microbiome differential abundance methods produce different results across 38 datasets', *Nature Communications*, 13(1), p. 342. Available at: https://doi.org/10.1038/s41467-022-28034-z.

Nemec, A. *et al.* (2003) 'Acinetobacter parvus sp. nov., a small-colony-forming species isolated from human clinical specimens', *International Journal of Systematic and Evolutionary Microbiology*, 53(5), pp. 1563–1567. Available at: https://doi.org/10.1099/ijs.0.02631-0.

Nenoff, P. *et al.* (2014) 'Mycology - an update. Part 1: Dermatomycoses: causative agents, epidemiology and pathogenesis', *Journal der Deutschen Dermatologischen Gesellschaft = Journal of the German Society of Dermatology : JDDG*, 12(3), pp. 188–210. Available at: https://doi.org/10.1111/DDG.12245.

Nguyen, A.V. and Soulika, A.M. (2019) 'The dynamics of the skin's immune system', *International Journal of Molecular Sciences*, 20(8), pp. 1–53. Available at: https://doi.org/10.3390/ijms20081811.

Niedźwiedzka, A. *et al.* (2024) 'The Role of the Skin Microbiome in Acne: Challenges and Future Therapeutic Opportunities', *International Journal of Molecular Sciences*, 25(21), p. 11422. Available at: https://doi.org/10.3390/ijms252111422.

Nikias, M.K., Fink, R. and Shapiro, S. (1975) 'Comparisons of Poverty and Nonpoverty Groups on Dental Status, Needs, and Practices*', *Journal of Public Health Dentistry*, 35(4), pp. 237–259. Available at: https://doi.org/10.1111/j.1752-7325.1975.tb01017.x.

Nikodemova, M. *et al.* (2023) 'The effect of low-abundance OTU filtering methods on the reliability and variability of microbial composition assessed by 16S rRNA amplicon sequencing', *Frontiers in Cellular and Infection Microbiology*, 13. Available at: https://doi.org/10.3389/fcimb.2023.1165295.

Nishimoto, Y. *et al.* (2023) 'Integrated analysis of the oral and intestinal microbiome and metabolome of elderly people with more than 26 original teeth: a pilot study', *Frontiers in Microbiology*, 14. Available at: https://doi.org/10.3389/fmicb.2023.1233460.

Nørskov-Lauritsen, N. and Kilian, M. (2006) 'Reclassification of Actinobacillus actinomycetemcomitans, Haemophilus aphrophilus, Haemophilus paraphrophilus and Haemophilus segnis as Aggregatibacter actinomycetemcomitans gen. nov., comb. nov., Aggregatibacter aphrophilus comb. nov. and Aggregatibacter segnis comb. nov., and emended description of Aggregatibacter aphrophilus to include V factor-dependent and V factor-independent isolates', *International Journal of Systematic and Evolutionary Microbiology*, 56(9), pp. 2135–2146. Available at: https://doi.org/10.1099/ijs.0.64207-0.

Nüse, B. *et al.* (2023) 'L-arginine metabolism as pivotal interface of mutual host–microbe interactions in the gut', *Gut Microbes*, 15(1), p. 2222961. Available at: https://doi.org/10.1080/19490976.2023.2222961.

Oberoi, S.S., Sharma, G. and Oberoi, A. (2016) 'A cross-sectional survey to assess the effect of socioeconomic status on the oral hygiene habits', *Journal of Indian Society of Periodontology*, 20(5), p. 531. Available at: https://doi.org/10.4103/0972-124X.201629.

O'Callaghan, A. and van Sinderen, D. (2016) 'Bifidobacteria and Their Role as Members of the Human Gut Microbiota', *Frontiers in Microbiology*, 7, p. 925. Available at: https://doi.org/10.3389/fmicb.2016.00925.

Office for National Statistics (2022) *Postcode to Output Area to Lower Layer Super Output Area to Middle Layer Super Output Area to Local Authority District (May 2022) Lookup in the UK*. Available at: https://geoportal.statistics.gov.uk/datasets/e7824b1475604212a2325cd373946235 (Accessed: 15 September 2023).

Office for National Statistics (2023) 'Sex-UK Census 2021'. Available at: https://www.ons.gov.uk/datasets/TS008/editions/2021/versions/4.

Ogai, K. *et al.* (2021) 'Amplicon-based skin microbiome profiles collected by tape stripping with different adhesive film dressings: a comparative study', *BMC Microbiology*, 21(1), pp. 1–9. Available at: https://doi.org/10.1186/S12866-021-02122-4/FIGURES/5.

Ogawa, T. *et al.* (2017) 'Characterizations of oral microbiota in elderly nursing home residents with diabetes', *Journal of Oral Science*, 59(4), pp. 549–555. Available at: https://doi.org/10.2334/JOSNUSD.16-0722.

O'Rawe, J.A., Ferson, S. and Lyon, G.J. (2015) 'Accounting for uncertainty in DNA sequencing data', *Trends in Genetics*, 31(2), pp. 61–66. Available at: https://doi.org/10.1016/j.tig.2014.12.002.

Oren, A. and Göker, M. (2023) 'Candidatus List. Lists of names of prokaryotic Candidatus phyla', *International Journal of Systematic and Evolutionary Microbiology*, 73(5), p. 005821. Available at: https://doi.org/10.1099/ijsem.0.005821.

Ottman, N. *et al.* (2012) 'The function of our microbiota: who is out there and what do they do?', *Frontiers in cellular and infection microbiology*, 2, p. 104. Available at: https://doi.org/10.3389/FCIMB.2012.00104/BIBTEX.

Pac, A. *et al.* (2019) 'Influence of Sociodemographic, Behavioral and Other Health-Related Factors on Healthy Ageing Based on Three Operative Definitions', *The journal of nutrition, health & aging*, 23(9), pp. 862–869. Available at: https://doi.org/10.1007/s12603-019-1243-5.

Palmer, J.D. and Foster, K.R. (2022) 'Bacterial species rarely work together', *Science*, 376(6593), pp. 581–582. Available at: https://doi.org/10.1126/science.abn5093.

Paniagua Voirol, L.R. *et al.* (2021) 'How the "kitome" influences the characterization of bacterial communities in lepidopteran samples with low bacterial biomass', *Journal of Applied Microbiology*, 130(6), pp. 1780–1793. Available at: https://doi.org/10.1111/jam.14919.

Pardo, L.M. *et al.* (2024) 'Bacterial Microbiota Composition in Hidradenitis Suppurativa Differs per Skin Layer', *Journal of Investigative Dermatology*, 144(2), pp. 426-430.e5. Available at: https://doi.org/10.1016/j.jid.2023.08.017.

Park, M.G., Cho, S. and Oh, M.M. (2023) 'Menopausal Changes in the Microbiome-A Review Focused on the Genitourinary Microbiome', *Diagnostics (Basel, Switzerland)*, 13(6), p. 1193. Available at: https://doi.org/10.3390/diagnostics13061193.

Park, O.-J. *et al.* (2020) 'Streptococcus gordonii: Pathogenesis and Host Response to Its Cell Wall Components', *Microorganisms*, 8(12), p. 1852. Available at: https://doi.org/10.3390/microorganisms8121852.

Parmenter, K., Waller, J. and Wardle, J. (2000) 'Demographic variation in nutrition knowledge in England', *Health Education Research*, 15(2), pp. 163–174. Available at: https://doi.org/10.1093/her/15.2.163.

Parrey, A.H. *et al.* (2016) 'Aerococcus viridans infection presenting as cutaneous vasculitis in an immunocompetent patient', *Reumatologia*, 54(6), pp. 318–320. Available at: https://doi.org/10.5114/reum.2016.64909.

Patel, J.B. (2001) '16S rRNA Gene Sequencing for Bacterial Pathogen Identification in the Clinical Laboratory', *Molecular Diagnosis*, 6(4), pp. 313–321. Available at: https://doi.org/10.1007/BF03262067.

Pedro Martinez Arbizu (2017) 'pairwiseAdonis: Pairwise Multilevel Comparison using Adonis'. Available at: https://github.com/pmartinezarbizu/pairwiseAdonis.

Pérez-Cobas, A.E., Gomez-Valero, L. and Buchrieser, C. (2020) 'Metagenomic approaches in microbial ecology: an update on whole-genome and marker gene sequencing analyses', *Microbial Genomics*, 6(8), p. mgen000409. Available at: https://doi.org/10.1099/mgen.0.000409.

Pessemier, B.D. *et al.* (2021) 'Gut–Skin Axis: Current Knowledge of the Interrelationship between Microbial Dysbiosis and Skin Conditions', *Microorganisms*, 9(2), pp. 1–33. Available at: https://doi.org/10.3390/MICROORGANISMS9020353.

Peters, B.A. *et al.* (2022) 'Spotlight on the Gut Microbiome in Menopause: Current Insights', *International Journal of Women's Health*, 14, pp. 1059–1072. Available at: https://doi.org/10.2147/IJWH.S340491.

Pfeiffer, F. *et al.* (2018) 'Systematic evaluation of error rates and causes in short samples in next-generation sequencing', *Scientific Reports*, 8(1), p. 10950. Available at: https://doi.org/10.1038/s41598-018-29325-6.

Pfisterer, K. *et al.* (2021) 'The Extracellular Matrix in Skin Inflammation and Infection', *Frontiers in Cell and Developmental Biology*, 9, p. 1578. Available at: https://doi.org/10.3389/FCELL.2021.682414/BIBTEX.

Piggott, D.A. and Tuddenham, S. (2020) 'The gut microbiome and frailty', *Translational Research*, 221, pp. 23–43. Available at: https://doi.org/10.1016/J.TRSL.2020.03.012.

Pinto, R.D. *et al.* (2024) 'Meta-analysis of oral microbiome reveals sex-based diversity in biofilms during periodontitis', *JCI Insight*, 9(17). Available at: https://doi.org/10.1172/jci.insight.171311.

Pistone, D. *et al.* (2021) 'A journey on the skin microbiome: Pitfalls and opportunities', *International Journal of Molecular Sciences*, 22(18). Available at: https://doi.org/10.3390/ijms22189846.

Prasanth, M. (2011) 'Antimicrobial Efficacy of Different Toothpastes and Mouthrinses: An In Vitro Study', *Dental Research Journal*, 8(2), pp. 85–94.

Praveen, Z. *et al.* (2024) 'Oral Microbiome and CPT1A Function in Fatty Acid Metabolism in Oral Cancer', *International Journal of Molecular Sciences*, 25(20), p. 10890. Available at: https://doi.org/10.3390/ijms252010890.

Priyanka, K. (2017) 'Impact of Alcohol Dependency on Oral Health – A Cross-sectional Comparative Study', *JOURNAL OF CLINICAL AND DIAGNOSTIC RESEARCH*, 11(6), p. ZC43. Available at: https://doi.org/10.7860/JCDR/2017/26380.10058.

Proctor, D.M. *et al.* (2020) 'Microbial biogeography and ecology of the mouth and implications for periodontal diseases', *Periodontology 2000*, 82(1), pp. 26–41. Available at: https://doi.org/10.1111/prd.12268.

Puciato, D. *et al.* (2018) 'Physical Activity of Working-Age People in View of Their Income Status', *BioMed Research International*, 2018, p. e8298527. Available at: https://doi.org/10.1155/2018/8298527.

Puri, P. *et al.* (2017) 'Effects of air pollution on the skin: A review', *Indian Journal of Dermatology, Venereology and Leprology*, 83, p. 415. Available at: https://doi.org/10.4103/0378-6323.199579.

Qian, X.-B. *et al.* (2020) 'A guide to human microbiome research: study design, sample collection, and bioinformatics analysis', *Chinese Medical Journal*, 133(15), pp. 1844–1855. Available at: https://doi.org/10.1097/CM9.0000000000000871.

'Qualtrics' (2022). Provo, Utah, USA: Qualtrics. Available at: https://www.qualtrics.com.

Rahmati, E. *et al.* (2017) 'Facklamia Species as an Underrecognized Pathogen', *Open Forum Infectious Diseases*, 4(1), p. ofw272. Available at: https://doi.org/10.1093/ofid/ofw272.

Ramos-Morcillo, A.J. *et al.* (2019) 'Social Determinants of Health, the Family, and Children's Personal Hygiene: A Comparative Study', *International Journal of Environmental Research and Public Health*, 16(23), p. 4713. Available at: https://doi.org/10.3390/ijerph16234713.

Ranjan, R. *et al.* (2016) 'Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing', *Biochemical and Biophysical Research Communications*, 469(4), pp. 967–977. Available at: https://doi.org/10.1016/j.bbrc.2015.12.083.

Ray, K.J. *et al.* (2019) 'High-throughput sequencing of pooled samples to determine community-level microbiome diversity', *Annals of Epidemiology*, 39, pp. 63–68. Available at: https://doi.org/10.1016/j.annepidem.2019.09.002.

Rea, J.N., Newhouse, M.L. and Halil, T. (1976) 'Skin disease in Lambeth. A community study of prevalence and use of medical care', *British Journal of Preventive & Social Medicine*, 30(2), pp. 107–114. Available at: https://doi.org/10.1136/jech.30.2.107.

Reddy, G.S.N. and Garcia-Pichel, F. (2007) 'Sphingomonas mucosissima sp. nov. and Sphingomonas desiccabilis sp. nov., from biological soil crusts in the Colorado Plateau, USA', *International Journal of Systematic and Evolutionary Microbiology*, 57(5), pp. 1028–1034. Available at: https://doi.org/10.1099/ijs.0.64331-0.

Reich, A.J. *et al.* (2020) 'What Does "Successful Aging" Mean to you? — Systematic Review and Cross-Cultural Comparison of Lay Perspectives of Older Adults in 13 Countries, 2010–2020', *Journal of Cross-Cultural Gerontology*, 35(4), pp. 455–478. Available at: https://doi.org/10.1007/s10823-020-09416-6.

Ren, J. *et al.* (2024) 'Insights into the role of Streptococcus oralis as an opportunistic pathogen in infectious diseases', *Frontiers in Cellular and Infection Microbiology*, 14, p. 1480961. Available at: https://doi.org/10.3389/fcimb.2024.1480961.

van Rensburg, J.J. *et al.* (2015) 'The Human Skin Microbiome Associates with the Outcome of and Is Influenced by Bacterial Infection', *mBio*, 6(5), p. 10.1128/mbio.01315-15. Available at: https://doi.org/10.1128/mbio.01315-15.

Renson, A. *et al.* (2019) 'Sociodemographic variation in the oral microbiome', *Annals of Epidemiology*, 35, pp. 73-80.e2. Available at: https://doi.org/10.1016/j.annepidem.2019.03.006.

Rimoldi, S.G. *et al.* (2023) 'Traditional Cultures versus Next Generation Sequencing for Suspected Orthopedic Infection: Experience Gained from a Reference Centre', *Antibiotics*, 12(11), p. 1588. Available at: https://doi.org/10.3390/antibiotics12111588.

Rittie, L. and Fisher, G.J. (2015) 'Natural and Sun-Induced Aging of Human Skin', *Cold Spring Harbor Perspectives in Medicine*, 5(1), pp. a015370–a015370. Available at: https://doi.org/10.1101/cshperspect.a015370.

Roberts, A.P. and Kreth, J. (2014) 'The impact of horizontal gene transfer on the adaptive ability of the human oral microbiome', *Frontiers in Cellular and Infection Microbiology*, 4(SEP), pp. 1–9. Available at: https://doi.org/10.3389/fcimb.2014.00124.

Rodrigues Hoffmann, A. (2017) 'The cutaneous ecosystem: the roles of the skin microbiome in health and its association with inflammatory skin conditions in humans and animals', *Veterinary Dermatology*, 28(1), pp. 60-e15. Available at: https://doi.org/10.1111/vde.12408.

Rooks, M.G. and Garrett, W.S. (2016) 'Gut microbiota, metabolites and host immunity', *Nature Reviews Immunology 2016 16:6*, 16(6), pp. 341–352. Available at: https://doi.org/10.1038/nri.2016.42.

Rosier, B.T. *et al.* (2020) 'Isolation and Characterization of Nitrate-Reducing Bacteria as Potential Probiotics for Oral and Systemic Health', *Frontiers in Microbiology*, 11, p. 555465. Available at: https://doi.org/10.3389/fmicb.2020.555465.

Ross, A.A., Doxey, A.C. and Neufeld, J.D. (2017) 'The Skin Microbiome of Cohabiting Couples', *mSystems*, 2(4). Available at: https://doi.org/10.1128/msystems.00043-17.

Rubin, L.G. (2008) '*Capnocytophaga* Species', in S.S. Long (ed.) *Principles and Practice of Pediatric Infectious Disease (Third Edition)*. Edinburgh: W.B. Saunders, pp. 874–875. Available at: https://doi.org/10.1016/B978-0-7020-3468-8.50171-1.

Rubiola, S. *et al.* (2022) 'Comparison Between Full-Length 16S rRNA Metabarcoding and Whole Metagenome Sequencing Suggests the Use of Either Is Suitable for Large-Scale Microbiome Studies', *Foodborne Pathogens and Disease*, 19(7), pp. 495–504. Available at: https://doi.org/10.1089/fpd.2022.0027.

Ruggiero, S. *et al.* (2018) 'Task force on design and analysis in oral health research: Medication-related osteonecrosis of the jaw', *JDR Clinical and Translational Research*, 3(3), pp. 222–225. Available at: https://doi.org/10.1177/2380084418770662.

Russo, E. *et al.* (2023) 'Facial Skin Microbiome: Aging-Related Changes and Exploratory Functional Associations with Host Genetic Factors, a Pilot Study', *Biomedicines*, 11(3), p. 684. Available at: https://doi.org/10.3390/biomedicines11030684.

Ruuskanen, M.O. *et al.* (2022) 'Towards standardized and reproducible research in skin microbiomes', *Environmental Microbiology* [Preprint]. Available at: https://doi.org/10.1111/1462-2920.15945.

Sabo, C.M. and Dumitrascu, D.L. (2021) 'Microbiota and the irritable bowel syndrome', *Minerva Gastroenterology*, 67(4), pp. 377–384. Available at: https://doi.org/10.23736/S2724-5985.21.02923-5.

Sacks, D., Canadian Paediatric Society, and Adolescent Health Committee (2003) 'Age limits and adolescents', *Paediatrics & Child Health*, 8(9), p. 577. Available at: https://doi.org/10.1093/pch/8.9.577.

Salter, S.J. *et al.* (2014) 'Reagent and laboratory contamination can critically impact sequence-based microbiome analyses', *BMC Biology*, 12(1), p. 87. Available at: https://doi.org/10.1186/s12915-014-0087-z.

SankeyMATIC: Make Beautiful Flow Diagrams (2023). Available at: https://sankeymatic.com/ (Accessed: 19 September 2023).

Santiago-Rodriguez, T.M. *et al.* (2023) 'The Skin Microbiome: Current Techniques, Challenges, and Future Directions', *Microorganisms*, 11(5), p. 1222. Available at: https://doi.org/10.3390/microorganisms11051222.

Santoro, A. *et al.* (2020) 'Microbiomes other than the gut: inflammaging and age-related diseases', *Seminars in Immunopathology 2020 42:5*, 42(5), pp. 589–605. Available at: https://doi.org/10.1007/S00281-020-00814-Z.

Sarkar, K. and Sil, P.C. (2021) 'Chapter 34 - Effect of diet, pharmaceuticals, and environmental toxicants on gut microbiota imbalance and increased intestinal membrane permeability', in A.M. Tsatsakis (ed.) *Toxicological Risk Assessment and Multi-System Health Impacts from Exposure*. Academic Press, pp. 403–413. Available at: https://doi.org/10.1016/B978-0-323-85215-9.00007-6.

Schiaffi, V., Barras, F. and Bouveret, E. (2024) 'Matching the β-oxidation gene repertoire with the wide diversity of fatty acids', *Current Opinion in Microbiology*, 77, p. 102402. Available at: https://doi.org/10.1016/j.mib.2023.102402.

Schluenzen, F. *et al.* (2000) 'Structure of Functionally Activated Small Ribosomal Subunit at 3.3 Å Resolution', *Cell*, 102(5), pp. 615–623. Available at: https://doi.org/10.1016/S0092-8674(00)00084-2.

Scholz, C.F.P. and Kilian, M. (2016) 'The natural history of cutaneous propionibacteria, and reclassification of selected species within the genus propionibacterium to the proposed novel genera acidipropionibacterium gen. Nov., cutibacterium gen. nov. and pseudopropionibacterium gen. nov.', *International Journal of Systematic and Evolutionary Microbiology*, 66(11), pp. 4422–4432. Available at: https://doi.org/10.1099/ijsem.0.001367.

Schulz, K.F. and Grimes, D.A. (2019) *Essential Concepts in Clinical Research,*. 2nd edn. Elsevier. Available at: https://www.clinicalkey.com/#!/content/book/3-s2.0-B9780702073946000055 (Accessed: 1 September 2023).

Schwartz, J.L. *et al.* (2021) 'Old age and other factors associated with salivary microbiome variation', *BMC Oral Health*, 21(1), p. 490. Available at: https://doi.org/10.1186/S12903-021-01828-1.

Sedghi, L. *et al.* (2021) 'The oral microbiome: Role of key organisms and complex networks in oral health and disease', *Periodontology 2000*, 87(1), pp. 107–131. Available at: https://doi.org/10.1111/prd.12393.

See-Too, W.S. *et al.* (2017) '*Pseudomonas versuta* sp. nov., isolated from Antarctic soil', *Systematic and Applied Microbiology*, 40(4), pp. 191–198. Available at: https://doi.org/10.1016/j.syapm.2017.03.002.

Selwitz, R.H., Ismail, A.I. and Pitts, N.B. (2007) 'Dental caries', *The Lancet*, 369(9555), pp. 51–59. Available at: https://doi.org/10.1016/S0140-6736(07)60031-2.

Serghiou, I.R., Webber, M.A. and Hall, L.J. (2023) 'An update on the current understanding of the infant skin microbiome and research challenges', *Current Opinion in Microbiology*, 75, p. 102364. Available at: https://doi.org/10.1016/j.mib.2023.102364.

Sfriso, R. *et al.* (2020) 'Revealing the secret life of skin - with the microbiome you never walk alone', *International Journal of Cosmetic Science*, 42(2), pp. 116–126. Available at: https://doi.org/10.1111/ICS.12594.

Shaffer, J.P. *et al.* (2022) 'A Comparison of Six DNA Extraction Protocols for 16S, ITS and Shotgun Metagenomic Sequencing of Microbial Communities', *BioTechniques*, 73(1), pp. 34–46. Available at: https://doi.org/10.2144/btn-2022-0032.

Sharma, G. *et al.* (2022) '*Prevotella*: An insight into its characteristics and associated virulence factors', *Microbial Pathogenesis*, 169, p. 105673. Available at: https://doi.org/10.1016/j.micpath.2022.105673.

Sheck, E. *et al.* (2023) 'Acinetobacter Non-baumannii Species: Occurrence in Infections in Hospitalized Patients, Identification, and Antibiotic Resistance', *Antibiotics*, 12(8), p. 1301. Available at: https://doi.org/10.3390/antibiotics12081301.

Shibagaki, N. *et al.* (2017) 'Aging-related changes in the diversity of women's skin microbiomes associated with oral bacteria', *Scientific Reports 2017 7:1*, 7(1), pp. 1–10. Available at: https://doi.org/10.1038/s41598-017-10834-9.

Shin, J.M. *et al.* (2016) 'Biomedical applications of nisin', *Journal of Applied Microbiology*, 120(6), pp. 1449–1465. Available at: https://doi.org/10.1111/JAM.13033.

Shoemark, D.K. and Allen, S.J. (2014) 'The Microbiome and Disease: Reviewing the Links between the Oral Microbiome, Aging, and Alzheimer's Disease', *Journal of Alzheimer's Disease*, 43(3), pp. 725–738. Available at: https://doi.org/10.3233/JAD-141170.

Siegwald, L. *et al.* (2019) 'The Impact of Bioinformatics Pipelines on Microbiota Studies: Does the Analytical "Microscope" Affect the Biological Interpretation?', *Microorganisms*, 7(10), p. 393. Available at: https://doi.org/10.3390/microorganisms7100393.

Simon Andrews (2023) *Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data*. Available at: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (Accessed: 10 November 2023).

Sinha, S., Lin, G. and Ferenczi, K. (2021) 'The skin microbiome and the gut-skin axis', *Clinics in Dermatology*, 39(5), pp. 829–839. Available at: https://doi.org/10.1016/j.clindermatol.2021.08.021.

Sipos, R. *et al.* (2007) 'Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targetting bacterial community analysis', *FEMS Microbiology Ecology*, 60(2), pp. 341–350. Available at: https://doi.org/10.1111/j.1574-6941.2007.00283.x.

Skowron, K. *et al.* (2021) 'Human skin microbiome: Impact of intrinsic and extrinsic factors on skin microbiota', *Microorganisms*, 9(3), pp. 1–20. Available at: https://doi.org/10.3390/microorganisms9030543.

Snider, E.J. *et al.* (2018) 'Barrett's esophagus is associated with a distinct oral microbiome', *Clinical and Translational Gastroenterology*, 9(3). Available at: https://doi.org/10.1038/S41424-018-0005-8.

Sogodogo, E. *et al.* (2019) 'First characterization of methanogens in oral cavity in Malian patients with oral cavity pathologies', *BMC oral health*, 19(1), p. 232. Available at: https://doi.org/10.1186/S12903-019-0929-8/FIGURES/1.

Stalsberg, R. and Pedersen, A.V. (2018) 'Are Differences in Physical Activity across Socioeconomic Groups Associated with Choice of Physical Activity Variables to Report?', *International Journal of Environmental Research and Public Health*, 15(5), p. 922. Available at: https://doi.org/10.3390/ijerph15050922.

Stefanie Peschel (2022) 'NetCoMi: Network Construction and Comparison for Microbiome Data'. Available at: https://netcomi.de.

Steffen, J. *et al.* (2021) 'Altered alcohol consumption during COVID-19 pandemic lockdown', *Nutrition Journal*, 20(1), p. 44. Available at: https://doi.org/10.1186/s12937-021-00699-0.

Sterndorff, E.B. *et al.* (2020) 'The T-shirt microbiome is distinct between individuals and shaped by washing and fabric type', *Environmental Research*, 185, p. 109449. Available at: https://doi.org/10.1016/J.ENVRES.2020.109449.

Stewart, A.G. *et al.* (2021) 'Completing the Picture—Capturing the Resistome in Antibiotic Clinical Trials', *Clinical Infectious Diseases*, 72(12), pp. e1122–e1129. Available at: https://doi.org/10.1093/CID/CIAA1877.

Sudhinaraset, M., Wigglesworth, C. and Takeuchi, D.T. (2016) 'Social and Cultural Contexts of Alcohol Use', *Alcohol Research : Current Reviews*, 38(1), pp. 35–45.

Sugita, T. and Cho, O. (2015) 'Significance of Changes in the Skin Fungal Microbiomes of Astronauts Staying on the International Space Station', *Journal of Disaster Research*, 10(6), pp. 1031–1034. Available at: https://doi.org/10.20965/JDR.2015.P1031.

Sukumar, S. *et al.* (2016) 'Metagenomic Insights into Transferable Antibiotic Resistance in Oral Bacteria', *Journal of Dental Research*, 95(9), pp. 969–976. Available at: https://doi.org/10.1177/0022034516648944.

Sun, Z. *et al.* (2023) 'Removal of false positives in metagenomics-based taxonomy profiling via targeting Type IIB restriction sites', *Nature Communications*, 14(1), p. 5321. Available at: https://doi.org/10.1038/s41467-023-41099-8.

Sureda, A. *et al.* (2020) 'Oral microbiota and Alzheimer's disease: Do all roads lead to Rome?', *Pharmacological Research*, 151(November 2019), p. 104582. Available at: https://doi.org/10.1016/j.phrs.2019.104582.

Szondi, D.C. *et al.* (2025) 'A role for arginase in skin epithelial differentiation and antimicrobial peptide production', *British Journal of Dermatology*, p. ljaf057. Available at: https://doi.org/10.1093/bjd/ljaf057.

Takeshita, T. *et al.* (2016) 'Bacterial diversity in saliva and oral health-related conditions: the Hisayama Study', *Scientific Reports*, 6, p. 22164. Available at: https://doi.org/10.1038/srep22164.

Tett, A. *et al.* (2021) 'Prevotella diversity, niches and interactions with the human host', *Nature reviews. Microbiology*, 19(9), pp. 585–599. Available at: https://doi.org/10.1038/s41579-021-00559-y.

Teufel, R. *et al.* (2010) 'Bacterial phenylalanine and phenylacetate catabolic pathway revealed', *Proceedings of the National Academy of Sciences of the United States of America*, 107(32), pp. 14390–14395. Available at: https://doi.org/10.1073/pnas.1005399107.

Thompson, L.A. and Chen, H. (2021) 'Physiology of Aging of Older Adults', *Dental Clinics of North America*, 65(2), pp. 275–284. Available at: https://doi.org/10.1016/j.cden.2020.11.002.

Tonelli, A., Lumngwena, E.N. and Ntusi, N.A.B. (2023) 'The oral microbiome in the pathophysiology of cardiovascular disease', *Nature Reviews Cardiology*, 20(6), pp. 386–403. Available at: https://doi.org/10.1038/s41569-022-00825-3.

Torres-Morales, J. *et al.* (2023) 'Site-specialization of human oral Gemella species', *Journal of Oral Microbiology*, 15(1), p. 2225261. Available at: https://doi.org/10.1080/20002297.2023.2225261.

Tourlousse, D.M. *et al.* (2017) 'Synthetic spike-in standards for high-throughput 16S rRNA gene amplicon sequencing', *Nucleic Acids Research*, 45(4), p. e23. Available at: https://doi.org/10.1093/nar/gkw984.

Tourlousse, D.M., Ohashi, A. and Sekiguchi, Y. (2018) 'Sample tracking in microbiome community profiling assays using synthetic 16S rRNA gene spike-in controls', *Scientific Reports*, 8(1), p. 9095. Available at: https://doi.org/10.1038/s41598-018-27314-3.

Tran, Q. and Phan, V. (2020) 'Assembling Reads Improves Taxonomic Classification of Species', *Genes*, 11(8), p. 946. Available at: https://doi.org/10.3390/genes11080946.

Trim, R.D. *et al.* (2011) 'Use of PCR to detect Entamoeba gingivalis in diseased gingival pockets and demonstrate its absence in healthy gingival sites', *Parasitology Research*, 109(3), pp. 857–864. Available at: https://doi.org/10.1007/S00436-011-2312-9/TABLES/3.

Tseng, Y.-C. *et al.* (2025) 'A human oral commensal-mediated protection against Sjögren's syndrome with maintenance of T cell immune homeostasis and improved oral microbiota', *npj Biofilms and Microbiomes*, 11(1), pp. 1–12. Available at: https://doi.org/10.1038/s41522-025-00654-5.

Van Hoogmoed, C.G. *et al.* (2008) 'Reduction of periodontal pathogens adhesion by antagonistic strains', *Oral Microbiology and Immunology*, 23(1), pp. 43–48. Available at: https://doi.org/10.1111/J.1399-302X.2007.00388.X.

Van Horn, K.G. *et al.* (2008) 'Comparison of 3 swab transport systems for direct release and recovery of aerobic and anaerobic bacteria', *Diagnostic Microbiology and Infectious Disease*, 62(4), pp. 471–473. Available at: https://doi.org/10.1016/j.diagmicrobio.2008.08.004.

Vandegrift, R. *et al.* (2017) 'Cleanliness in context: reconciling hygiene with a modern microbial perspective', *Microbiome 2017 5:1*, 5(1), pp. 1–12. Available at: https://doi.org/10.1186/S40168-017-0294-2.

Venturelli, R. *et al.* (2021) 'Use of AUDIT-C alcohol screening tool in NHS general dental practices in North London', *British Dental Journal* [Preprint]. Available at: https://doi.org/10.1038/s41415-021-2964-5.

Verbanic, S. *et al.* (2019) 'Improved single-swab sample preparation for recovering bacterial and phage DNA from human skin and wound microbiomes', *BMC Microbiology*, 19(1), p. 214. Available at: https://doi.org/10.1186/s12866-019-1586-4.

Verma, D., Garg, P.K. and Dubey, A.K. (2018) 'Insights into the human oral microbiome', *Archives of Microbiology*, 200(4), pp. 525–540. Available at: https://doi.org/10.1007/s00203-018-1505-3.

Vich Vila, A. *et al.* (2020) 'Impact of commonly used drugs on the composition and metabolic function of the gut microbiota', *Nature Communications*, 11, p. 362. Available at: https://doi.org/10.1038/s41467-019-14177-z.

Visontay, R. *et al.* (2020) 'Changes over time in young adults' harmful alcohol consumption: A cross-temporal meta-analysis using the AUDIT', *Drug and Alcohol Dependence*, 214, p. 108172. Available at: https://doi.org/10.1016/j.drugalcdep.2020.108172.

Vitetta, L., Chen, J. and Clarke, S. (2019) 'The vermiform appendix: An immunological organ sustaining a microbiome inoculum', *Clinical Science*, 133(1), pp. 1–8. Available at: https://doi.org/10.1042/CS20180956.

Wade, W.G. (2013) 'The oral microbiome in health and disease', *Pharmacological Research*, 69(1), pp. 137–143. Available at: https://doi.org/10.1016/J.PHRS.2012.11.006.

Wagner Mackenzie, B., Waite, D.W. and Taylor, M.W. (2015) 'Evaluating variation in human gut microbiota profiles due to DNA extraction method and inter-subject differences', *Frontiers in Microbiology*, 6. Available at: https://doi.org/10.3389/fmicb.2015.00130.

Walker, M. and McKinney-Freeman, S. (2019) 'Adult Hematopoietic Stem Cell Engagement with the Myeloablated Bone Marrow Niche', in R.L. Reis (ed.) *Encyclopedia of Tissue Engineering and Regenerative Medicine*. Oxford: Academic Press, pp. 221–227. Available at: https://doi.org/10.1016/B978-0-12-801238-3.65462-4.

Walker, M.Y. *et al.* (2018) 'Role of oral and gut microbiome in nitric oxide-mediated colon motility', *Nitric Oxide - Biology and Chemistry*, 73, pp. 81–88. Available at: https://doi.org/10.1016/j.niox.2017.06.003.

Walker, S.P. *et al.* (2020) 'Non-specific amplification of human DNA is a major challenge for 16S rRNA gene sequence analysis', *Scientific Reports*, 10(1), p. 16356. Available at: https://doi.org/10.1038/s41598-020-73403-7.

Wallen-Russell, C. (2018) 'The Role of Every-Day Cosmetics in Altering the Skin Microbiome: A Study Using Biodiversity', *Cosmetics*, 6(1), p. 2. Available at: https://doi.org/10.3390/cosmetics6010002.

Walsh, A.M. *et al.* (2018) 'Species classifier choice is a key consideration when analysing low-complexity food microbiome data', *Microbiome*, 6(1), p. 50. Available at: https://doi.org/10.1186/s40168-018-0437-0.

Walters, S.J. *et al.* (2017) 'Recruitment and retention of participants in randomised controlled trials: a review of trials funded and published by the United Kingdom Health Technology Assessment Programme', *BMJ Open*, 7(3), p. e015276. Available at: https://doi.org/10.1136/bmjopen-2016-015276.

Wang, L. *et al.* (2021) 'Facial Skin Microbiota-Mediated Host Response to Pollution Stress Revealed by Microbiome Networks of Individual', *mSystems*, 6(4). Available at: https://doi.org/10.1128/msystems.00319-21.

Wang, Q. and Cole, J.R. (2024) 'Updated RDP taxonomy and RDP Classifier for more accurate taxonomic classification', *Microbiology Resource Announcements*, 13(4), pp. e01063-23. Available at: https://doi.org/10.1128/mra.01063-23.

Ward, G. *et al.* (2023) 'Alcohol consumption and oral microbiome composition in a sample of healthy young adults', *Alcohol and Alcoholism*, 58(6), pp. 573–577. Available at: https://doi.org/10.1093/alcalc/agad048.

Waterman, S.R. and Small, P.L.C. (1998) 'Acid-Sensitive Enteric Pathogens Are Protected from Killing under Extremely Acidic Conditions of pH 2.5 when They Are Inoculated onto Certain Solid Food Sources', *Applied and Environmental Microbiology*, 64(10), pp. 3882–3886. Available at: https://doi.org/10.1128/AEM.64.10.3882-3886.1998.

Welch, J.L.M. *et al.* (2016) 'Biogeography of a human oral microbiome at the micron scale', *Proceedings of the National Academy of Sciences of the United States of America*, 113(6), pp. E791–E800. Available at: https://doi.org/10.1073/PNAS.1522149113/-/DCSUPPLEMENTAL/PNAS.201522149SI.PDF.

Wenger, K. *et al.* (2021) 'Factors associated with the counts of selected oral microorganisms in nursing home residents', *Special Care in Dentistry*, 41(1), pp. 32–40. Available at: https://doi.org/10.1111/SCD.12530.

Wetzel, D. and McBride, S.M. (2020) 'The Impact of pH on Clostridioides difficile Sporulation and Physiology', *Applied and Environmental Microbiology*, 86(4). Available at: https://doi.org/10.1128/AEM.02706-19.

Weyrich, L.S. *et al.* (2019) 'Laboratory contamination over time during low-biomass sample analysis', *Molecular Ecology Resources*, 19(4), pp. 982–996. Available at: https://doi.org/10.1111/1755-0998.13011.

Wickham, H. *et al.* (2019) 'Welcome to the Tidyverse', *Journal of Open Source Software*, 4(43), p. 1686. Available at: https://doi.org/10.21105/joss.01686.

Wilsnack, R.W. *et al.* (2009) 'Gender and alcohol consumption: patterns from the multinational GENACIS project', *Addiction*, 104(9), pp. 1487–1500. Available at: https://doi.org/10.1111/j.1360-0443.2009.02696.x.

Wilson, J.A. *et al.* (2023) 'Conservative management versus tonsillectomy in adults with recurrent acute tonsillitis in the UK (NATTINA): a multicentre, open-label, randomised controlled trial', *The Lancet*, 401(10393), pp. 2051–2059. Available at: https://doi.org/10.1016/S0140-6736(23)00519-6.

Wilson, J.W. *et al.* (2002) 'Mechanisms of bacterial pathogenicity', *Postgraduate Medical Journal*, 78(918), pp. 216–224. Available at: https://doi.org/10.1136/pmj.78.918.216.

Wilson, M. (2008) *Bacteriology of Humans an Ecological Perspective*. Blackwell Publishing Ltd.

Wong, J.D. and Janda, J.M. (1992) 'Association of an important Neisseria species, Neisseria elongata subsp. nitroreducens, with bacteremia, endocarditis, and osteomyelitis', *Journal of Clinical Microbiology*, 30(3), pp. 719–720. Available at: https://doi.org/10.1128/jcm.30.3.719-720.1992.

Woo, J., Leung, J. and Morley, J.E. (2012) 'Comparison of Frailty Indicators Based on Clinical Phenotype and the Multiple Deficit Approach in Predicting Mortality and Physical Limitation', *Journal of the American Geriatrics Society*, 60(8), pp. 1478–1486. Available at: https://doi.org/10.1111/j.1532-5415.2012.04074.x.

Worthington, H.V. *et al.* (2019) 'Home use of interdental cleaning devices, in addition to toothbrushing, for preventing and controlling periodontal diseases and dental caries', *Cochrane Database of Systematic Reviews* [Preprint], (4). Available at: https://doi.org/10.1002/14651858.CD012018.pub2.

Wu, R.Q. *et al.* (2014) 'The mucosal immune system in the oral cavity-an orchestra of T cell diversity', *International Journal of Oral Science*, 6(3), pp. 125–132. Available at: https://doi.org/10.1038/ijos.2014.48.

Xu, Q. *et al.* (2025) 'The oral-gut microbiota axis: a link in cardiometabolic diseases', *npj Biofilms and Microbiomes*, 11(1), pp. 1–13. Available at: https://doi.org/10.1038/s41522-025-00646-5.

Yamada, Y. *et al.* (2018) 'The Effect of Improving Oral Hygiene through Professional Oral Care to Reduce the Incidence of Pneumonia Post-esophagectomy in Esophageal Cancer', *The Keio Journal of Medicine*, 68(1), pp. 17–25. Available at: https://doi.org/10.2302/kjm.2017-0017-OA.

Yanagawa, R. and Honda, E. (1978) 'Corynebacterium pilosum and Corynebacterium cystitidis, Two New Species from Cows', *International Journal of Systematic Bacteriology*, 28(2), pp. 209–216. Available at: https://doi.org/10.1099/00207713-28-2-209.

Yang, I. *et al.* (2021) 'The oral microbiome and inflammation in mild cognitive impairment', *Experimental Gerontology*, 147(February), p. 111273. Available at: https://doi.org/10.1016/j.exger.2021.111273.

Yang, S. *et al.* (2025) 'Associations between the oral microbiome, number of teeth and frailty among American adults: A cross-sectional study from NHANES 2009–2012', *Experimental Gerontology*, 203, p. 112727. Available at: https://doi.org/10.1016/j.exger.2025.112727.

Yang, S.F. *et al.* (2022) 'Host Genetic Associations with Salivary Microbiome in Oral Cancer', *Journal of Dental Research*, 101(5), pp. 590–598. Available at: https://doi.org/10.1177/00220345211051967.

Ye, X. *et al.* (2017) 'Fusobacterium Nucleatum Subspecies Animalis Influences Proinflammatory Cytokine Expression and Monocyte Activation in Human Colorectal Tumors', *Cancer Prevention Research*, 10(7), pp. 398–409. Available at: https://doi.org/10.1158/1940-6207.CAPR-16-0178.

Yokoro, M. *et al.* (2023) 'Associations between Dietary Variety and Frailty in Community-Dwelling Older People Who Live Alone: Gender Differences', *The Journal of Frailty & Aging*, 12(1), pp. 30–36. Available at: https://doi.org/10.14283/jfa.2021.49.

Zachary Kurtz *et al.* (2025) 'SpiecEasi: Sparse Inverse Covariance for Ecological Statistical Inference'. Available at: https://github.com/zdk123/SpiecEasi.

Zaremba, M.L. *et al.* (2006) 'Incidence rate of Candida species in the oral cavity of middle-aged and elderly subjects', *Advances in Medical Sciences*, 51 Suppl 1, pp. 233–236.

Zargar, K. *et al.* (2016) 'In vitro Characterization of Phenylacetate Decarboxylase, a Novel Enzyme Catalyzing Toluene Biosynthesis in an Anaerobic Microbial Community', *Scientific Reports*, 6(1), p. 31362. Available at: https://doi.org/10.1038/srep31362.

Zasloff, M. (2002) 'Innate immunity, antimicrobial peptides, and protection of the oral cavity', *The Lancet*, 360(9340), pp. 1116–1117. Available at: https://doi.org/10.1016/S0140-6736(02)11239-6.

Zaura, E. *et al.* (2021) 'Optimizing the quality of clinical studies on oral microbiome: A practical guide for planning, performing, and reporting', *Periodontology 2000*, 85(1), pp. 210–236. Available at: https://doi.org/10.1111/prd.12359.

Zell, J.A. *et al.* (2008) 'Survival for patients with invasive cutaneous melanoma among ethnic groups: The effects of socioeconomic status and treatment', *Journal of Clinical Oncology*, 26(1), pp. 66–75. Available at: https://doi.org/10.1200/JCO.2007.12.3604.

Zhang, Z. *et al.* (2020) 'Towards a General Understanding of Bacterial Interactions', *Trends in Microbiology*, 28(10), pp. 783–785. Available at: https://doi.org/10.1016/j.tim.2020.05.010.

Zhong, S.R. *et al.* (2021) 'Functional roles of the microbiota-gut-brain axis in Alzheimer's disease: Implications of gut microbiota-targeted therapy', *Translational Neuroscience*, 12(1), pp. 581–600. Available at: https://doi.org/10.1515/TNSCI-2020-0206.

Zhou, P. *et al.* (2021) 'Veillonellae: Beyond Bridging Species in Oral Biofilm Ecology', *Frontiers in Oral Health*, 2. Available at: https://doi.org/10.3389/froh.2021.774115.

Zhou, W. *et al.* (2023) 'Skin microbiome attributes associate with biophysical skin ageing', *Experimental Dermatology*, 32(9), pp. 1546–1556. Available at: https://doi.org/10.1111/exd.14863.

Zhou, X. and Li, Y. (eds) (2020) *Atlas of Oral Microbiology: From Healthy Microflora to Disease*. Singapore: Springer Singapore. Available at: https://doi.org/10.1007/978-981-15-7899-1.

Chapter 9. Appendix

9.1. Supplemental Material for Chapter 2

Table 9.1 Observations of colonies and phenotypic test results obtained from skin sampling methods on different growth media. *Tests were only carried out on colonies with unique morphology when first encountered.*

Sampling method	Media	Sample ID	Colony morphology	Other descriptive features	Count	Gram stain	Microscopic morphology	Catalase	Oxidase	Presumptive ID
Skin Swab	MAC3	PS01	No Growth							
		PS02	Circular, possibly umbonate	Semi-transparent	2	-ve	Poor smear so no observations	+ve	-ve	
		PS03	No Growth							
		PS04	No Growth							
		PS05	No Growth							
		PS06	No Growth							
		PS07	No Growth							
		PS08	No Growth							

		PS09	No Growth							
		PS10	No Growth							
	MAN	PS01	No Growth							
		PS02	No Growth							
		PS03	Small, circular, raised	Cream coloured	1			Inconclusive	Inconclusive	
		PS04	No Growth							
		PS05	Circular, raised	Cream coloured	5	+ve	Rod- connected	+ve	-ve	Bacillus
		PS06	No Growth							
		PS07	No Growth							
		PS08	No Growth							
		PS09	No Growth							
		PS10	No Growth							

	CX SA	PS01	No Growth							
		PS02	No Growth							
		PS03	No Growth							
		PS04	No Growth							
		PS05	Circular, raised	Blue	3	+ve	rod	-ve	-ve	Bacillus?
		PS06	No Growth							
		PS07	No Growth							
		PS08	No Growth							
		PS09	No Growth							
		PS10	No Growth							
	SDA	PS01	No Growth							
		PS02	Filamentous, raised	White, furry	1			Inconclusive	-ve	

		PS03	Filamentous, raised	White with green centre	1		Poor smear so no observations	-ve	-ve	Fungi
		PS04	No Growth							
		PS05	No Growth							
		PS06	No Growth							
		PS07	No Growth							
		PS08	No Growth							
		PS09	No Growth							
		PS10	No Growth							
	CBA	PS01	No Growth							
		PS02	No Growth							
		PS03	No Growth							
		PS04	No Growth							

		PS05	Circular, flat	beta	1	+ve	Rods	+ve	-ve	Bacilli
			Circular flat	alpha	1					
			Filamentous	Alpha/beta	1					Fungi
		PS06	No Growth							
		PS07	No Growth							
		PS08	Irregular, raised	White, Gama haemolysis, might be plate I had to flip.	1-ish	+ve	Bacilli	+ve	+ve	Bacillus
		PS09	Round, raised, small	White	2			Inconclusive		Staph
		PS10	Round, raised, small	White	1			Inconclusive		Staph
Skin Scrape	MAC3	PS01	No Growth							
		PS02	No Growth							
		PS03	No Growth							

		PS04	No Growth							
		PS05	No Growth							
		PS06	No Growth							
		PS07	No Growth							
		PS08	No Growth							
		PS09	No Growth							
		PS10	No Growth							
	MAN	PS01	No Growth							
		PS02	No Growth							
		PS03	No Growth							
		PS04	No Growth							
		PS05	No Growth							

		PS06	No Growth							
		PS07	No Growth							
		PS08	No Growth							
		PS09	No Growth							
		PS10	Irregular, raised	Same colour as medium	1	+ve	Bacili, in chains	+ve	-ve	Bacillus
	CX SA	PS01	No Growth							
		PS02	No Growth							
		PS03	No Growth							
		PS04	No Growth							
		PS05	Circular, raised	Blue	1			-ve	-ve	
			Circular, raised	Cream coloured	1	+ve	Rod-bunched	+ve	-ve	Staph
		PS06	No Growth							

		PS07	No Growth							
		PS08	No Growth							
		PS09	No Growth							
		PS10	No Growth							
	SDA	PS01	No Growth							
		PS02	No Growth							
		PS03	No Growth							
		PS04	No Growth							
		PS05	No Growth							
		PS06	No Growth							
		PS07	No Growth							
		PS08	No Growth							

		PS09	No Growth							
		PS10	No Growth							
	CBA	PS01	Iregular, flat	beta	1	+ve	Rods	+ve	-ve	Bacilli
		PS02								
		PS03								
		PS04	Circular, flat	alpha	3			+ve	-ve	
			irregular	beta	1	+ve	Rods, with halos of stuff round them likely dead rods	+ve	-ve	Bacilli
		PS05								
		PS06								
		PS07								

		PS08	Round, flat	White, beta haemolysis	1	+ve	Cocci, clustered	+ve	-ve	Staph
			Round, umbonate	Cream, Gama haemolysis	1			+ve	-ve	
		PS09	Round, raised, small	White	1			Inconclusive		Staph
		PS10								

Table 9.2 Observations of colonies and phenotypic test results obtained from oral sampling methods on different growth media. *Tests were only carried out on colonies with unique morphology when first encountered.*

Sampling method	Media	Sample ID	Colony morphology	Other descriptive features	Count	Gram stain	Microscopic morphology	Catalase	Oxidase	Presumptive ID
Oral Rinse	MAC3	PS01	Small, irregular, crateriform	Pink/purple	108			-ve	-ve	E.coli/Kleb?
		PS02	Circular, raised	Non-Lactose fermenter	70	-ve	Poor smear so no observations	+ve	-ve	A.baumanii?
			Circular, possibly umbonate	Semi-transparent	2		Poor smear so no observations	+ve	-ve	
		PS03	No Growth							
		PS04	Small, irregular, crateriform	Pink/purple, Lactose fermenter	135		Poor smear so no observations	-ve	-ve	E.coli/Kleb?

		PS05	No Growth							
		PS06	No Growth							
		PS07	No Growth							
		PS08	No Growth							
		PS09	Round, raised	Purple shade	1	-ve	Cocci paired	+ve	-ve	Veillonella?
			Round, raised	Non-lactose fermenting	30			+ve	-ve	
		PS10	Irregular, small	purple	5		Won't stain	-ve	-ve	E.coli/kleb
	MAN	PS01	Circular, yellow, raised	Golden halo round colony	165	+ve	Cocci	+ve	-ve	Staph
			Filamentous, convex	White, visibly furry	2	-ve	Na fungi	-ve	-ve	
		PS02	No Growth							
		PS03	Circular, yellow, raised	dehydrated plate	1			+ve	Inconclusive	
		PS04	Circular, yellow, raised	Golden halo round colony	7			+ve	-ve	

		PS05	No Growth							
		PS06	No Growth							
		PS07	No Growth							
		PS08	Round, raised	White with golden halo	1					
			Round, raised	Yellow with golden halo	3					
		PS09	No Growth							
		PS10	Round, raised	White with golden halo	3	+ve	Cocci, clustered	+ve	-ve	Staph
			Round, raised	Yellow with golden halo	4	+ve	Cocci	+ve	-ve	Staph
			Round, raised	White without halo	3	+ve	Cocci	+ve	-ve	Staph
	CX CAND	PS01	No Growth							
		PS02	Circular, raised	purple	3			-ve	Inconclusive	
		PS03	Circular, raised	Green	1			-ve	Inconclusive	C. albicans

		PS04	No Growth							
		PS05	No Growth							
		PS06	Round, Raised	Green	16					C.albicans
			Round, raised	Pink/purple	1					C.glabrata
		PS07	No Growth							
		PS08	No Growth							
		PS09	No Growth							
		PS10	No Growth							
	SDA	PS01	Circular, raised	Yellow	20			+ve	-ve	
		PS02	Circular, raised	White	16	+ve	Cocci	+ve	-ve	Staph
			Circular, raised	yellow	19	+ve	Rods	+ve	-ve	Staph, or bacillus
		PS03	No Growth							

		PS04	Circular, raised	white	1			+ve	-ve	
			Circular, raised	yellow	4			+ve	-ve	
		PS05	No Growth							
		PS06	Round, convex, large	white	7					
			Round, raised, small	white	TMTC					
		PS07	Round, raised, small	White	TMTC					
		PS08	Round, raised, small	White	TMTC					
		PS09	No Growth							
		PS10	Round, raised, small	white	TMTC					
			Round, raised	yellow	TMTC	-ve	Bacilli	+ve	+ve	Unknown negative
			Irregular, raised,	White/cream	3	+ve	Cocci/coccibacili / paired cocci	+ve	-ve	Unknown positive
	Strep	PS01	Lawn colonising entire plate	Alpha haemolysis	TMTC	+ve	Cocci/coccobacilli			Strep

		PS02	Lawn colonising entire plate	Alpha haemolysis	TMTC			-ve	-ve	
		PS03	Lawn colonising entire plate	Alpha haemolysis	TMTC					
		PS04	Lawn colonising entire plate	Alpha haemolysis	TMTC					
		PS05	Lawn colonising entire plate	Alpha haemolysis	TMTC					
		PS06	Round, raised, small	White/grey, Alpha haemolysis	TMTC					
		PS07	Round, raised, small	White/grey, Alpha haemolysis	TMTC					
		PS08	Round, raised, small	White/grey, Alpha haemolysis	TMTC					
		PS09	Round, raised, small	White/grey, Alpha haemolysis	TMTC					
		PS10	Round, raised, small	White/grey, Alpha haemolysis	TMTC					
	FAA	PS01	Small, circular, raised	Beta haemolysis Growth round disc	TMTC					
		PS02	Small, circular, raised	Beta haemolysis Growth round disc	TMTC					
		PS03	No Growth							

		PS04	Small, circular, raised	Beta haemolysis Growth round disc	TMTC					
		PS05	Small, circular, raised	Beta haemolysis Growth round disc	TMTC					
		PS06	Round, raised, small	White, Grey, Beta haemolysis?	TMTC					
		PS07	Round, raised, small	White, Grey, Beta haemolysis?	TMTC					
		PS08	Round, raised, small	White, Grey, Beta haemolysis?	TMTC					
		PS09	Round, raised, small	White, Grey, Beta haemolysis	TMTC					
		PS10	Round, raised, small	White, Grey, Beta haemolysis	TMTC					
	CBA	PS01	Lawn, circular, raised	Alpha/beta	TMTC					
		PS02	Lawn, circular, raised	Alpha/beta	TMTC					
		PS03	Lawn, circular, raised	Alpha/beta	TMTC					
		PS04	Lawn, circular, raised	Alpha/beta	TMTC					
		PS05	Lawn, small, circular raised	Alpha haemolysis	TMTC					

		PS06	Round, raised, small	White/grey, Alpha haemolysis	TMTC	+ve	Bacili/coccibacili	-ve	-ve	Bacillus
			Irregular, raised,	Yellow/white,	TMTC	-ve	Cocci	+ve	+ve	Neisseria
		PS07	Round, raised, small	White/grey, Alpha haemolysis	TMTC					
			Irregular, raised,	Yellow/white,	12					
		PS08	Round, raised, small	White/grey, Alpha haemolysis	TMTC					
			Irregular, raised,	Yellow/white,	8					
		PS09	Round, raised, small	White/grey, Alpha haemolysis	TMTC					
			Irregular, raised,	Yellow/white,	2					
		PS10	Round, raised, small	White/grey, Alpha haemolysis	TMTC					
			Irregular, raised,	Yellow/white,	12					
Oral Swab	MAC3	PS01	No Growth							
		PS02	No Growth							

		PS03	No Growth							
		PS04	No Growth							
		PS05	No Growth							
		PS06	No Growth							
		PS07	No Growth							
		PS08	No Growth							
		PS09	No Growth							
		PS10	Round, raised	Non-lactose fermenting	1	-ve	Baccilli Or chains of cocci	+ve	-ve	A.baumannii ?
	MAN	PS01	No Growth							
		PS02	No Growth	Circular, raised, purple	1			-ve	Inconclusive	c. glabrata
		PS03	No Growth							
		PS04	No Growth							

		PS05	No Growth							
		PS06	No Growth							
		PS07	Round, raised	White with golden halo	2					
		PS08	No Growth							
		PS09	No Growth							
		PS10	Round, raised	White with golden halo	1					
	CX CAND	PS01	No Growth							
		PS02	No Growth							
		PS03	No Growth							
		PS04	No Growth							
		PS05	No Growth							
		PS06	Round, raised	Green	18					C.albicans

			Round, raised	Pink/purple	3					C.glabrata
		PS07	No Growth							
		PS08	No Growth							
		PS09	No Growth							
		PS10	No Growth							
	SDA	PS01	No Growth							
		PS02	No Growth							
		PS03	No Growth							
		PS04	No Growth							
		PS05	No Growth							
		PS06	Round, convex, large	white	TMTC	NA	Fungi, spores	+ve	-ve	candida
			Round, raised, small	white	TMTC	+ve	Cocci	-ve	-ve	Strep

			Filamentous	White edge with green centre	1					Fungi
		PS07	Round, raised, small	White	TMTC					
		PS08	Round, raised, small	White	TMTC					
		PS09	Round, raised	White/cream	13					
			Round, raised, small	white	TMTC					
			Round, raised	yellow	1					
		PS10	Round, raised, small	white	TMTC					
	Strep	PS01	Small, circular, raised	Alpha haemolysis	TMTC					
		PS02	Small, circular, raised	Alpha haemolysis	45			-ve	-ve	
		PS03	No Growth							
		PS04	Small, circular, raised	Alpha haemolysis	TMTC					
		PS05	Small, circular, raised	Alpha haemolysis	TMTC					

		PS06	Round, raised, small	White/grey, Alpha haemolysis	TMT C	+ve	Cocci	-ve	-ve	Strep
		PS07	Round, raised, small	White/grey, Alpha haemolysis	TMT C					
		PS08	Round, raised, small	White/grey, Alpha haemolysis	TMT C					
		PS09	Round, raised, small	White/grey, Alpha haemolysis	TMT C					
		PS10	Round, raised, small	White/grey, Alpha haemolysis	TMT C					
	FAA	PS01	Small, circular, raised	Beta haemolysis Growth round disc	TMT C					
		PS02	Small, circular, raised	Beta haemolysis Growth round disc	60	+ve	Cocci	-ve	-ve	Strep
		PS03	Small, circular, raised	Growth away from disc	1					
		PS04	Small, circular, raised	Beta haemolysis Growth round disc	TMT C					
		PS05	Small, circular, raised	Beta haemolysis Growth round disc	TMT C					
		PS06	Round, raised, small	White, Grey, Beta haemolysis?	TMT C					
		PS07	Round, raised, small	White, Grey, Beta haemolysis?	TMT C					

		PS08	Round, raised, small	White, Grey, Beta haemolysis	TMTC					
		PS09	Round, raised, small	White, Grey, Beta haemolysis	TMTC	-ve/+ve	Cocci	-ve	-ve	Strep or multiple
		PS10	Round, raised, small	White, Grey, Beta haemolysis	TMTC					
	CBA	PS01	Small, circular, raised	Beta/alpha haemolysis	TMTC	+ve	Cocci, singular	+ve	-ve	Staph
		PS02	Small, circular, raised	Beta/alpha haemolysis	130					
		PS03	Irregular craterform	White, alpha haemolysis	9	+ve	Cocci, grouped	-ve	-ve	unknown
		PS04	Small, circular, raised	Beta/alpha haemolysis	TMTC					
		PS05	Small, circular, raised	Beta Haemolysis	TMTC					
		PS06	Round, raised, small	White/grey, Alpha haemolysis	TMTC					
		PS07	Round, raised, small	White/grey, Alpha haemolysis	TMTC					
			Irregular, raised,	Yellow/white,	10					
		PS08	Round, raised, small	White/grey, Alpha haemolysis	TMTC					

		PS09	Round, raised, small	White/grey, Alpha haemolysis	TMTC					
			Round, raised, small	White/grey, Alpha haemolysis	TMTC					
		PS10	Irregular, raised,	Yellow/white,	20					

9.2. Supplemental Material for Chapter 3

9.2.1. Participant Questionnaires

9.2.1.1. Recruitment Questionnaire

Participant screening

Start of Block: Basic information

1 What is your first name?

_____

1 What is your last name?

_____

*

2 What is your contact email address?

_____

*

3 What is your birth-date?

_____

4 What is your sex?

- ○ Male (1)
- ○ Female (2)
- ○ Non-binary / third gender (3)
- ○ Prefer not to say (4)

*

5 What is the postcode of your permanent address?

_____

6 How would you describe your ethnicity?

- ○ English/Welsh/Scottish/Northern Irish/British (1)
- ○ Irish (2)
- ○ Gypsy or Irish Traveler (3)
- ○ Any other White background, please describe (4) _____
- ○ White and Black Caribbean (5)
- ○ White and Black African (6)
- ○ White and Asian (7)
- ○ Any other Mixed/Multiple ethnic background, please describe (8) _____
- ○ Indian (9)
- ○ Pakistani (10)
- ○ Bangladeshi (11)
- ○ Chinese (12)
- ○ Any other Asian background, please describe (13) _____
- ○ African (14)
- ○ Caribbean (15)
- ○ Any other Black/African/Caribbean background, please describe (16) _____
- ○ Arab (17)
- ○ Any other ethnic group, please describe (18) _____

7 Are you a current member of Aston University's student or staff body?

- Yes (1)
- No (2)

Display This Question:
If Are you a current member of Aston University's student or staff body? = No

7a Are you capable of traveling to Aston University by your own means?

- Yes (1)
- No (2)

Display This Question:
If Are you capable of traveling to Aston University by your own means? = No

7b Do you live within a 60 minute drive of Aston University?

- Yes (1)
- No (2)
- It's just over 60 minutes, Please provide a time in minutes. (3) _____

End of Block: Basic information

Participant health metadata

Start of Block: Default Question Block

1 Participant code

End of Block: Default Question Block

Start of Block: Overview of health

2 Do you suffer from any of the following conditions

- [ ] Arthritis (1)
- [ ] Dementia/Alzheimer's (2)
- [ ] Diabetes type 1 (3)
- [ ] Diabetes type 2 (4)
- [ ] High blood pressure/heart disease (5)
- [ ] Asthma (6)
- [ ] Chronic bronchitis (7)
- [ ] Parkinson's (8)
- [ ] Stroke (9)
- [ ] Depression (10)
- [ ] Huntington's (11)
- [ ] Other, please describe (12) _____
- [ ] Prefer not to say (13)
- [ ] None (14)

D. J. Newman, PhD Thesis, Aston University 2025.

3 Do you smoke or vape, or have you given up within the last 6 months? - if yes how many per day?

- Yes - Smoke (1) _____
- Yes - Vape (2)
- No (3)
- Given up within the last 6 months (4) _____
- Prefer not to say (5)

---

4 Do you drink alcohol or have you regularly drunk but given it up recently? - if yes how many units per week?

- Yes (1) _____
- No (2)
- Recently given up (3) _____
- Prefer not to say (4)

Display This Question:

If Do you drink alcohol or have you regularly drunk but given it up recently? - if yes how many unit... = Yes

Or Do you drink alcohol or have you regularly drunk but given it up recently? - if yes how many unit... = Recently given up

5 Did you ever receive medical treatment or interventions from the amount you consumed?

- No (1)
- Prefer not to say (2)
- Yes (3)

---

6 Are you actively taking any antibiotic or have been within the past 6 weeks? - if yes do you know which you were prescribed.

- ○ Yes (1) _____
- ○ Prefer not to say (2)
- ○ No (3)

---

7 Are you undergoing any form of hormone treatment?

- ○ Yes (1)
- ○ Prefer not to say (2)
- ○ No (3)

---

8 Are you currently considered immune suppressed (e.g. due to cancer treatments or actively taking immune suppressants etc.)?

- ○ Yes (1)
- ○ Prefer not to say (2)
- ○ No (3)

---

9 Are you for any reason (e.g. personal or religious) following any particular food restricted diet, such as vegetarian, dairy free etc. and what is it?

- ○ Yes (1) _____
- ○ Prefer not to say (2)
- ○ No (3)

---

10 Would you consider yourself an active person? (e.g. do you go outside regularly for exercise or hobbies) - if yes does this involve swimming in either a pool or in open water?

- Yes - no swimming (1)
- Yes - swimming in a pool (2)
- Yes - swimming in open water (3)
- No (4)
- Prefer not to say (5)

End of Block: Overview of health

Start of Block: Dental and skin health

11 Do you have any fillings and have any been within the last 6 months?

- No (1)
- Prefer not to say (2)
- Yes but none within the last 6 months (3)
- Yes including some within the last 6 months (4)

12 Aside from your wisdom teeth have you had any of your adult teeth removed?

- Yes (1)
- Prefer not to say (2)
- No (3)

13 Do you have any of the following? Tick all appropriate

- ○ Implants (1)
- ○ Root canals (2)
- ○ Crowns (3)
- ○ Bridge work (4)
- ○ Dentures (5)
- ○ Retainer (6)
- ○ Braces (7)
- ○ Other permanent/semipermanent artificial implant e.g. piercing please describe (8) _____
- ○ No (9)
- ○ Prefer not to say (10)

---

14 Have you had your tonsils removed

- ○ No (1)
- ○ Prefer not to say (2)
- ○ Yes (3)

Page Break

15 Do you currently have any skin condition? - if yes what is it?

- Yes (1) _____
- Yes not sure which type (2)
- Prefer not to say (3)
- No (4)

---

16 Have you suffered from Acne, or other skin conditions in the past?

- Yes acne (1)
- Yes other (2) _____
- No (3)
- Prefer not to say (4)

---

17 Have you been prescribed any topical creams or treatments for skin conditions within the past 6 weeks? - if yes do you know what type.

- Yes (1) _____
- Prefer not to say (2)
- No (3)

---

18 Are there any tattoos, implants or other topical features (e.g. scar tissue) of the area about to be sampled? - if yes have they occurred within the last 4 weeks?

- ○ Yes within the last 4 weeks (1)
- ○ Yes but more than 4 weeks ago (2)
- ○ No (3)
- ○ Prefer not to say (4)

End of Block: Dental and skin health

Participant hygiene routines metadata

Start of Block: Default Question Block

1 Participant code

End of Block: Default Question Block

Start of Block: Dental hygiene routine

2
Do you brush you teeth?

- Yes (1)
- Yes but also has dentures (4)
- Prefer not to say (2)
- No (3)

Display This Question:
If Do you brush you teeth? = Yes

3
How often per day?

- 1 (1)
- 2 (2)
- 3 (3)
- 4+ (4)
- Prefer not to say (5)

Display This Question:
If Do you brush you teeth? = Yes

4 Do you use an electric or manual tooth brush?

- Electric (1)
- Manual (2)
- Prefer not to say (3)

Display This Question:
If Do you brush you teeth? = Yes

5 Which brand/product of toothpaste do you use?

_____

Display This Question:
If Do you brush you teeth? = No
And Do you brush you teeth? = Yes but also has dentures

6 If this is due to dentures if so which brand/product of denture cleaners do you use?

- ○ Dentures (1) _____
- ○ No (2)
- ○ Other (3) _____
- ○ Prefer not to say (4)

---

7 Do you use mouthwash? if yes which brand/product do you use?

- ○ Yes (1) _____
- ○ Prefer not to say (2)
- ○ No (3)

---

8 Do you floss or use interdentin brushes?

- ○ Floss (1)
- ○ Interdentin brush (2)
- ○ Prefer not to say (3)
- ○ No (4)

Display This Question:

If Do you floss or use interdentin brushes? = Floss

Or Do you floss or use interdentin brushes? = Interdentin brush

9 How often per day?

- 1 (1)
- 2 (2)
- 3 (3)
- 4+ (4)
- Prefer not to say (5)

---

10 Have you visited a dental hygienist in the last 6 months?

- No (1)
- Prefer not to say (2)
- Yes (3)

End of Block: Dental hygiene routine

Start of Block: Skin hygiene routine

11 Do you shower take a bath or perform some other kind of skin hygiene routine regularly?

- Shower (1)
- Bath (2)
- Other please describe (3) _____
- No (4)
- Prefer not to say (5)

Display This Question:

If Do you shower take a bath or perform some other kind of skin hygiene routine regularly? = Shower

Or Do you shower take a bath or perform some other kind of skin hygiene routine regularly? = Bath

Or Do you shower take a bath or perform some other kind of skin hygiene routine regularly? = Other please describe

12 How often per week?

- 1 (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (5)
- 6 (6)
- 7 (7)
- 8 (8)
- 9 (9)
- 10+ (10)
- Prefer not to say (11)

Display This Question:

If Do you shower take a bath or perform some other kind of skin hygiene routine regularly? = Shower

Or Do you shower take a bath or perform some other kind of skin hygiene routine regularly? = Bath

Or Do you shower take a bath or perform some other kind of skin hygiene routine regularly? = Other please describe

13 Do you use any shower gel/body wash, hard soap, or other detergent? if yes which brand/product do you use?

- ○ Shower gel/ body wash (1) _____
- ○ Hard soap (2) _____
- ○ Other (3) _____
- ○ Prefer not to say (4)
- ○ No (5)

---

14 Do you apply moisturiser or another skin cream regularly? if yes which brand/product do you use?

- ○ Yes (1) _____
- ○ No (2)
- ○ Prefer not to say (3)

End of Block: Skin hygiene routine

9.2.2. Definition Table

Table 9.3 Glossary of descriptive terms assigned to various sociodemographic codes.

Demographic Classification System	Code	Descriptor
Output Area Code	***1***	***Rural residents***
	1A	**Farming communities**
	1A1	Rural workers and families
	1A2	Established farming communities
	1A3	Agricultural communities
	1A4	Older farming communities
	1B	**Rural tenants**
	1B1	Rural life
	1B2	Rural white-collar workers
	1B3	Rural white-collar workers
	1C	**Ageing rural dwellers**
	1C1	Rural employment and retirees
	1C2	Renting rural retirement
	1C3	Detached rural retirement
	2	***Cosmopolitans***
	2A	**Students around campus**
	2A1	Student communal living
	2A2	Student digs
	2A3	Students and professionals
	2B	**Inner city students**
	2B1	Students and commuters
	2B2	Multicultural student neighbourhoods
	2C	**Comfortable cosmopolitan**
	2C1	Migrant families
	2C2	Migrant commuters
	2C3	Professional service cosmopolitans
	2D	**Aspiring and affluent**
	2D1	Urban cultural mix
	2D2	Highly-qualified quaternary workers
	2D3	EU white-collar workers
	3	***Ethnicity central***
	3A	**Ethnic family life**
	3A1	Established renting families
	3A2	Young families and students
	3B	**Endeavouring Ethnic Mix**
	3B1	Striving service workers
	3B2	Bangladeshi mixed employment
	3B3	Multi-ethnic professional service workers
	3C	**Ethnic dynamics**
	3C1	Constrained neighbourhoods
	3C2	Constrained commuters
	3D	**Aspirational techies**
	3D1	New EU tech workers
	3D2	Established tech workers

	3D3	Old EU tech workers
	4	***Multicultural metropolitans***
	4A	**Rented family living**
	4A1	Social renting young families
	4A2	Private renting new arrivals
	4A3	Commuters with young families
	4B	**Challenged Asian terraces**
	4B1	Asian terraces and flats
	4B2	Pakistani communities
	4C	**Asian traits**
	4C1	Achieving minorities
	4C2	Multicultural new arrivals
	4C3	Inner city ethnic mix
	5	***Urbanites***
	5A	**Urban professionals and families**
	5A1	White professionals
	5A2	Multi-ethnic professionals with families
	5A3	Families in terraces and flats
	5B	**Ageing urban living**
	5B1	Delayed retirement
	5B2	Communal retirement
	5B3	Self-sufficient retirement
	6	***Suburbanites***
	6A	**Suburban achievers**
	6A1	Indian tech achievers
	6A2	Comfortable suburbia
	6A3	Detached retirement living
	6A4	Ageing in suburbia
	6B	**Semi-detached suburbia**
	6B1	Multi-ethnic suburbia
	6B2	White suburban communities
	6B3	Semi-detached ageing
	6B4	Older workers and retirement
	7	***Constrained city dwellers***
	7A	**Challenged diversity**
	7A1	Transitional Eastern European neighbourhoods
	7A2	Hampered aspiration
	7A3	Multi-ethnic hardship
	7B	**Constrained flat dwellers**
	7B1	Eastern European communities
	7B2	Deprived neighbourhoods
	7B3	Endeavouring flat dwellers
	7C	**White communities**
	7C1	Challenged transitionaries
	7C2	Constrained young families
	7C3	Outer city hardship
	7D	**Ageing city dwellers**
	7D1	Ageing communities and families
	7D2	Retired independent city dwellers

	7D3	Retired communal city dwellers
	7D4	Retired city hardship
	8	***Hard-pressed living***
	8A	**Industrious communities**
	8A1	Industrious transitions
	8A2	Industrious hardship
	8B	**Challenged terraced workers**
	8B1	Deprived blue-collar terraces
	8B2	Hard pressed rented terraces
	8C	**Hard pressed ageing workers**
	8C1	Ageing industrious workers
	8C2	Ageing rural industry workers
	8C3	Renting hard-pressed workers
	8D	**Migration and churn**
	8D1	Young hard-pressed families
	8D2	Hard-pressed ethnic mix
	8D3	Hard-Pressed European Settlers
Super Output Area Code	Code	Descriptor
	1	**Cosmopolitan student neighbourhoods**
	1A	Cosmopolitan student neighbourhoods
	2	**Countryside living**
	2A	Ageing rural neighbourhoods
	2B	Prospering countryside life
	2C	Remoter communities
	2D	Rural traits
	3	**Ethnically diverse professionals**
	3A	Achieving neighbourhoods
	3B	Asian traits
	3C	Highly qualified professionals
	3D	Households in terraces and flats
	4	**Hard-pressed communities**
	4A	Challenged white communities
	4B	Constrained renters
	4C	Hampered neighbourhoods
	4D	Hard-pressed flat dwellers
	5	**Industrious communities**
	5A	Ageing urban communities
	5B	Aspiring urban households
	5C	Comfortable neighbourhoods
	5D	Endeavouring social renters
	5E	Primary sector workers
	6	**Inner city cosmopolitan**
	6A	Inner city cosmopolitan
	7	**Multicultural living**
	7A	Urban cultural mix
	7B	Young ethnic communities
	8	**Suburban living**
	8A	Affluent communities
	8B	Ageing suburbanites

	8C	Comfortable suburbia
Workplace Area Code	Code	Descriptor
	1	**Retail**
	1A	Low density retail and wholesale
	1B	Market squares
	1C	Multicultural urban high streets
	1D	Traditional high streets
	1E	Shop until you drop
	1F	Eat, drink, shop and be merry
	2	**Top jobs**
	2A	Global business
	2B	Administrative centres
	2C	Big city life
	2D	Regional business centres
	2E	Science and business parks
	3	**Metro suburbs**
	3A	Metro suburban distribution
	3B	Cosmopolitan metro suburban mix
	3C	Independent professional metro services
	3D	Suburban metro infrastructure
	4	**Suburban services**
	4A	Non-metropolitan suburban areas
	4B	Primarily residential suburbs
	5	**Manufacturing and distribution**
	5A	Mining and quarrying facilities
	5B	Industrial units
	5C	Business parks
	5D	Manufacturing, energy and utilities
	6	**Rural**
	6A	Rural with core services
	6B	Rural with non-local workers
	6C	Rural with mining or quarrying
	6D	Traditional countryside
	7	**Servants of society**
	7A	Large scale education
	7B	Public administration
	7C	Major hospitals
	7D	Highly qualified workforces and professional services
Local Authority Code	Code	Descriptor
	1	***Affluent England***
	1A	**Rural-urban fringe**
	1A1	Rural-urban fringe
	1B	**Thriving rural**
	1B1	Affluent rural
	1B2	Rural growth areas
	2	***Business, education and heritage centres***
	2A	Larger towns and cities
	2A1	Larger towns and cities
	2B	**University towns and cities**

	2B1	University towns and cities
	3	***Countryside living***
	3A	**English and Welsh countryside**
	3A1	English and Welsh countryside
	3B	**Remoter coastal living**
	3B1	Ageing coastal living
	3B2	Seaside living
	3C	**Scottish countryside**
	3C1	Scottish countryside
	4	***Ethnically diverse metropolitan living***
	4A	**Ethnically diverse metropolitan living**
	4A1	Ethnically diverse metropolitan living
	5	***London cosmopolitan***
	5A	**London cosmopolitan**
	5A1	London cosmopolitan
	6	***Services and industrial legacy***
	6A	**Services, manufacturing and mining legacy**
	6A1	Manufacturing legacy
	6A2	Mining legacy
	6A3	Service economy
	6B	**Scottish industrial legacy**
	6B1	Scottish industrial legacy
	7	***Town and country living***
	7A	***Country living***
	7A1	Country living
	7B	**Northern Ireland countryside**
	7B1	Northern Ireland countryside
	7C	**Town living**
	7C1	Town living
	8	***Urban settlements***
	8A	**Manufacturing traits**
	8A1	Industrial and multi-ethnic
	8A2	Urban living
	8B	**Suburban traits**
	8B1	City periphery
	8B2	Expanded areas

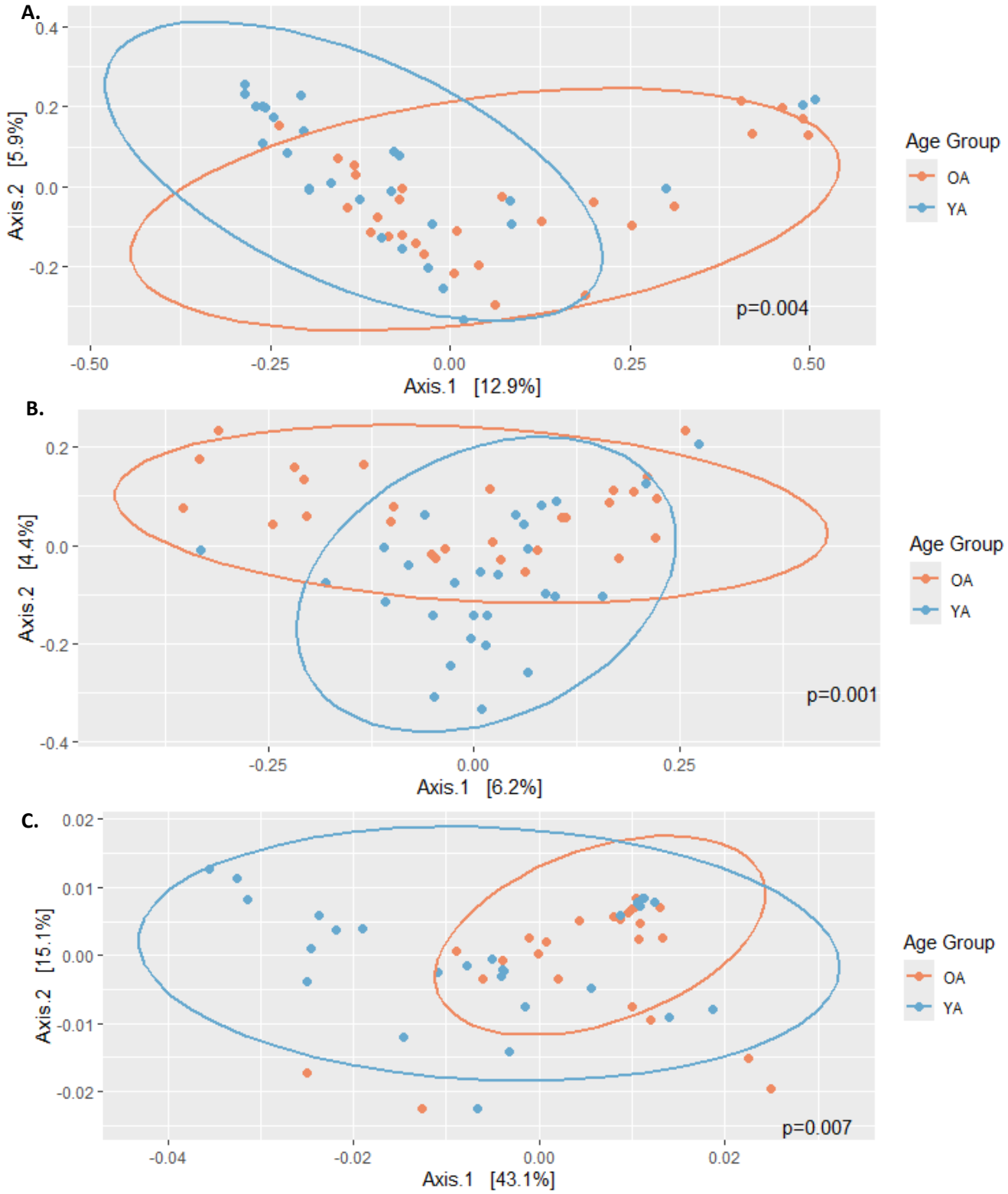


Figure 9.1 Beta diversity as an ordination plot of arm samples grouped by age group. *A. Jaccard Index. B. Unweighted UniFrac. C. Weighted UniFrac.*

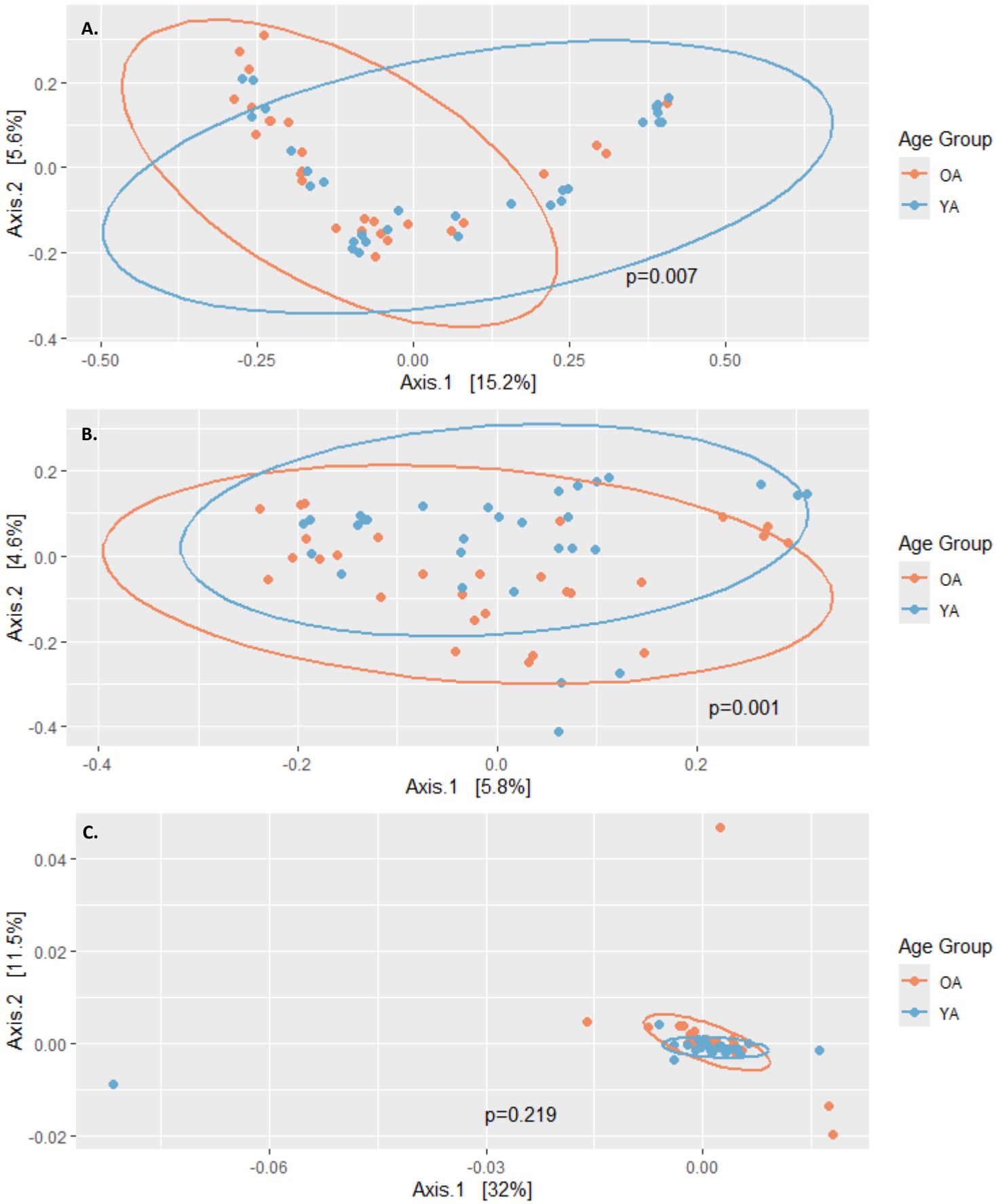


Figure 9.2 Beta diversity as an ordination plot of face samples grouped by age group. *A. Jaccard Index. B. Unweighted UniFrac. C. Weighted UniFrac.*

D. J. Newman, PhD Thesis, Aston University 2025.

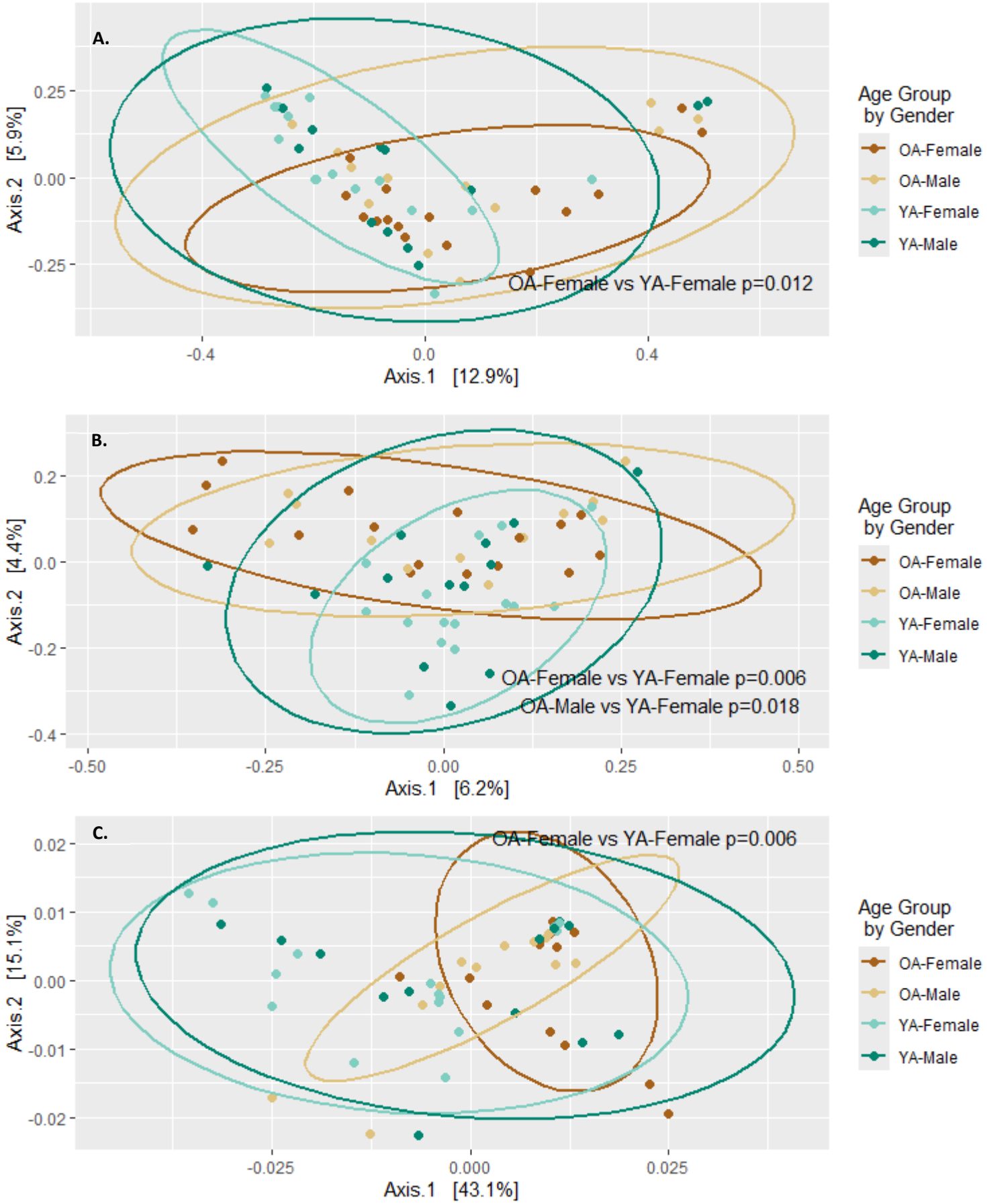


Figure 9.3 Beta diversity as an ordination plot of arm samples grouped by age group and gender. *A. Jaccard Index. B. UniFrac. C. Weighted UniFrac.*

D. J. Newman, PhD Thesis, Aston University 2025.

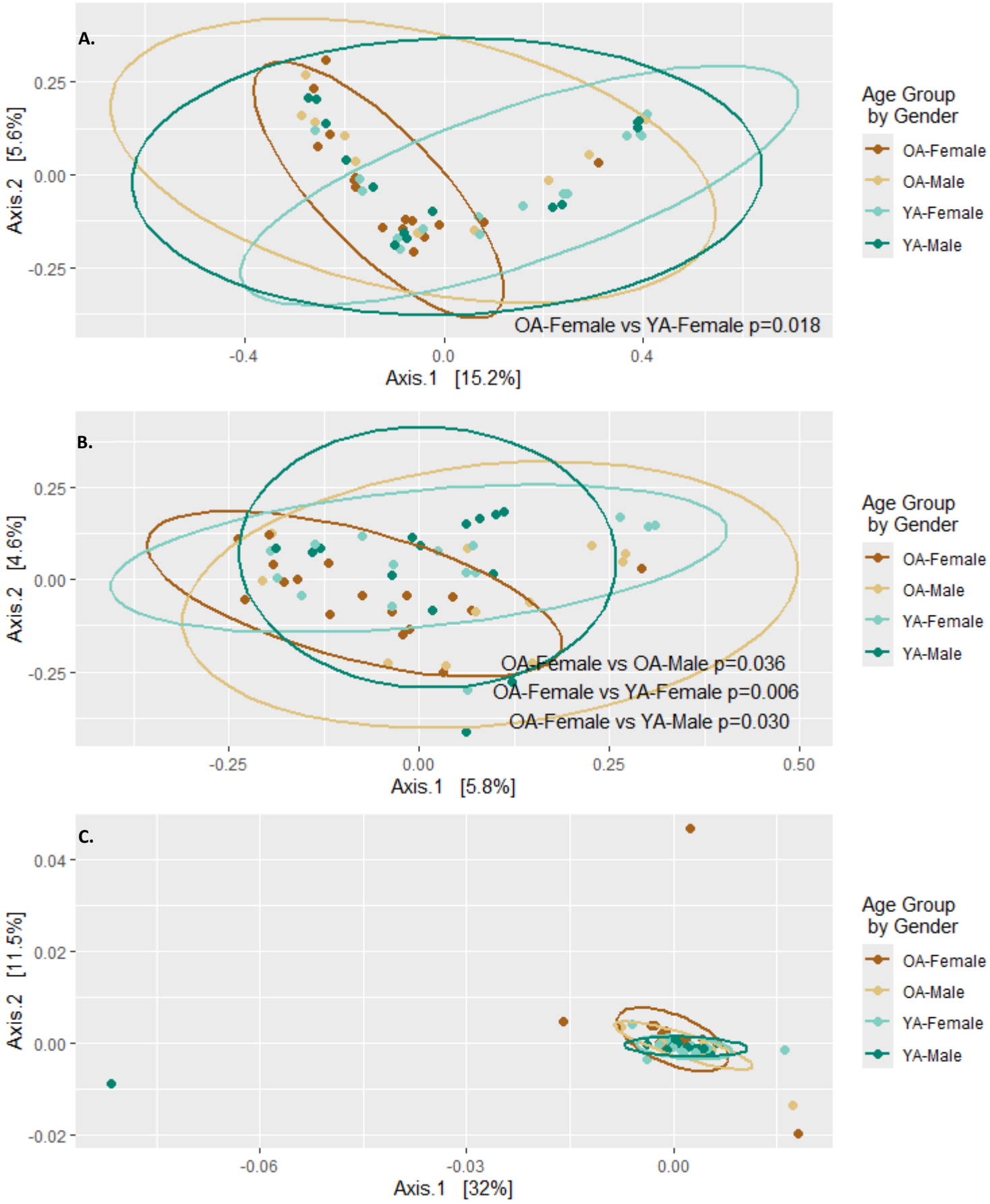


Figure 9.4 Beta diversity as an ordination plot of face samples grouped by age group and gender. *A. Jaccard Index. B. UniFrac. C. Weighted UniFrac.*

D. J. Newman, PhD Thesis, Aston University 2025.

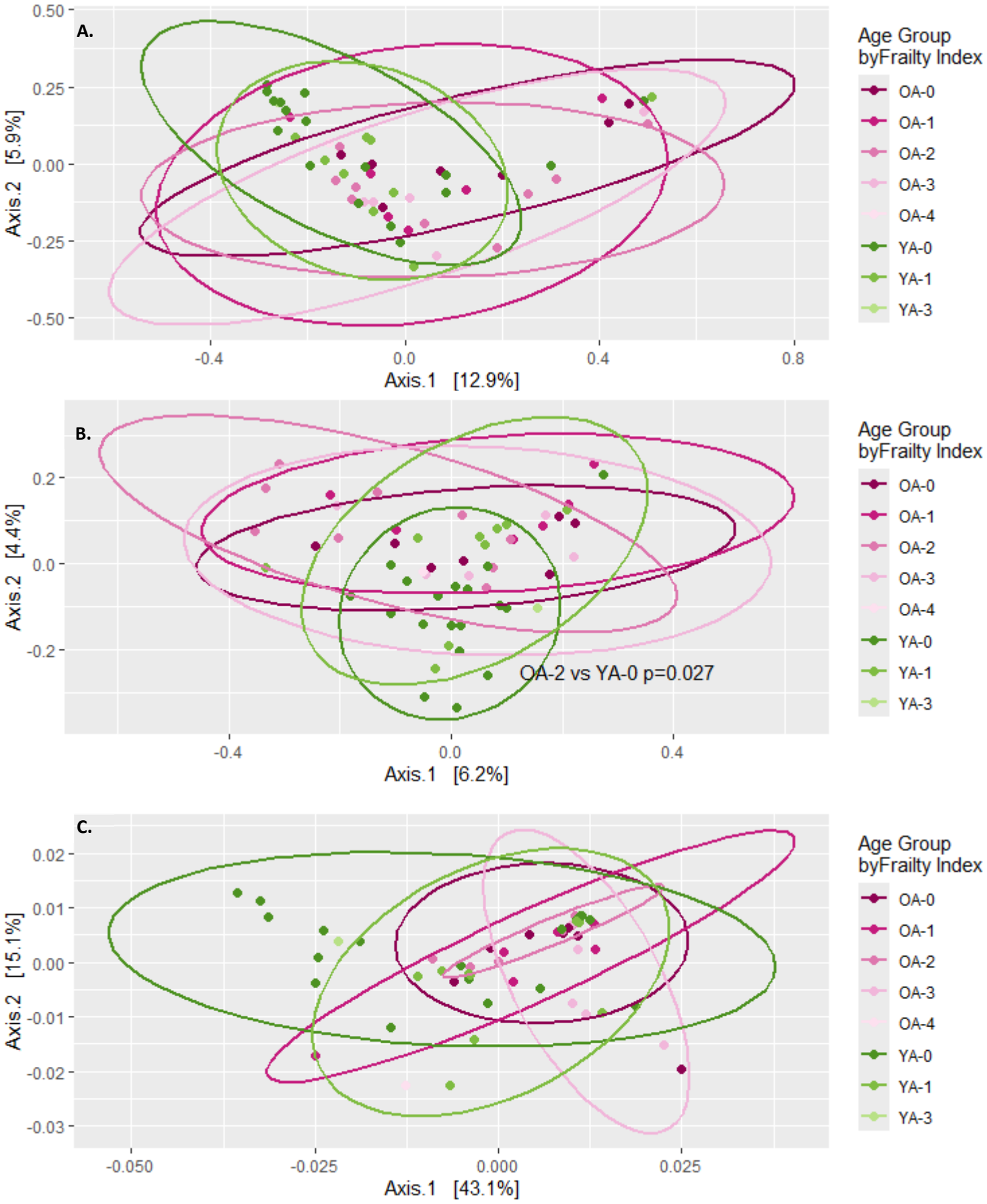


Figure 9.5 Beta diversity as an ordination plot of arm samples grouped by age group and FI score. *A. Jaccard Index. B. UniFrac. C. Weighted UniFrac.*

D. J. Newman, PhD Thesis, Aston University 2025.

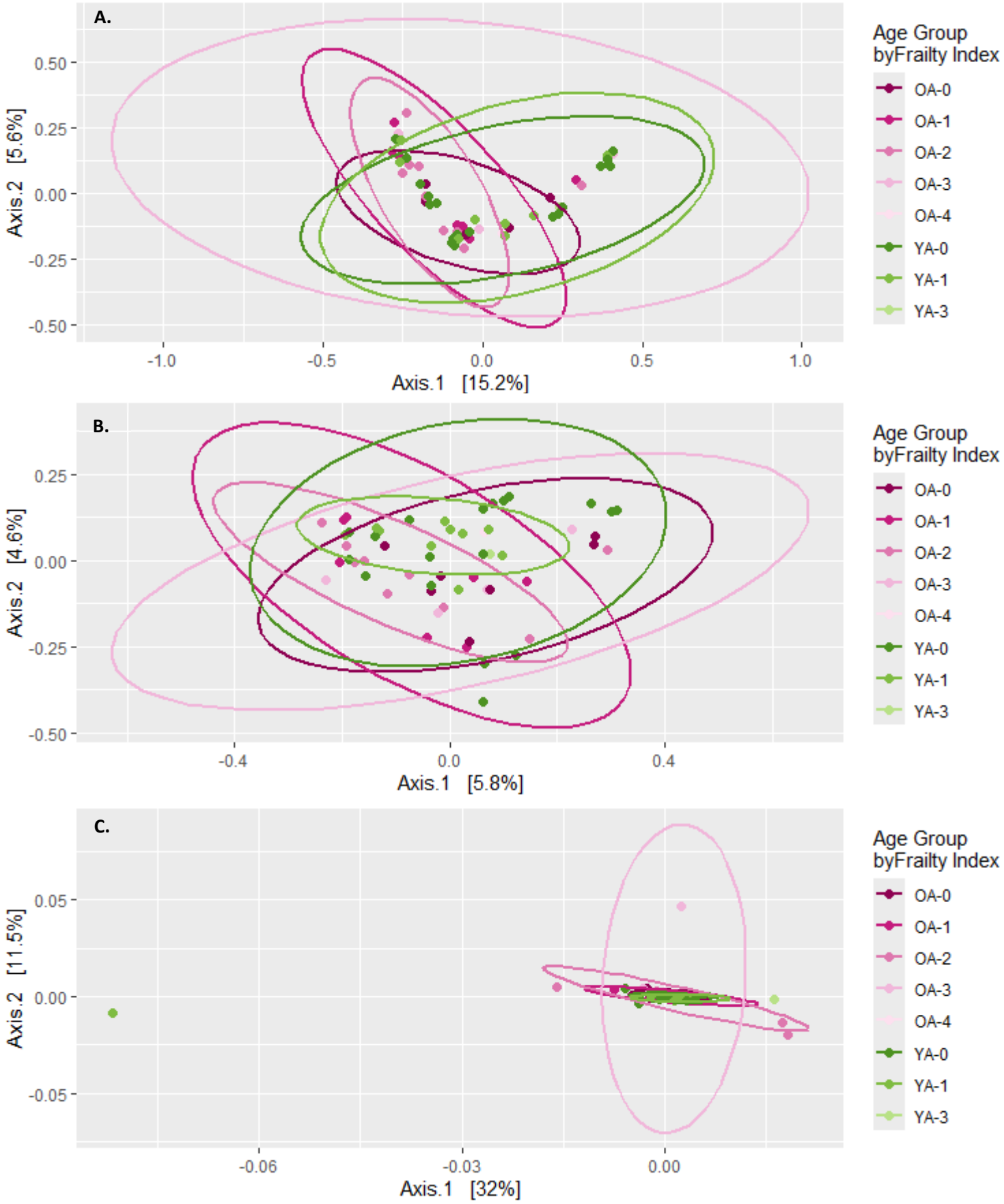


Figure 9.6 Beta diversity as an ordination plot of face samples grouped by age group and FI score. *A. Jaccard Index. B. UniFrac. C. Weighted UniFrac.*

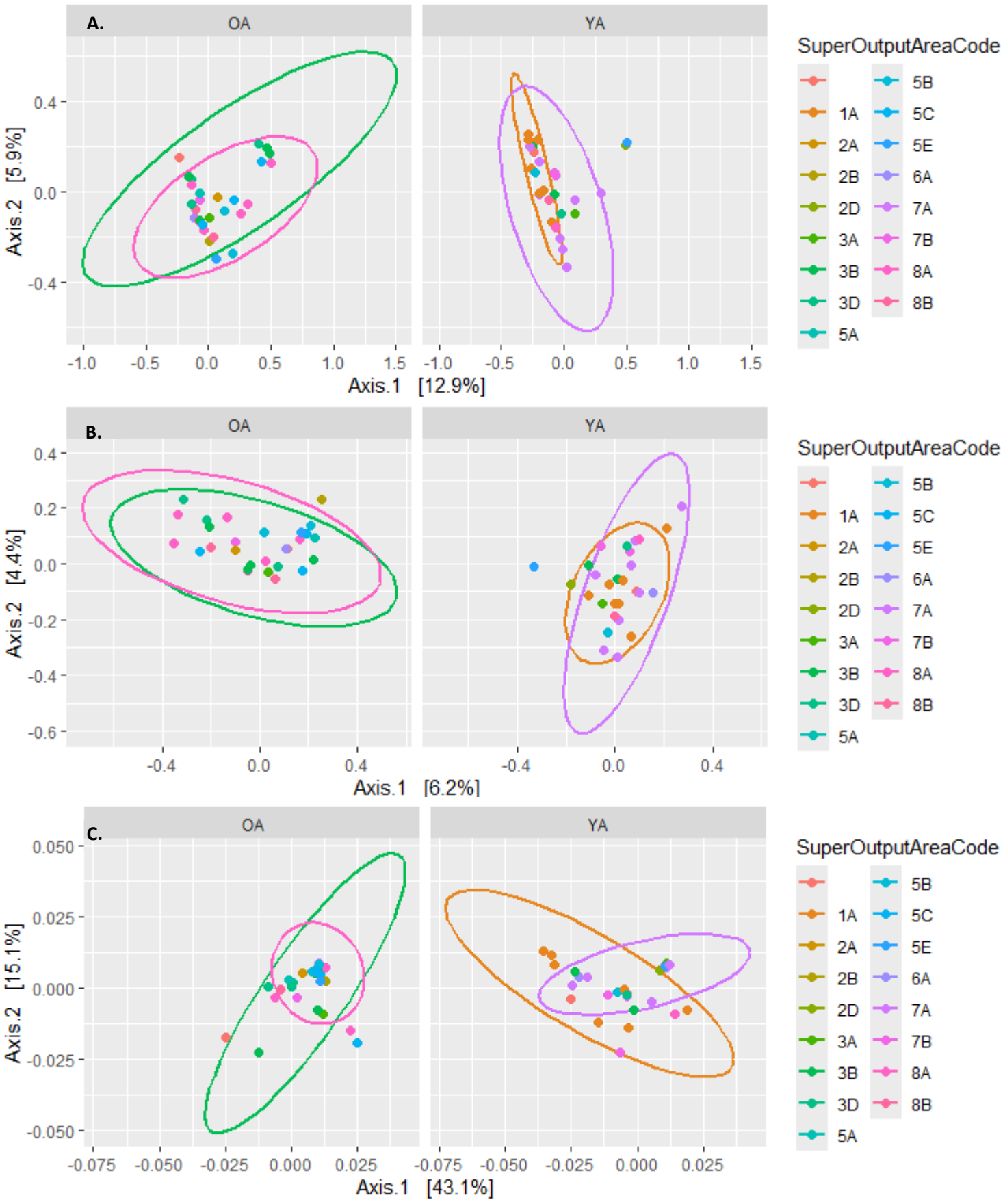


Figure 9.7 Beta diversity as an ordination plot of arm samples grouped by age group and Super Output Area Code. *A. Jaccard Index. B. UniFrac. C. Weighted UniFrac.*

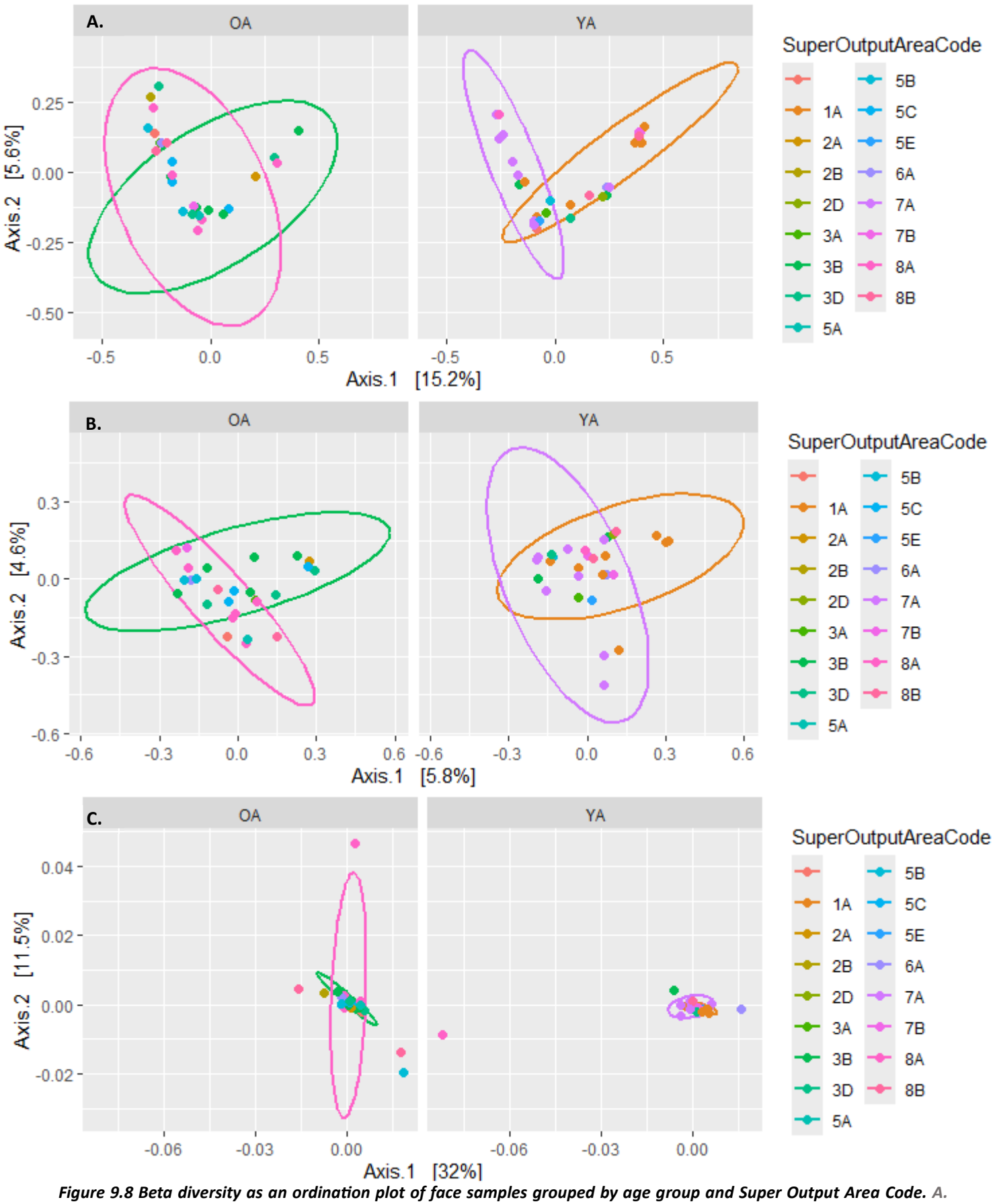


Figure 9.8 Beta diversity as an ordination plot of face samples grouped by age group and Super Output Area Code. *A. Jaccard Index. B. UniFrac. C. Weighted UniFrac.*

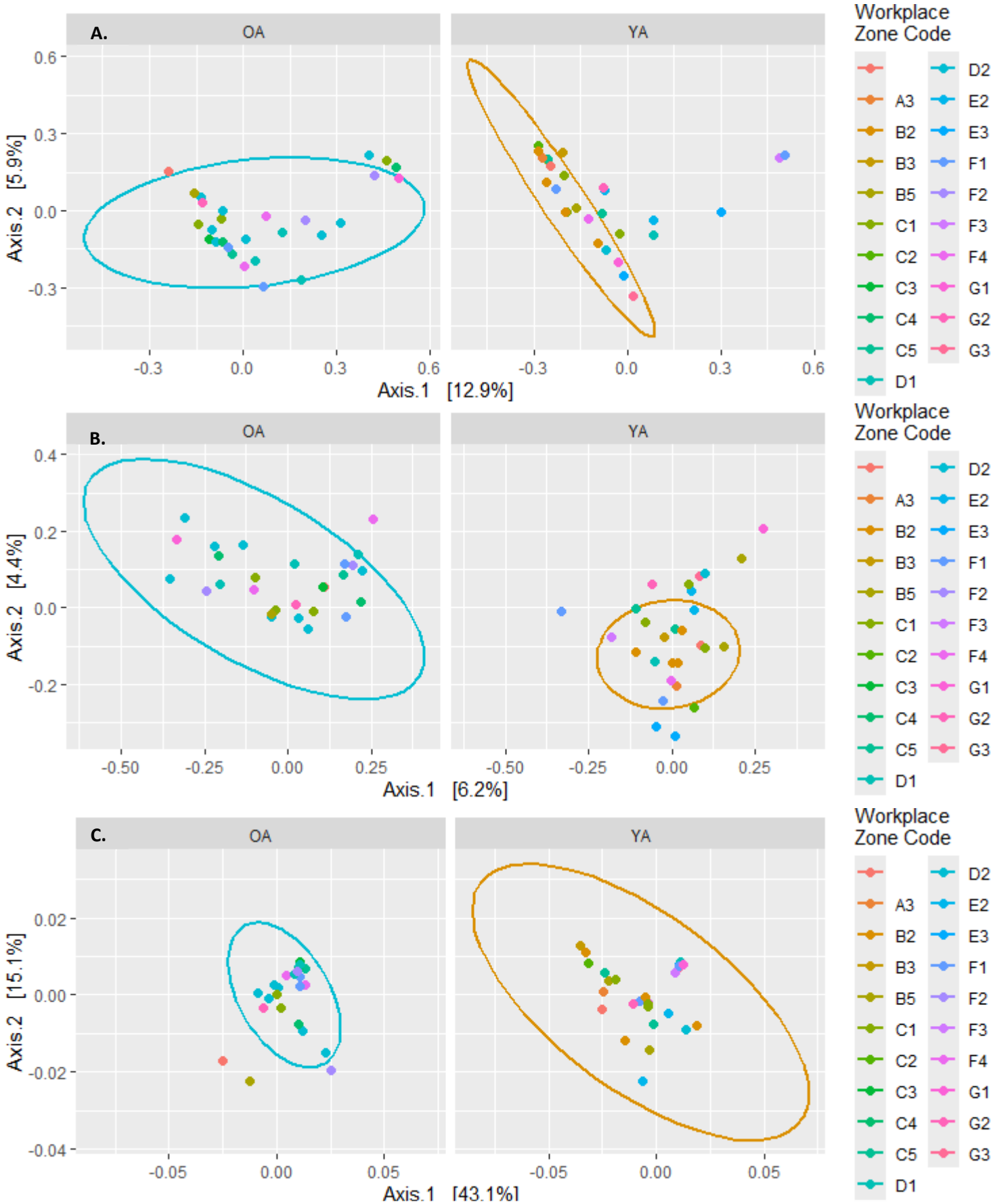


Figure 9.9 Beta diversity as an ordination plot of arm samples grouped by age group and Workplace Zone Code. *A. Jaccard Index. B. UniFrac. C. Weighted UniFrac.*

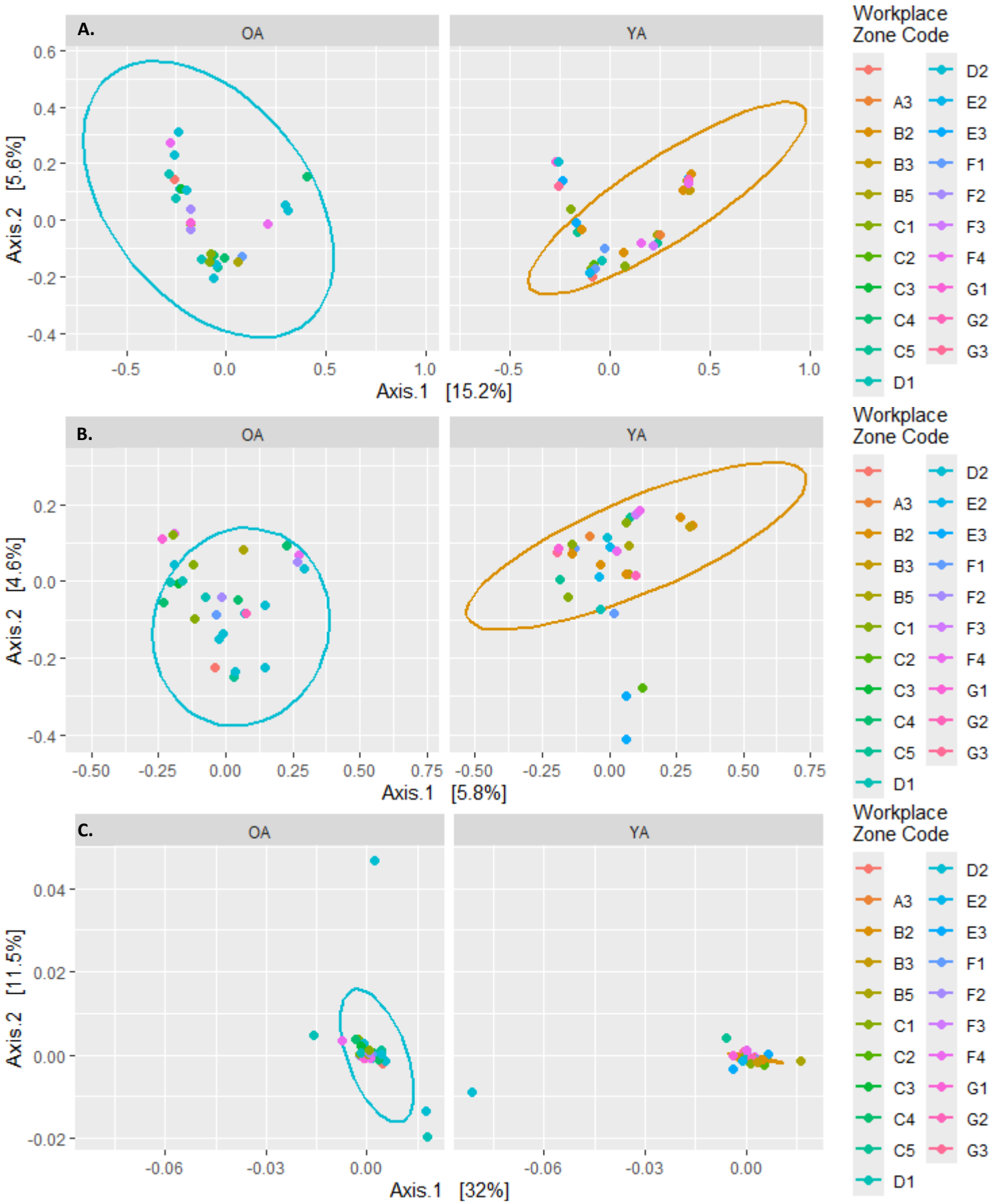


Figure 9.10 Beta diversity as an ordination plot of face samples grouped by age group and Workplace Zone Code. *A. Jaccard Index. B. UniFrac. C. Weighted UniFrac.*

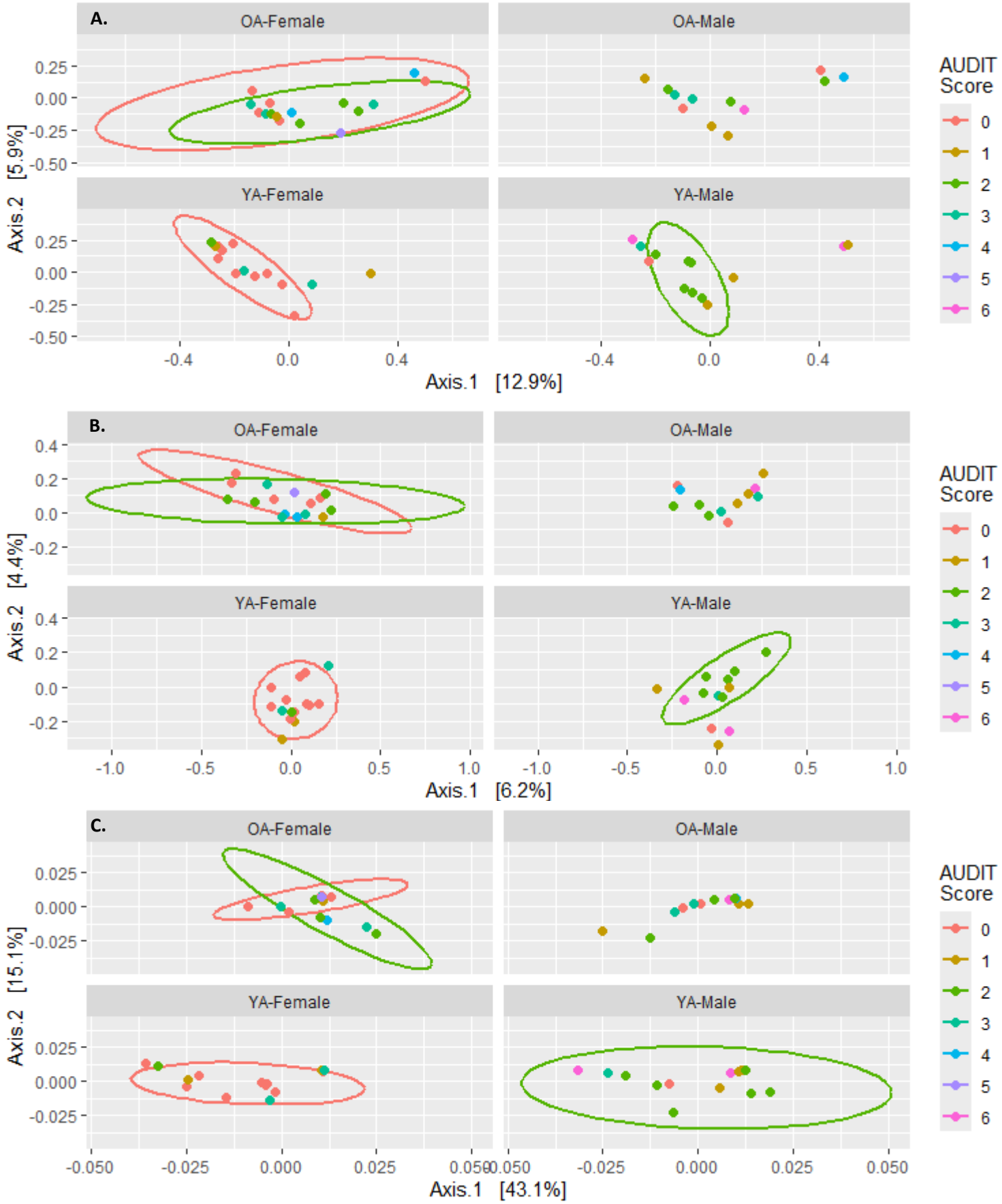


Figure 9.11 Beta diversity as an ordination plot of arm samples grouped by age group and AUDIT-C score. *A. Jaccard Index. B. UniFrac. C. Weighted UniFrac.*

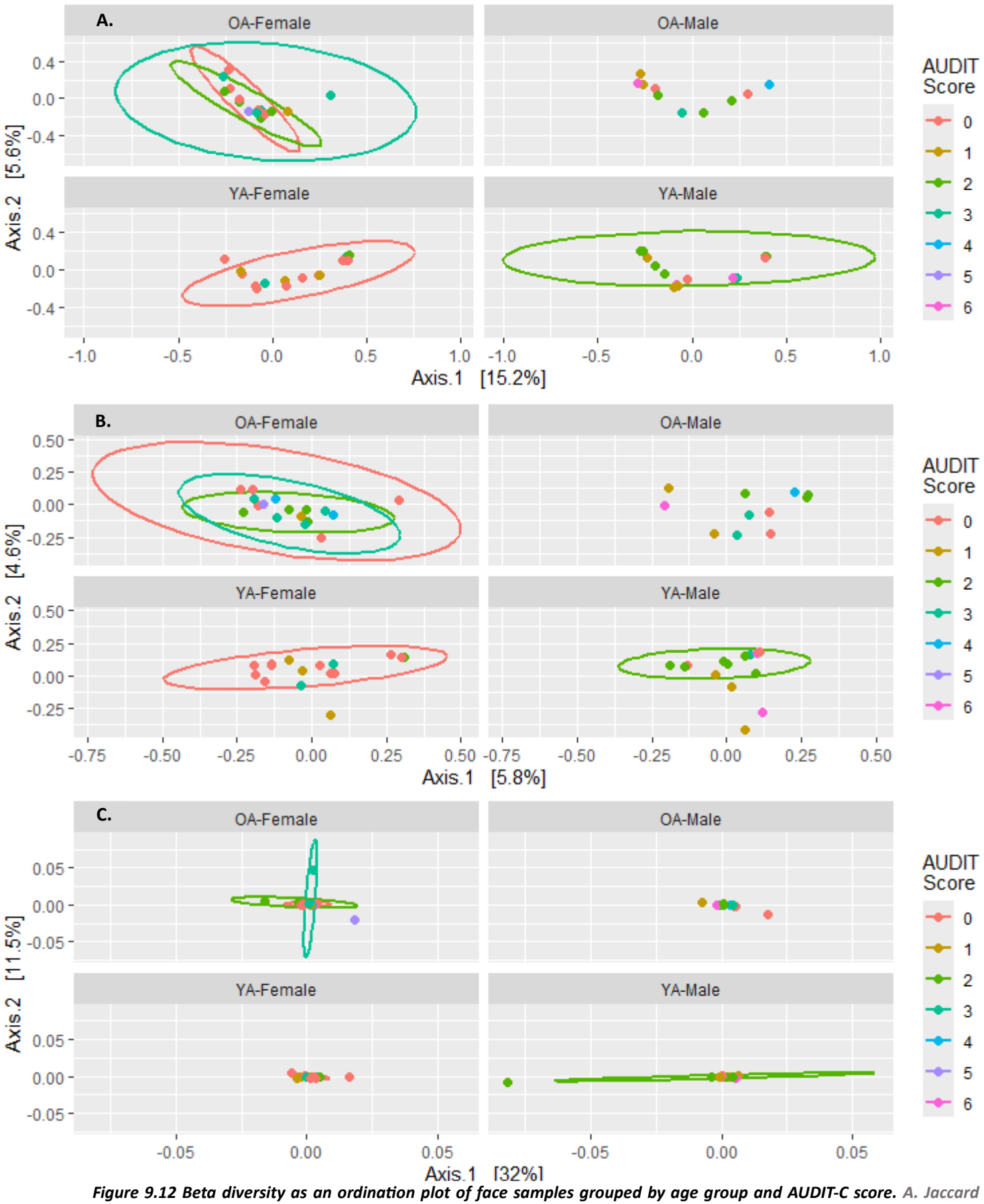


Figure 9.12 Beta diversity as an ordination plot of face samples grouped by age group and AUDIT-C score. *A. Jaccard Index. B. UniFrac. C. Weighted UniFrac.*

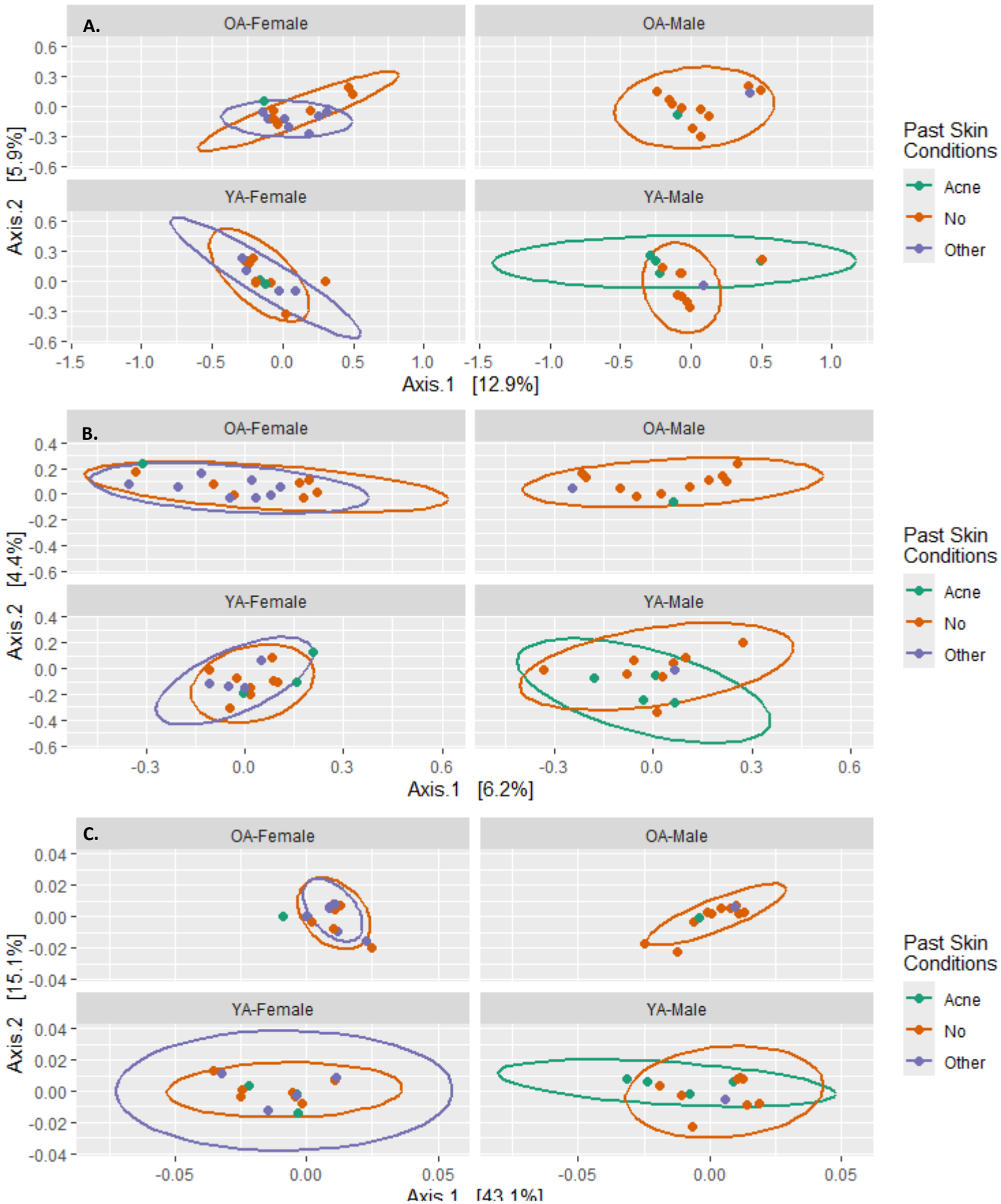


Figure 9.13 Beta diversity as an ordination plot of arm samples grouped by age group and Past skin conditions. *A. Jaccard Index. B. UniFrac. C. Weighted UniFrac.*

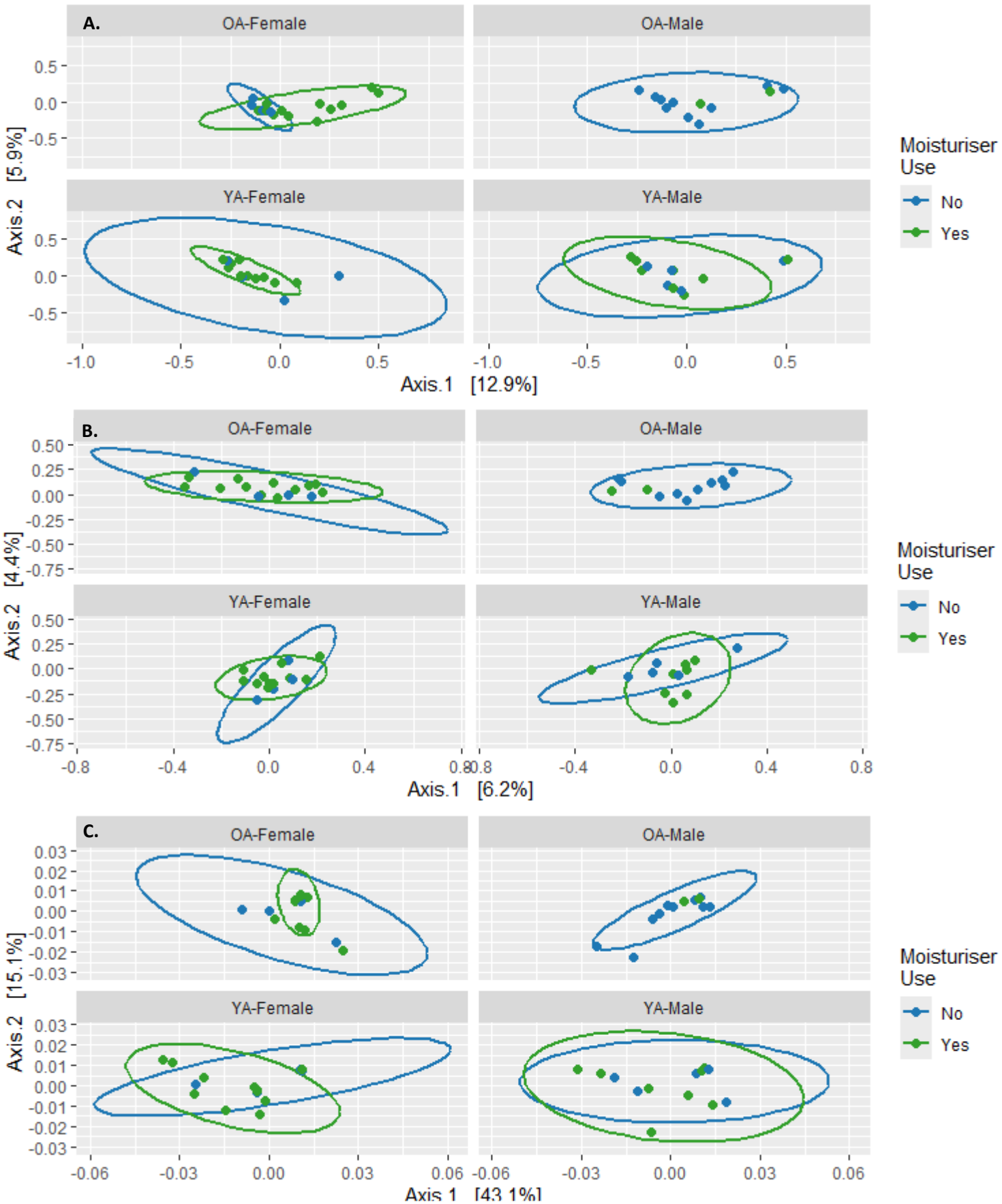


Figure 9.14 Beta diversity as an ordination plot of arm samples grouped by age group and Moisturiser use. *A. Jaccard Index. B. UniFrac. C. Weighted UniFrac.*

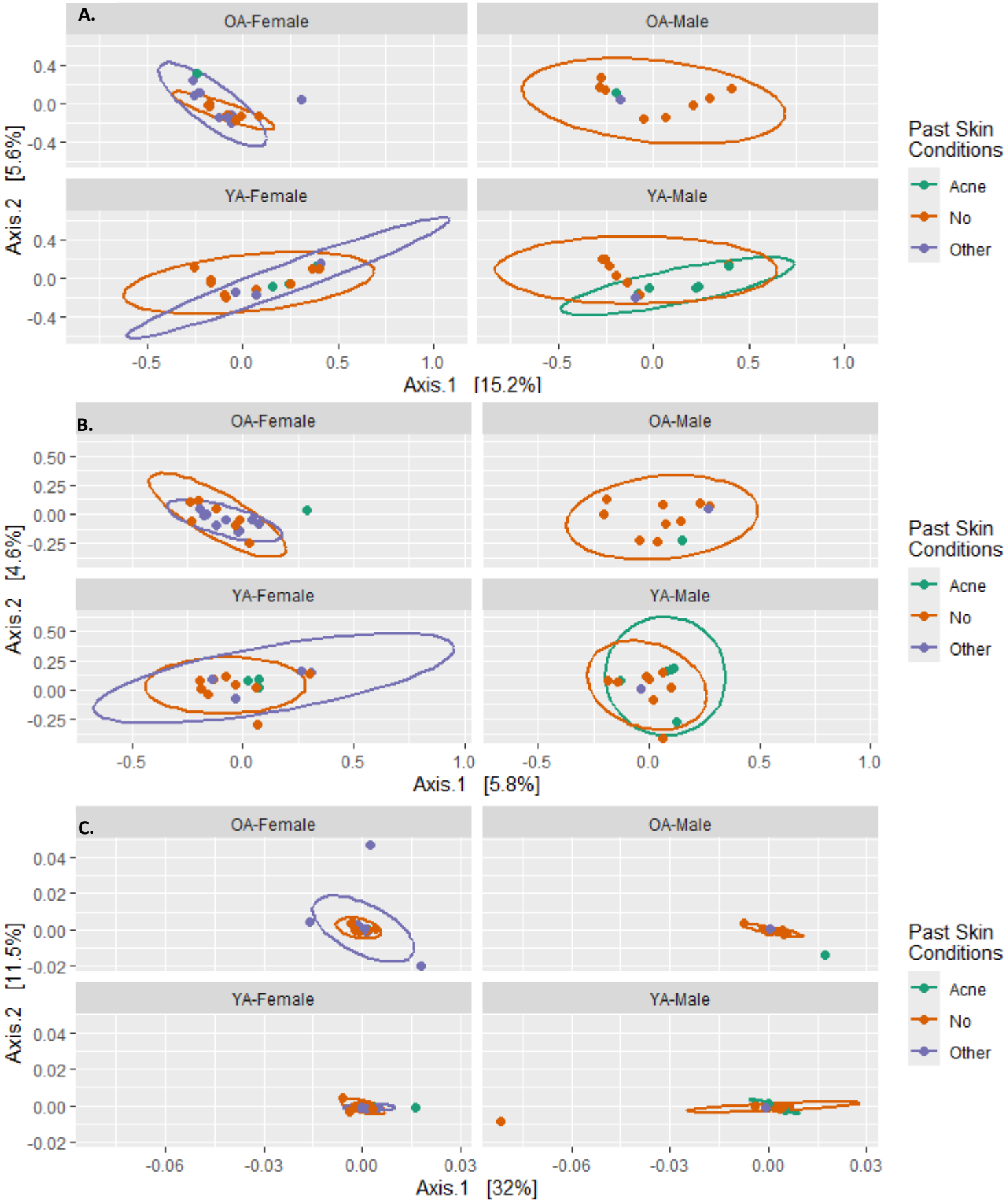


Figure 9.15 Beta diversity as an ordination plot of face samples grouped by age group and Past skin conditions. *A. Jaccard Index. B. UniFrac. C. Weighted UniFrac.*

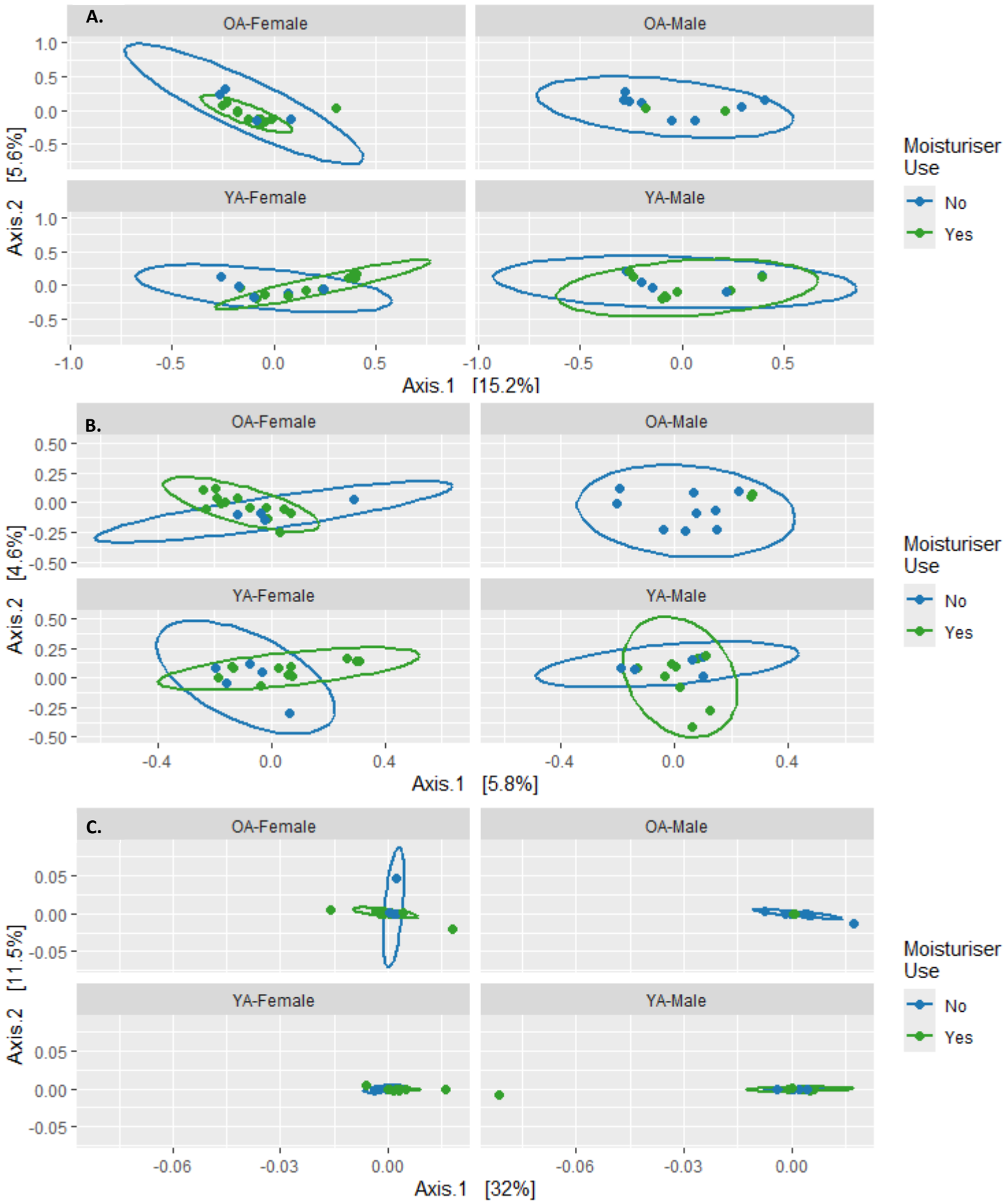


Figure 9.16 Beta diversity as an ordination plot of face samples grouped by age group and Moisturiser use. *A. Jaccard Index. B. UniFrac. C. Weighted UniFrac.*

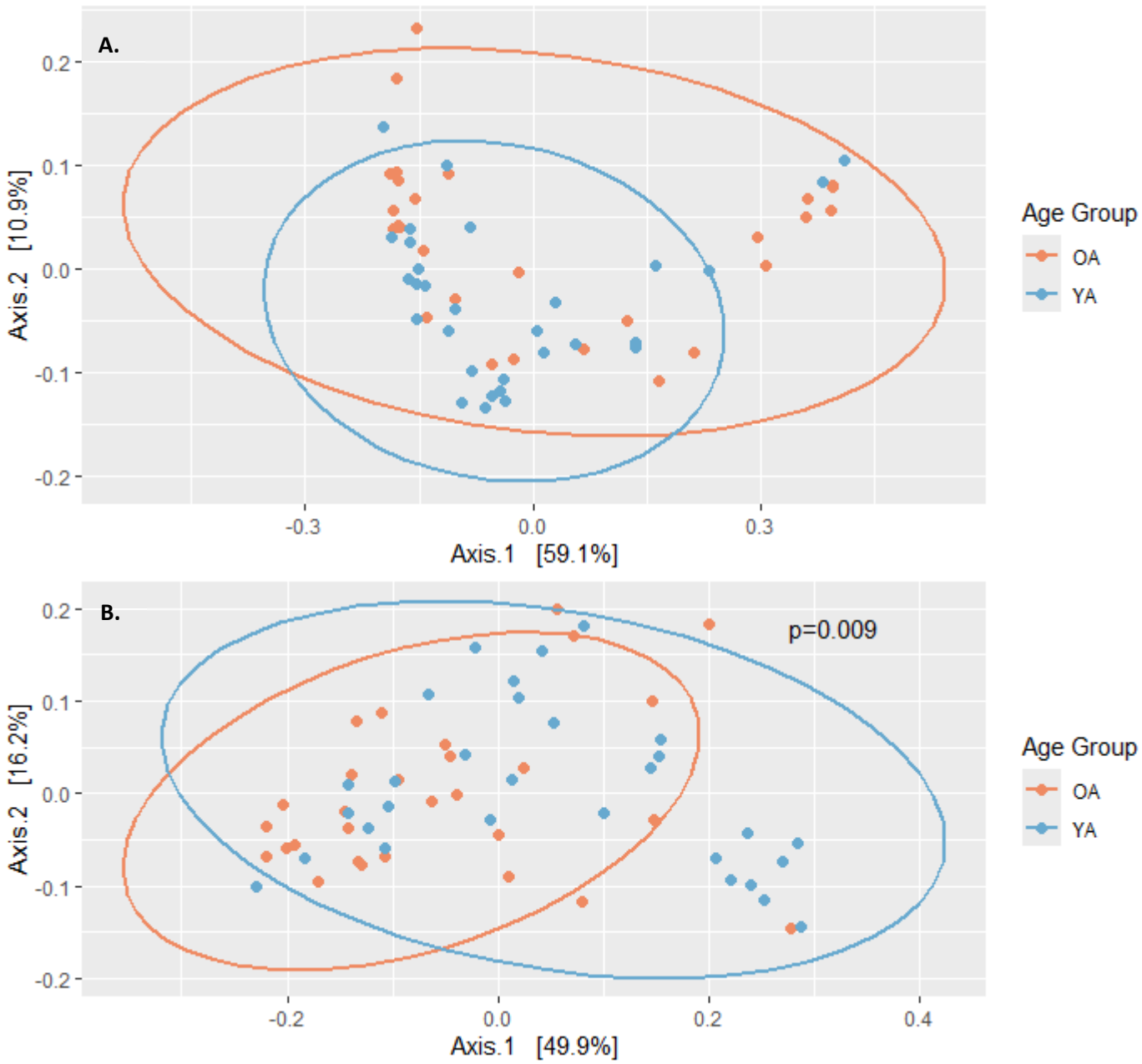


Figure 9.17 Jaccard Index beta diversity as an ordination plot for metacyc pathway abundance data grouped by age group. *A. Arm samples. B. Face samples.*

Table 9.4 Metacyc pathways identified as abundant in more than 3 of the 5 abundance tests carried out on the arm predictive metabolomics dataset, when grouped based on age group. ***Any pathway identified through all five methods and have an agreed abundant group are highlighted in green.***

Descriptor	Abundant Group	Methods Identified by					Count
		Ancombc	Deseq	Edger	Lefse	Limma-Voom	
L-arginine biosynthesis I (via L-ornithine)	OA	✓	✓				3
	YA				✓		
superpathway of branched chain amino acid biosynthesis	OA	✓	✓			✓	3
fatty acid elongation -- saturated	OA	✓	✓			✓	3
D-galactarate degradation I	YA	✓	✓	✓		✓	4
D-glucarate degradation I	YA	✓	✓	✓		✓	4
superpathway of D-glucarate and D-galactarate degradation	YA	✓	✓	✓		✓	4
superpathway of β-D-glucuronides degradation (to pyruvate)	YA		✓	✓	✓	✓	4
3-phenylpropanoate and 3-(3-hydroxyphenyl)propanoate degradation to 2-hydroxypentadienoate	YA	✓	✓			✓	3
heme *b* biosynthesis I (aerobic)	OA	✓	✓				3
	YA					✓	
ectoine biosynthesis	YA		✓		✓	✓	3
pyruvate fermentation to propanoate I	YA		✓	✓	✓		3
3-phenylpropanoate degradation	YA	✓	✓			✓	
L-isoleucine biosynthesis II	OA	✓	✓		✓	✓	4
superpathway of menaquinol-8 biosynthesis I	YA		✓	✓	✓	✓	4
superpathway of demethylmenaquinol-8 biosynthesis I	YA		✓	✓	✓	✓	4
superpathway of heme *b* biosynthesis from glutamate	OA	✓	✓				4
	YA				✓	✓	
GDP-D-*glycero*-α-D-*manno*-heptose biosynthesis	YA	✓	✓	✓		✓	4
spirilloxanthin and 2,2'-diketo-spirilloxanthin biosynthesis	OA	✓	✓	✓			3
adenine and adenosine salvage III	OA	✓	✓				3
	YA					✓	
cinnamate and 3-hydroxycinnamate degradation to 2-hydroxypentadienoate	YA	✓	✓			✓	3
queuosine biosynthesis I (*de novo*)	OA	✓	✓			✓	3
thiamine diphosphate salvage II	OA	✓	✓				3
	YA					✓	
pyruvate fermentation to isobutanol (engineered)	OA	✓	✓		✓	✓	4
pyrimidine deoxyribonucleotides *de novo* biosynthesis II	OA	✓	✓				3
	YA				✓		
superpathway of pyrimidine ribonucleosides salvage	OA	✓	✓				3
	YA					✓	
pyrimidine deoxyribonucleosides salvage	OA	✓	✓				3
	YA					✓	

superpathway of pyrimidine deoxyribonucleoside salvage	OA	✓	✓				3
	YA					✓	
D-fructuronate degradation	YA		✓	✓	✓	✓	4
L-arginine biosynthesis IV (archaea)	OA	✓	✓				3
	YA				✓		
gondoate biosynthesis (anaerobic)	OA	✓	✓			✓	3
assimilatory sulfate reduction I	OA	✓	✓		✓	✓	4
superpathway of sulfate assimilation and cysteine biosynthesis	OA	✓	✓		✓	✓	4
tRNA charging	OA	✓	✓				3
	YA					✓	
L-tryptophan biosynthesis	OA	✓	✓		✓	✓	4

Table 9.5 Metacyc pathways identified as abundant in more than 3 of the 5 abundance tests carried out on the face predictive metabolomics dataset, when grouped based on age group. ***Any pathway identified through all five methods and have an agreed abundant group are highlighted in green.***

Descriptor	Abundant Group	Methods Identified by					Count
		Ancombc	Deseq	Edger	Lefse	Limma-Voom	
4-hydroxyphenylacetate degradation	OA	✓		✓	✓	✓	4
homolactic fermentation	OA					✓	4
	YA	✓	✓	✓			
glycolysis III (from glucose)	YA	✓	✓	✓	✓		4
L-arginine biosynthesis II (acetyl cycle)	OA					✓	5
	YA	✓	✓	✓	✓		
L-arginine biosynthesis I (via L-ornithine)	OA					✓	5
	YA	✓	✓	✓	✓		
chorismate biosynthesis I	OA					✓	4
	YA	✓	✓	✓			
L-arginine degradation II (AST pathway)	OA	✓	✓	✓	✓	✓	5
Calvin-Benson-Bassham cycle	YA	✓	✓	✓			3
catechol degradation to β-ketoadipate	OA		✓	✓		✓	3
3,8-divinyl-chlorophyllide *a* biosynthesis (from uroporphyrinogen-III) I (aerobic, light-dependent)	OA	✓	✓	✓		✓	4
coenzyme A biosynthesis I (bacteria)	OA					✓	5
	YA	✓	✓	✓	✓		
superpathway of adenosylcobalamin salvage from cobinamide I	OA		✓	✓	✓		3
colanic acid building blocks biosynthesis	OA		✓	✓	✓	✓	4
superpathway of aromatic amino acid biosynthesis (from phospho*enol*pyruvate and D-erythrose 4-phosphate)	OA					✓	4
	YA	✓	✓	✓			
L-lysine biosynthesis I	YA	✓	✓	✓	✓		4
nitrate reduction I (denitrification)	OA	✓	✓	✓	✓	✓	5
superpathway of purine nucleotides *de novo* biosynthesis II	OA					✓	5
	YA	✓	✓	✓	✓		

glucose degradation (oxidative)	OA		✓	✓	✓	✓	4
dTDP-L-rhamnose biosynthesis	OA		✓	✓	✓		3
enterobactin biosynthesis	OA		✓	✓	✓		3
fatty acid β-oxidation I (generic)	OA		✓	✓	✓		3
superpathway of tetrahydrofolate biosynthesis and salvage	OA					✓	5
	YA	✓	✓	✓	✓		
superpathway of *N*-acetylneuraminate, *N*-acetylglucosamine, and *N*-acetylmannosamine degradation to β-D-fructofuranose 6-phosphate	YA	✓	✓	✓			3
gluconeogenesis I	OA					✓	4
	YA	✓	✓	✓			
glucose and glucose-1-phosphate degradation	OA					✓	4
	YA	✓	✓	✓			
superpathway of β-D-glucuronides degradation (to pyruvate)	YA	✓	✓	✓			3
L-ornithine biosynthesis I	YA	✓	✓	✓			3
glycogen degradation I	OA					✓	4
	YA	✓	✓	✓			
glycogen biosynthesis I (from ADP-D-Glucose)	OA					✓	4
	YA	✓	✓	✓			
heme *b* biosynthesis I (aerobic)	OA					✓	5
	YA	✓	✓	✓	✓		
heme *b* biosynthesis II (oxygen-independent)	YA	✓	✓	✓	✓		4
L-histidine degradation I	OA					✓	5
	YA	✓	✓	✓	✓		
L-histidine biosynthesis	OA					✓	5
	YA	✓	✓	✓	✓		
lipid IV_A biosynthesis (*E. coli*)	OA		✓	✓	✓		3
methylerythritol phosphate pathway I	YA	✓	✓	✓	✓		4
pentose phosphate pathway (non-oxidative branch) I	OA					✓	4
	YA	✓	✓	✓			
O-antigen building blocks biosynthesis (*E. coli*)	OA		✓	✓		✓	3
superpathway of ornithine degradation	OA		✓	✓	✓		3
pyruvate fermentation to propanoate I	OA					✓	5
	YA	✓	✓	✓	✓		
incomplete reductive TCA cycle	OA					✓	5
	YA	✓	✓	✓	✓		
myo-inositol degradation I	OA	✓		✓	✓	✓	4
pentose phosphate pathway	YA	✓	✓	✓			3
peptidoglycan biosynthesis I (*meso*-diaminopimelate containing)	OA					✓	5
	YA	✓	✓	✓	✓		
superpathway of phospholipid biosynthesis III (*E. coli*)	OA					✓	4
	YA	✓	✓	✓			
polyisoprenoid biosynthesis (*E. coli*)	OA					✓	5
	YA	✓	✓	✓	✓		
superpathway of histidine, purine, and pyrimidine biosynthesis	YA	✓	✓	✓			3
ADP-L-*glycero*-β-D-*manno*-heptose biosynthesis	OA		✓	✓	✓	✓	4
superpathway of purine deoxyribonucleosides degradation	YA	✓	✓	✓			3
CDP-diacylglycerol biosynthesis II	OA					✓	4
	YA	✓	✓	✓			

methylphosphonate degradation I	OA		✓	✓	✓		3
peptidoglycan maturation (*meso*-diaminopimelate containing)	OA		✓	✓	✓		3
superpathway of pyrimidine ribonucleotides *de novo* biosynthesis	OA					✓	5
	YA	✓	✓	✓	✓		
superpathway of pyrimidine deoxyribonucleotides *de novo* biosynthesis (*E. coli*)	OA					✓	5
	YA	✓	✓	✓	✓		
phenylacetate degradation I (aerobic)	OA	✓	✓	✓	✓	✓	5
superpathway of pyridoxal 5'-phosphate biosynthesis and salvage	OA		✓	✓	✓		3
(5Z)-dodecenoate biosynthesis I	OA		✓	✓	✓		3
CMP-3-deoxy-D-*manno*-octulosonate biosynthesis	OA		✓	✓	✓		3
vitamin E biosynthesis (tocopherols)	OA	✓	✓	✓		✓	4
formaldehyde assimilation II (assimilatory RuMP Cycle)	YA	✓	✓	✓			3
L-lysine biosynthesis II	YA	✓	✓	✓			3
L-lysine biosynthesis III	YA	✓	✓	✓	✓		4
glycine betaine degradation I	OA	✓		✓	✓	✓	4
aerobic respiration I (cytochrome c)	YA	✓	✓	✓			3
L-lysine biosynthesis VI	YA	✓	✓	✓	✓		4
superpathway of geranylgeranyl diphosphate biosynthesis II (via MEP)	OA					✓	5
	YA	✓	✓	✓	✓		
toluene degradation III (aerobic) (via *p*-cresol)	OA	✓	✓	✓	✓	✓	5
uroporphyrinogen-III I (from glutamate)	OA					✓	5
	YA	✓	✓	✓	✓		
uroporphyrinogen-III II (from glycine)	OA					✓	5
	YA	✓	✓	✓	✓		
factor 420 biosynthesis II (mycobacteria)	OA	✓	✓	✓		✓	4
superpathway of sulphur oxidation (*Acidianus ambivalens*)	YA	✓	✓		✓		3
sucrose degradation IV (sucrose phosphorylase)	OA					✓	4
	YA	✓	✓	✓			
catechol degradation III (ortho-cleavage pathway)	OA			✓	✓	✓	3
aromatic compounds degradation via 3-oxoadipate	OA			✓	✓	✓	3
glycolysis II (from fructose 6-phosphate)	OA					✓	4
	YA	✓	✓	✓			
adenosylcobalamin biosynthesis from adenosylcobinamide-GDP I	OA		✓	✓	✓	✓	4
superpathway of bacteriochlorophyll *a* biosynthesis (from glycine)	OA	✓	✓	✓		✓	4
GDP-mannose biosynthesis	OA					✓	4
	YA	✓	✓	✓			
CDP-diacylglycerol biosynthesis I	OA					✓	4
	YA	✓	✓	✓			
UMP biosynthesis I	OA					✓	5
	YA	✓	✓	✓	✓		
inosine 5'-phosphate degradation	OA					✓	4
	YA	✓	✓	✓			
1,4-dihydroxy-2-naphthoate biosynthesis	YA	✓	✓	✓	✓		4
superpathway of menaquinol-7 biosynthesis	YA	✓	✓	✓	✓		4
ubiquinol-7 biosynthesis (prokaryotic)	OA		✓	✓	✓		3
ubiquinol-9 biosynthesis (prokaryotic)	OA		✓	✓	✓		3

ubiquinol-10 biosynthesis (prokaryotic)	OA		✓	✓	✓		3
superpathway of phylloquinol biosynthesis	YA	✓	✓	✓	✓		4
superpathway of menaquinol-11 biosynthesis	YA	✓	✓	✓	✓		4
superpathway of menaquinol-12 biosynthesis	YA	✓	✓	✓	✓		4
superpathway of menaquinol-13 biosynthesis	YA	✓	✓	✓	✓		4
superpathay of heme biosynthesis from glutamate	OA					✓	5
	YA	✓	✓	✓	✓		
stearate biosynthesis II (bacteria and plants)	OA		✓	✓	✓		3
superpathway of phenylethylamine degradation	OA		✓	✓	✓	✓	4
5-aminoimidazole ribonucleotide biosynthesis I	OA					✓	5
	YA	✓	✓	✓	✓		
5-aminoimidazole ribonucleotide biosynthesis II	OA					✓	5
	YA	✓	✓	✓	✓		
inosine-5'-phosphate biosynthesis I	YA	✓	✓	✓	✓		4
superpathway of guanosine nucleotides *de novo* biosynthesis II	OA					✓	5
	YA	✓	✓	✓	✓		
superpathway of adenosine nucleotides de novo biosynthesis II	YA	✓	✓	✓	✓		4
6-hydroxymethyl-dihydropterin diphosphate biosynthesis I	YA	✓	✓	✓	✓		4
S-adenosyl-L-methionine cycle I	OA					✓	5
	YA	✓	✓	✓	✓		
chorismate biosynthesis from 3-dehydroquinate	OA					✓	5
	YA	✓	✓	✓	✓		
superpathway of salicylate degradation	OA		✓	✓	✓	✓	4
4-methylcatechol degradation (*ortho* cleavage)	OA	✓	✓	✓	✓	✓	5
sucrose degradation III (sucrose invertase)	OA					✓	4
	YA	✓	✓	✓			
adenosylcobalamin salvage from cobinamide II	OA		✓	✓	✓	✓	4
superpathway of 5-aminoimidazole ribonucleotide biosynthesis	OA					✓	5
	YA	✓	✓	✓	✓		
palmitoleate biosynthesis I (from (5Z)-dodec-5-enoate)	OA		✓	✓	✓		3
mono-*trans*, poly-*cis* decaprenyl phosphate biosynthesis	YA	✓	✓	✓	✓		4
peptidoglycan biosynthesis III (mycobacteria)	OA					✓	5
	YA	✓	✓	✓	✓		
UDP-*N*-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)	OA					✓	5
	YA	✓	✓	✓	✓		
UDP-*N*-acetylmuramoyl-pentapeptide biosynthesis I (*meso*-diaminopimelate containing)	OA					✓	5
	YA	✓	✓	✓	✓		
Kdo transfer to lipid IV_A III (Chlamydia)	OA		✓	✓	✓		3
8-amino-7-oxononanoate biosynthesis I	OA		✓	✓	✓		3
norspermidine biosynthesis	OA		✓	✓	✓		3
superpathway of polyamine biosynthesis III	OA		✓	✓	✓		3
adenine and adenosine salvage III	OA					✓	5
	YA	✓	✓	✓	✓		
superpathway of tetrahydrofolate biosynthesis	OA					✓	5
	YA	✓	✓	✓	✓		
ubiquinol-8 biosynthesis (prokaryotic)	OA		✓	✓	✓		3
starch degradation III	OA	✓		✓		✓	3
starch degradation V	OA					✓	4
	YA	✓	✓	✓			

thiazole biosynthesis II (aerobic bacteria)	YA	✓	✓	✓			3
superpathway of thiamine diphosphate biosynthesis II	YA	✓	✓	✓			3
thiamine salvage II	OA					✓	5
	YA	✓	✓	✓	✓		
androstenedione degradation	OA	✓	✓	✓		✓	4
TCA cycle V (2-oxoglutarate:ferredoxin oxidoreductase)	YA	✓	✓	✓	✓		4
fatty acid salvage	OA		✓	✓	✓		3
pyrimidine deoxyribonucleotides *de novo* biosynthesis I	OA					✓	5
	YA	✓	✓	✓	✓		
pyrimidine deoxyribonucleotides *de novo* biosynthesis II	OA					✓	5
	YA	✓	✓	✓	✓		
superpathway of pyrimidine ribonucleosides salvage	YA	✓	✓	✓	✓		4
pyrimidine deoxyribonucleotide phosphorylation	OA					✓	5
	YA	✓	✓	✓	✓		
pyrimidine deoxyribonucleosides salvage	OA					✓	5
	YA	✓	✓	✓	✓		
superpathway of pyrimidine deoxyribonucleoside salvage	OA					✓	5
	YA	✓	✓	✓	✓		
superpathway of pyrimidine nucleobases salvage	YA	✓	✓	✓	✓		4
adenosine ribonucleotides *de novo* biosynthesis	YA	✓	✓	✓	✓		4
adenosine deoxyribonucleotides *de novo* biosynthesis II	YA	✓	✓	✓	✓		4
guanosine ribonucleotides *de novo* biosynthesis	OA					✓	5
	YA	✓	✓	✓	✓		
guanosine deoxyribonucleotides *de novo* biosynthesis II	YA	✓	✓	✓	✓		4
superpathway of guanosine nucleotides *de novo* biosynthesis I	YA	✓	✓	✓	✓		4
superpathway of adenosine nucleotides *de novo* biosynthesis I	YA	✓	✓	✓	✓		4
inosine-5'-phosphate biosynthesis III	YA	✓	✓	✓	✓		4
myo-, *chiro*- and *scillo*-inositol degradation	OA		✓	✓	✓		3
D-fructuronate degradation	YA	✓	✓	✓			
TCA cycle VII (acetate-producers)	OA					✓	4
	YA	✓	✓	✓			
taxadiene biosynthesis (engineered)	OA					✓	3
	YA		✓	✓			
L-arginine biosynthesis IV (archaebacteria)	OA					✓	5
	YA	✓	✓	✓	✓		
aromatic biogenic amine degradation (bacteria)	YA	✓	✓	✓			3
6-hydroxymethyl-dihydropterin diphosphate biosynthesis III (Chlamydia)	OA					✓	5
	YA	✓	✓	✓	✓		
methylerythritol phosphate pathway II	YA		✓	✓	✓		3
oleate biosynthesis IV (anaerobic)	OA		✓	✓	✓		3
superpathway of purine nucleotides *de novo* biosynthesis I	OA					✓	4
	YA	✓	✓	✓			
mycolate biosynthesis	OA		✓	✓	✓		3
NAD salvage pathway I	OA					✓	4
	YA	✓	✓	✓			
pyridoxal 5'-phosphate biosynthesis I	OA		✓	✓	✓	✓	4
TCA cycle VIII (helicobacter)	YA	✓	✓	✓			3
flavin biosynthesis I (bacteria and plants)	OA					✓	5
	YA	✓	✓	✓	✓		

formaldehyde oxidation I	YA	✓	✓	✓			3
superpathway of L-serine and glycine biosynthesis I	OA					✓	4
	YA	✓	✓	✓			
superpathway of L-threonine biosynthesis	YA	✓	✓	✓	✓		4
superpathway of L-threonine biosynthesis	OA					✓	5
	YA	✓	✓	✓	✓		
tRNA charging	YA	✓	✓	✓	✓		4
L-tyrosine degradation I	OA		✓	✓	✓		3
superpathway of ubiquinol-8 biosynthesis (prokaryotic)	OA		✓	✓	✓		3
UDP-N-acetyl-D-glucosamine biosynthesis I	OA					✓	5
	YA	✓	✓	✓	✓		

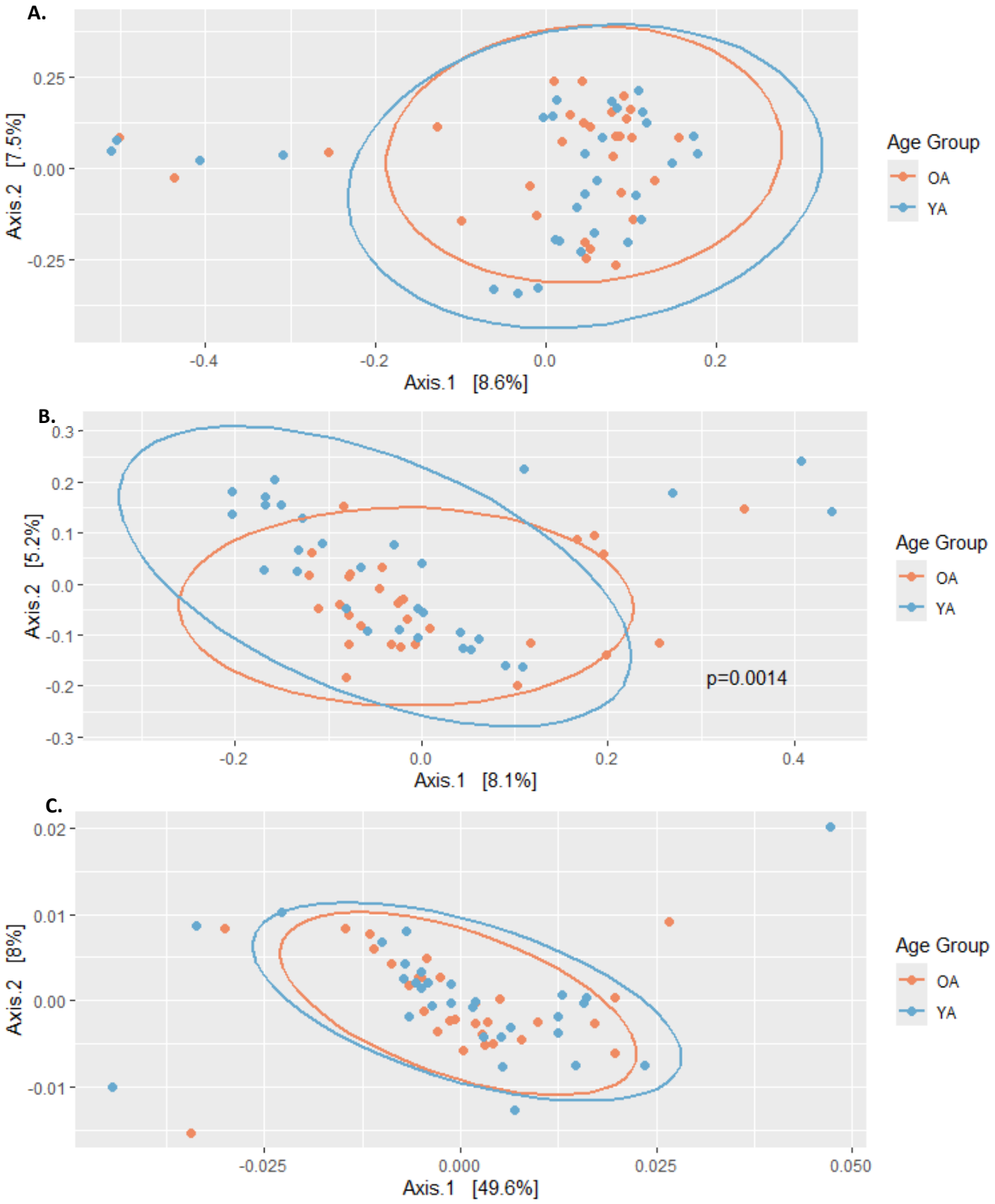


Figure 9.18 Beta diversity as an ordination plot of oral samples grouped by age group. *A. Jaccard Index. B. Unweighted UniFrac. C. Weighted UniFrac.*

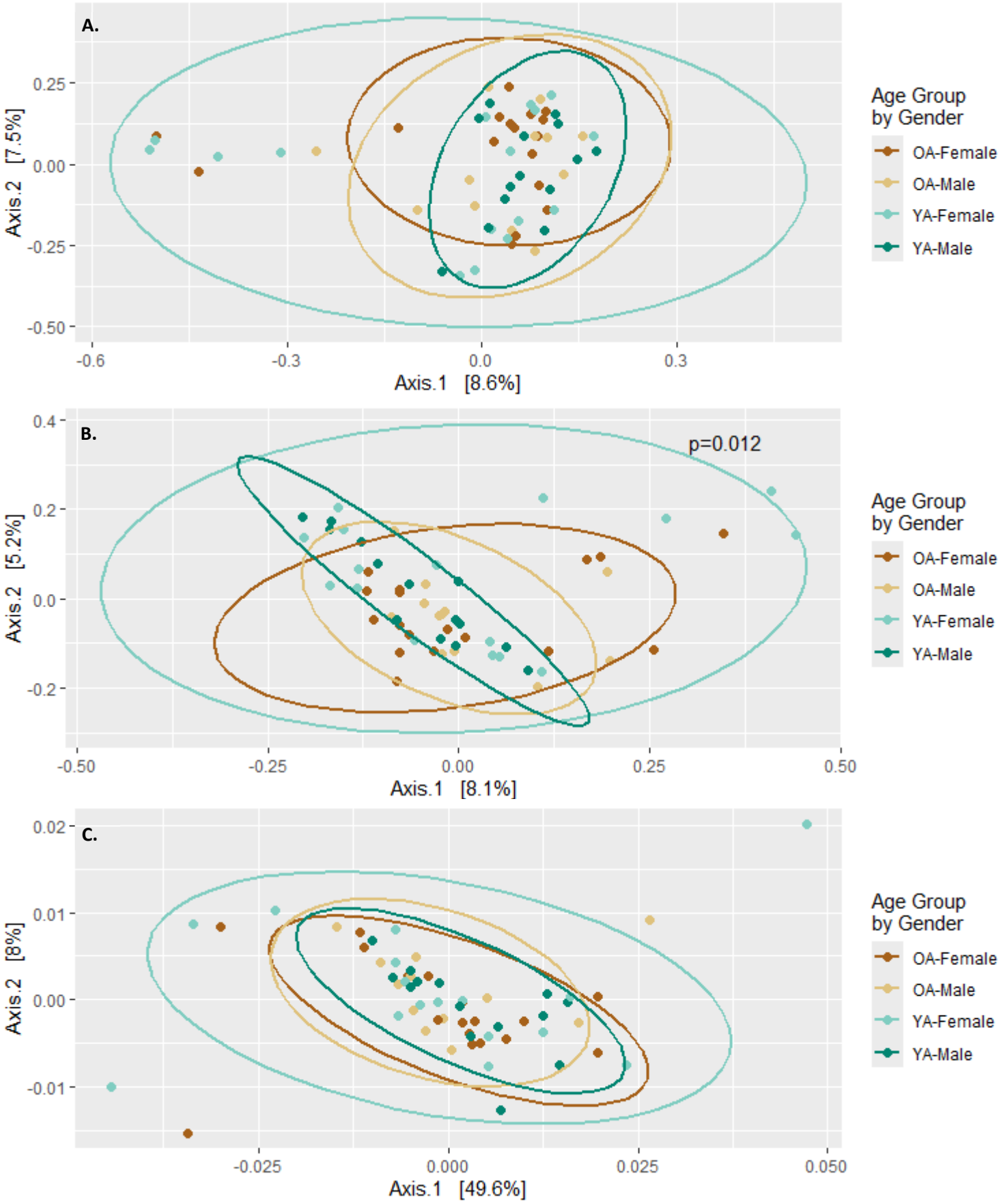


Figure 9.19 Beta diversity as an ordination plot of oral samples grouped by age group and gender. *A. Jaccard Index. B. Unweighted UniFrac. C. Weighted UniFrac.*

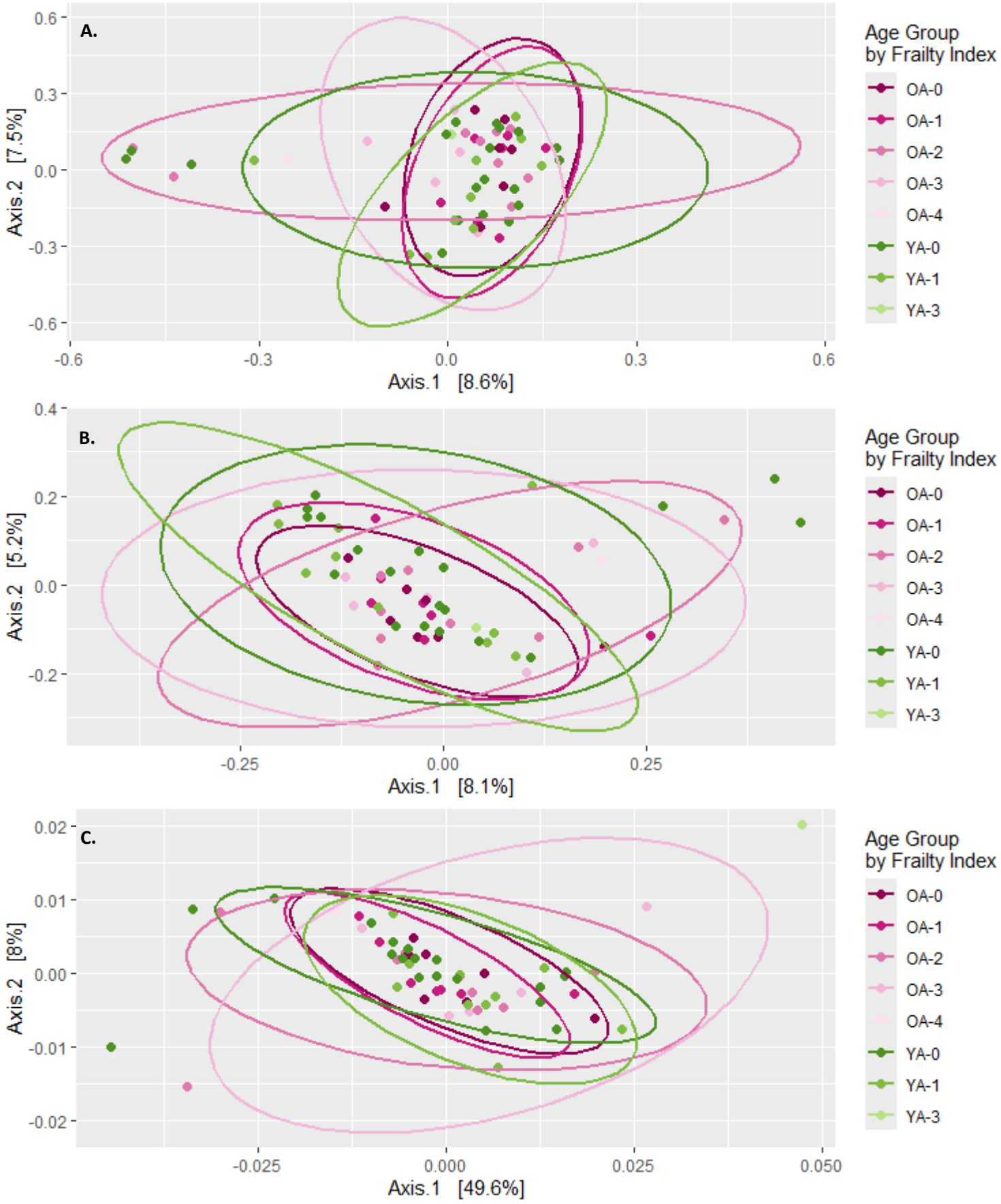


Figure 9.20 Beta diversity as an ordination plot of oral samples grouped by age group and Frailty Index. *A. Jaccard Index. B. Unweighted UniFrac. C. Weighted UniFrac.*

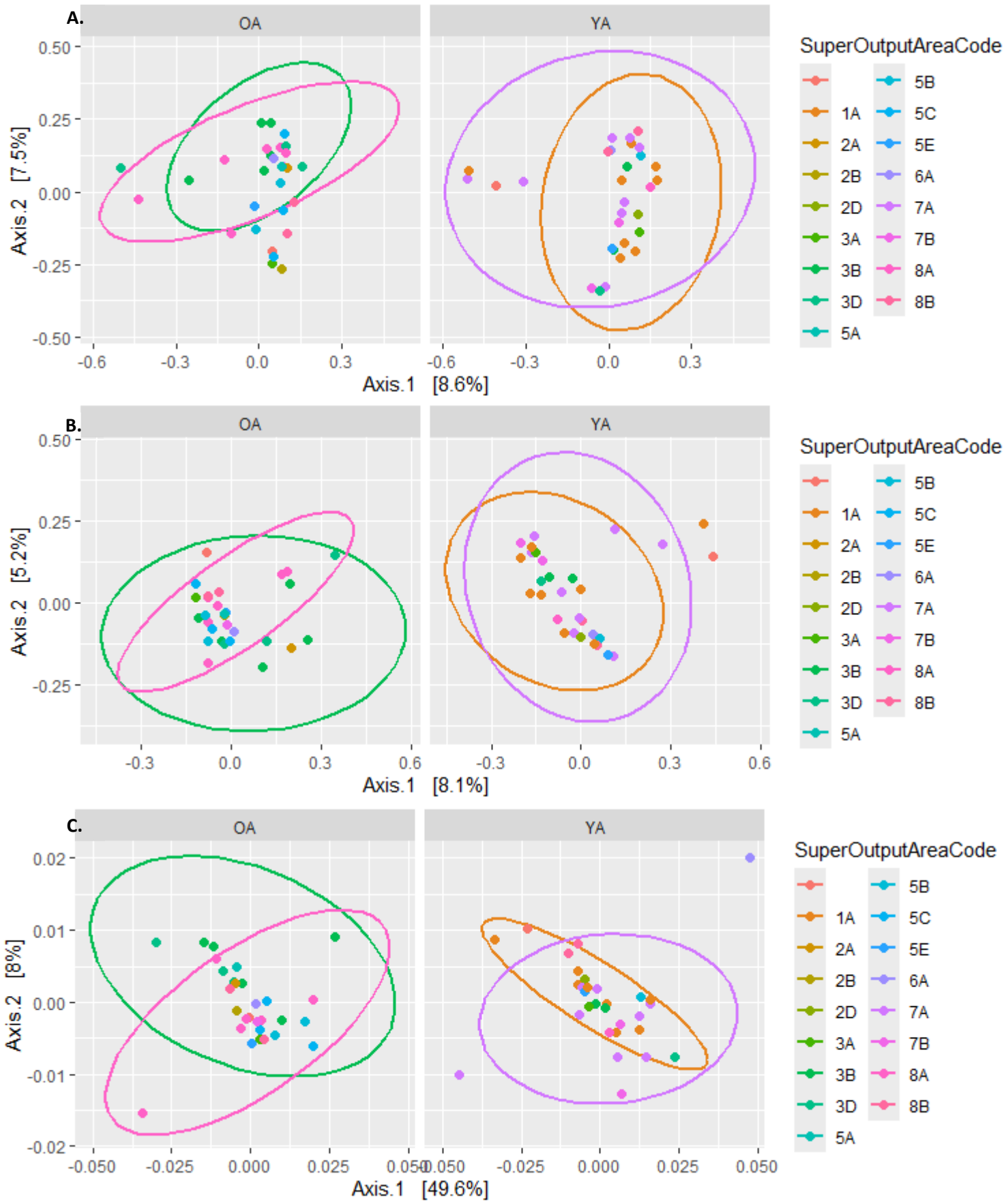


Figure 9.21 Beta diversity as an ordination plot of oral samples grouped by age group and Super Output Area Code. *A. Jaccard Index. B. Unweighted UniFrac. C. Weighted UniFrac.*

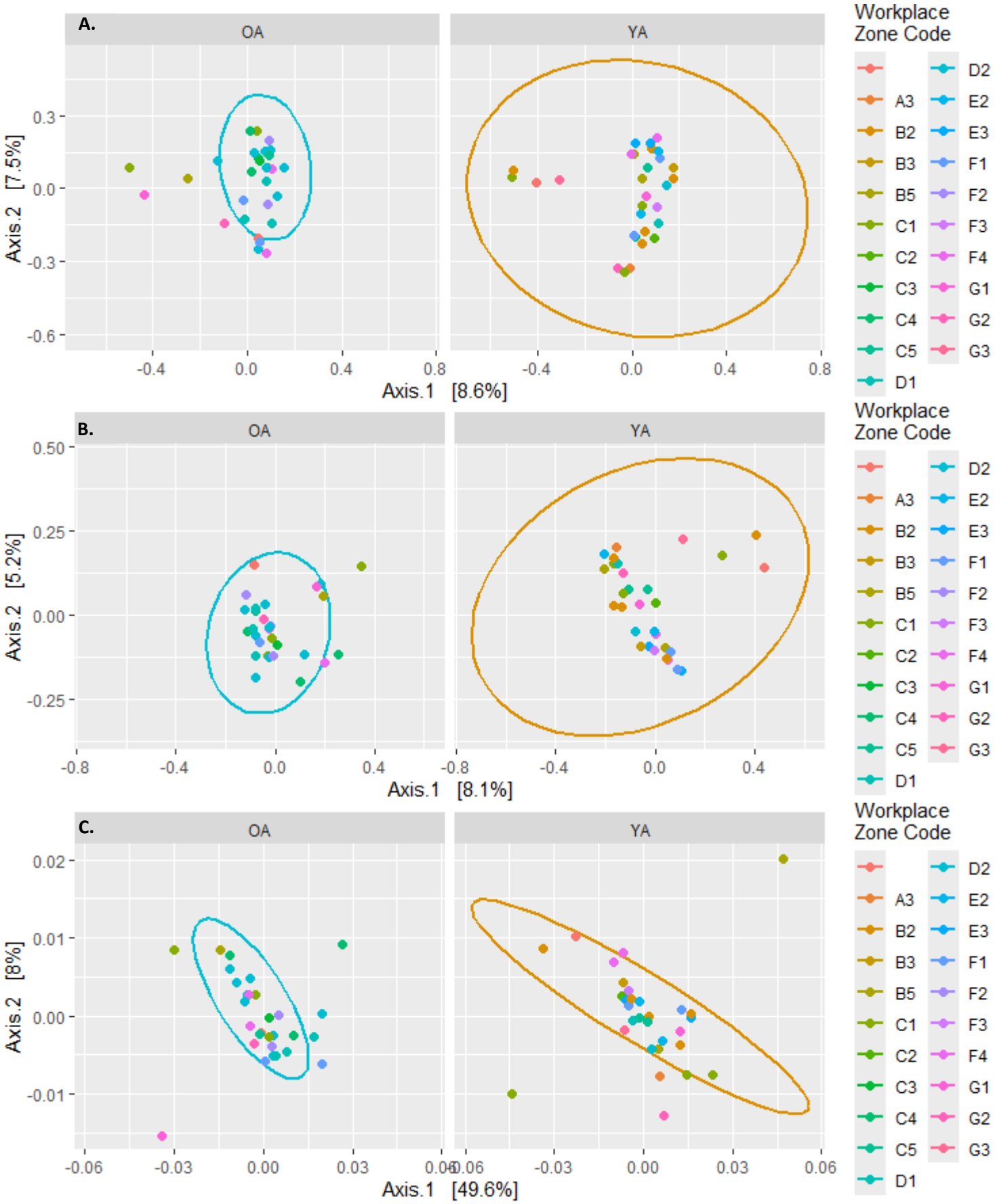


Figure 9.22 Beta diversity as an ordination plot of oral samples grouped by age group and Workplace Zone Code. *A. Jaccard Index. B. Unweighted UniFrac. C. Weighted UniFrac.*

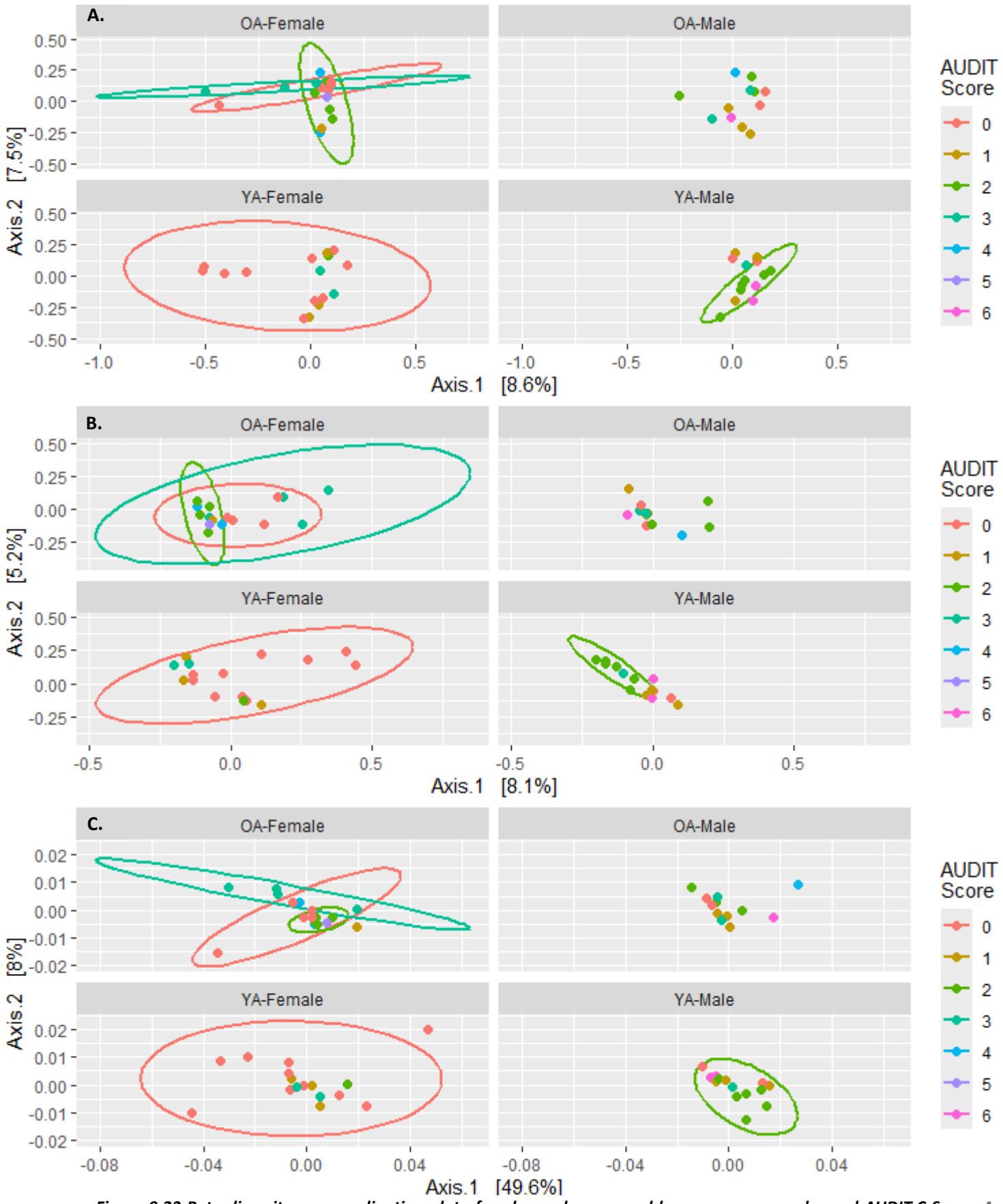


Figure 9.23 Beta diversity as an ordination plot of oral samples grouped by age group, gender and AUDIT-C Score. *A. Jaccard Index. B. Unweighted UniFrac. C. Weighted UniFrac.*

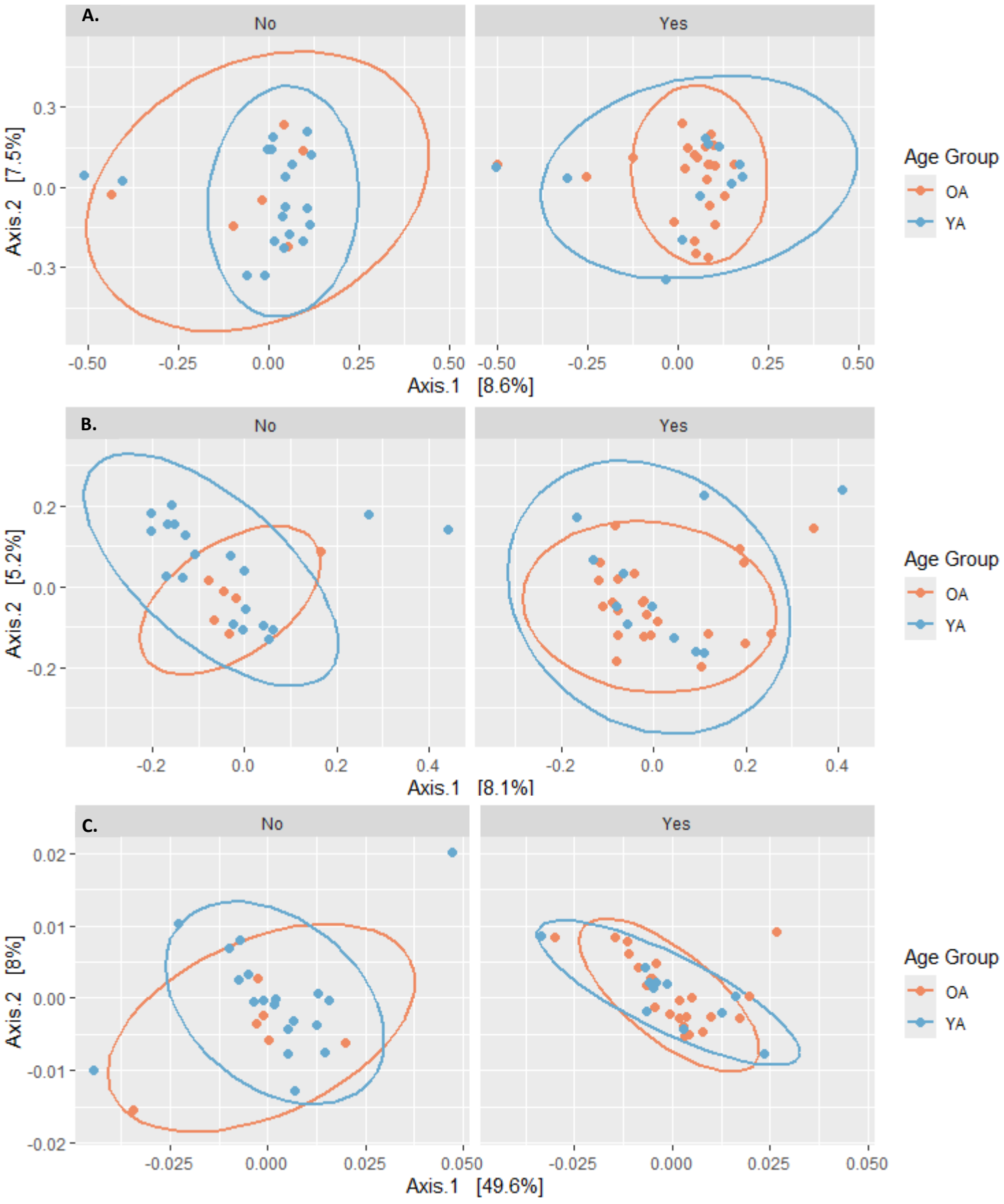


Figure 9.24 Beta diversity as an ordination plot of oral samples grouped by age group and if the participants are missing any non-wisdom adult teeth. *A. Jaccard Index. B. Unweighted UniFrac. C. Weighted UniFrac.*

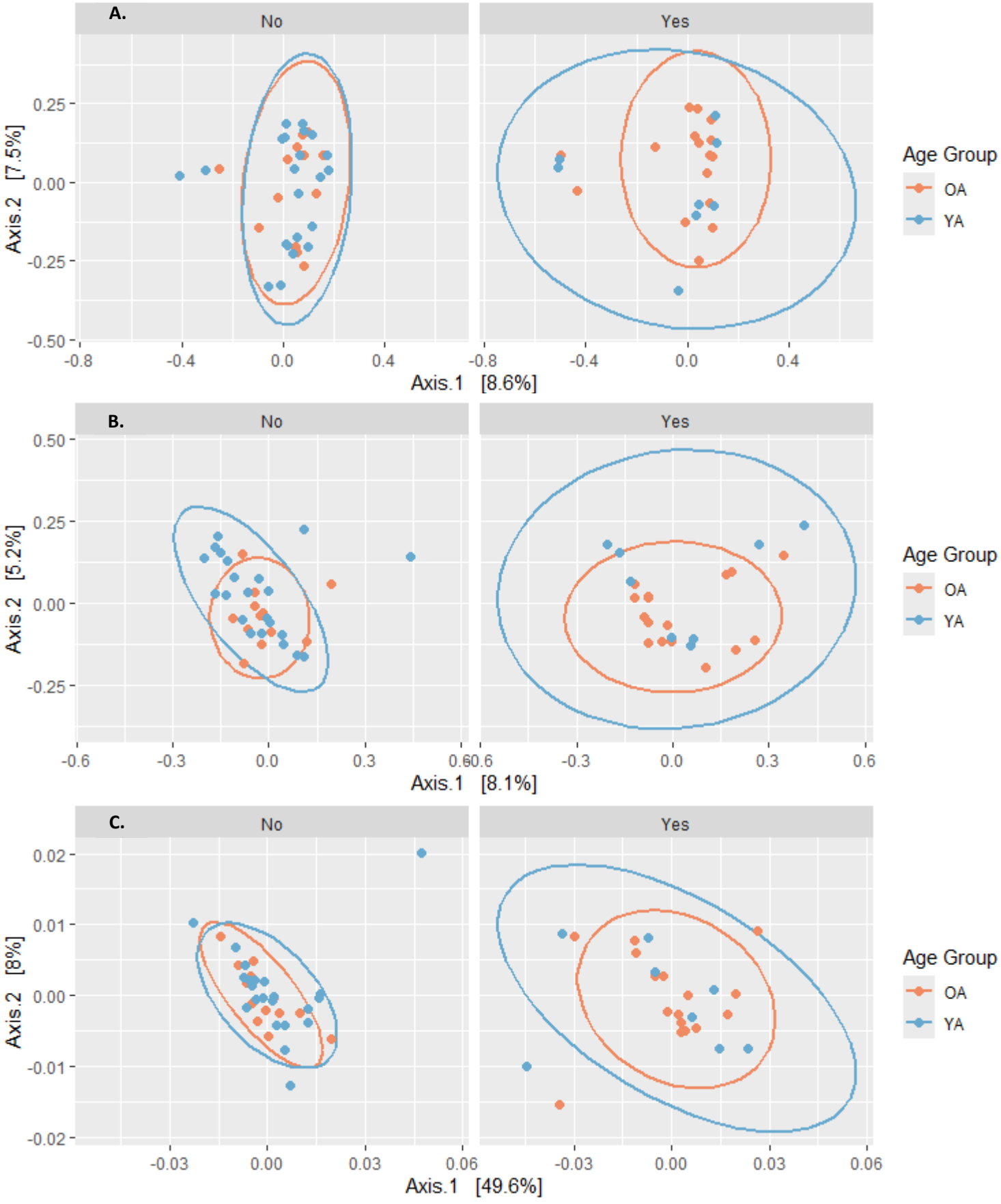


Figure 9.25 Beta diversity as an ordination plot of oral samples grouped by age group and if the participants had visited a dental hygienist within the 6 months prior to sampling. *A. Jaccard Index. B. Unweighted UniFrac. C. Weighted UniFrac.*

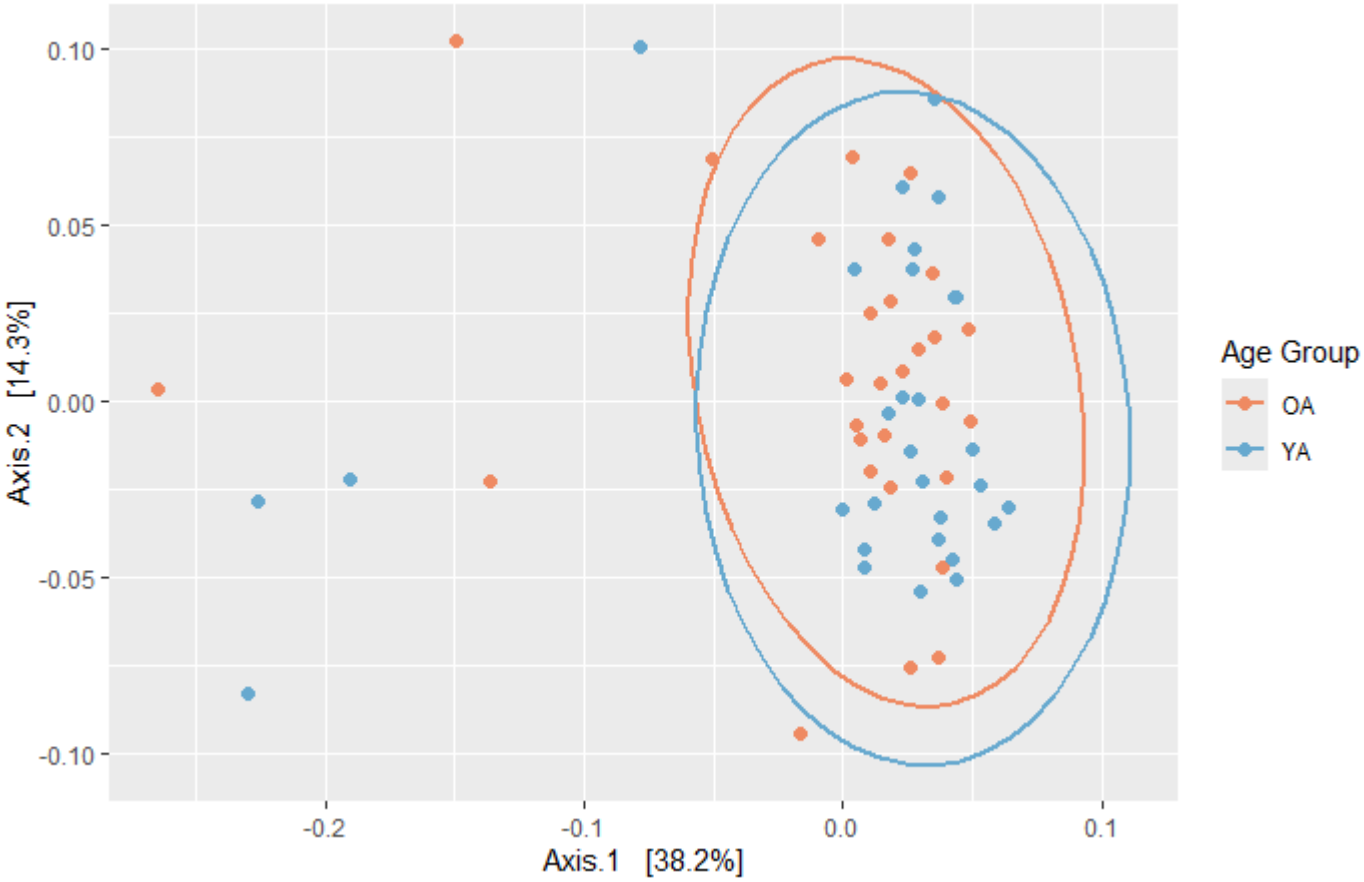


Figure 9.26 Jaccard Index beta diversity as an ordination plot for metacyc pathway abundance data from the oral dataset grouped by age group.

Table 9.6 Metacyc pathways identified as abundant in more than 3 of the 5 abundance tests carried out on the face predictive metabolomics dataset, when grouped based on age group. ***Any pathway identified through all five methods and have an agreed abundant group are highlighted in green.***

Descriptor	Abundant Group	Methods Identified by					Count
		Ancombc	Deseq	Edger	Lefse	Limma-Voom	
superpathway of arginine and polyamine biosynthesis (from L-glutamate)	OA	✓	✓	✓	✓	✓	5
superpathway of L-arginine, putrescine, and 4-aminobutanoate degradation	OA	✓	✓	✓		✓	4
fatty acid β-oxidation I (generic)	YA	✓	✓	✓	✓	✓	5
superpathway of hexuronide and hexuronate degradation	OA	✓	✓	✓	✓	✓	5
superpathway of L-arginine and L-ornithine degradation	OA	✓	✓	✓		✓	4
protocatechuate degradation I (*meta*-cleavage pathway)	OA	✓	✓	✓		✓	4
superpathway of polyamine biosynthesis I	OA	✓	✓	✓	✓	✓	5
superpathway of pyrimidine deoxyribonucleotides *de novo* biosynthesis (*E. coli*)	YA		✓	✓	✓	✓	4
L-arginine biosynthesis III (via *N*-acetyl-L-citrulline)	YA		✓	✓		✓	3
glutaryl-CoA degradation	YA	✓	✓	✓	✓	✓	5
superpathway of sulphur oxidation (*Acidianus ambivalens*)	OA		✓	✓	✓		3
sucrose degradation IV (sucrose phosphorylase)	YA		✓	✓	✓	✓	4
superpathway of guanosine nucleotides *de novo* biosynthesis II	YA		✓	✓	✓		3
GDP-D-*glycero*-α-D-*manno*-heptose biosynthesis	OA	✓	✓	✓	✓	✓	5
pyrimidine deoxyribonucleotides *de novo* biosynthesis I	YA		✓	✓	✓		3
pyrimidine deoxyribonucleotides *de novo* biosynthesis II	YA		✓	✓	✓	✓	4
superpathway of pyrimidine ribonucleosides salvage	YA		✓	✓	✓	✓	4
pyrimidine deoxyribonucleotide phosphorylation	YA		✓	✓	✓		3
superpathway of guanosine nucleotides *de novo* biosynthesis I	YA		✓	✓	✓		3
inosine-5'-phosphate biosynthesis III	YA		✓	✓		✓	3
TCA cycle VII (acetate-producers)	OA		✓	✓	✓	✓	4
CA cycle VI (Helicobacter)	OA		✓	✓	✓	✓	4