Impact of Chewing Betel Nut on the Oral Microbiome

BY

Scott D. Drucker
B.A., University of Pennsylvania, 2010
D.M.D., University of Pennsylvania, 2013

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Master of Science in Oral Sciences
in the Graduate College of the
University of Illinois at Chicago, 2016

Chicago, IL

Defense Committee:

Lin Tao, Chair and Advisor
Herve Y Sroussi, Oral Medicine and Diagnostic Sciences
Joel L Schwartz, Oral Medicine and Diagnostic Sciences

This thesis is dedicated to my family. Thank you for your loving support and guidance.

ACKNOWLEDGEMENTS

I would like to thank my committee members Dr. Lin Tao, Dr. Joel Schwartz, and Dr. Herve Sroussi. This project would not have been possible without their expertise and guidance. Specifically, I would like to thank Dr. Tao for his unwavering guidance and confidence in my ability to earn a master's degree.

I would like to thank Ashley Burton, who helped in reviewing and submitting the manuscript. Thank you!

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

LIST OF ABBREVIATIONS

ANOSIM	Analysis of Similarity
MDS	Multidimensional Scaling
NGS	Next Generation Sequencing
OSF	Oral Submucosal Fibrosis
OTU	Operational Taxonomic Unit
PCR	Polymerase Chain Reaction
SIMPER	Similarity of Percentages
UIC	University of Illinois at Chicago
WHO	World Health Organization

SUMMARY

Consumption of betel nut products dates back thousands of years, and is current common practice across South Asia and Oceania. In recent years, the World Health Organization (WHO) has publicly stated that betel quid products pose a substantial health threat to the populations that consume them. The International Agency for Research on Cancer has long determined that betel quid without tobacco is carcinogenic to humans, and areca nut is carcinogenic to humans. Conclusions were drawn from strong evidence of betel nut causing precancerous oral submucosal fibrosis (OSF) conditions, as well as sufficient evidence of carcinogenicity in animals.

Importance of local microbiome composition and its contribution to disease status in humans has become a topic of increased investigation, specifically with regard to gastrointestinal cancer. The effects of chewing betel nut on the composition of the oral microbiome have not been studied. Due to recent advancements in technology, oral microbiomes can be analyzed effectively with a combination of the Polymerase Chain Reaction (PCR) technique and next-generation sequencing hardware and software.

This study assessed the microbiome diversity within samples taken from betel nut chewers and controls, and additionally compared microbial composition between samples. The initial results of α-diversity tests indicate that chewing betel quid appears to increase the species diversity in the oral microbiome. Comparisons between samples for β-diversity analysis indicated that the control samples contained oral microbiomes that were much more closely related than the samples taken from betel quid chewers. Early findings suggest chewing betel quid does correlate with a microbial composition shift, which may play a role in development of oral cancer.

I. REVIEW OF LITERATURE

A. BETEL NUT NOMENCLATURE

Despite its near-ubiquity in certain parts of the world, the term "betel nut" is used to refer to various different products and substances. The areca nut is the seed of the areca palm (*Areca catechu*). It naturally grows in much of the tropical Pacific, Asia, and parts of east Africa, and is the central component of a mixture generally referred to as "betel nut." The *Piper betel* is a vine distinct from areca palms. It belongs to the same family as pepper and kava plants, and produces wide, distinct leaves.

The areca nut is typically consumed with the leaf of the betel vine, leading many to refer to the nut itself as "betel nut," despite the misleading terminology. The nut and leaf are mixed with smeared slaked lime and other additional items, often tobacco. The complete mixture is often referred to as "betel quid," of which the areca nut is the defining ingredient (Gupta et al. 2014). The use of slaked lime (calcium hydroxide) acts to significantly raise oral pH, which may bolster the effect of the nervous system-stimulating chemicals, such as arecoline in the areca nut and nicotine in tobacco (Cawte 1985).

The betel quid mixture can vary dramatically from location to location and include other ingredients ranging from acacia extract to local spices. Colloquially, betel quid is simply referred to as *betel nut*, owing to the fact that areca nuts are the defining ingredient and areca nuts are typically consumed in such fashion. Nonetheless, additional names for betel quid abound in various regions of the world, and vary based on the language of the location as well as the composition of the particular mixture. In India, for

instance, the mixture is referred to as *gutkha* or pan masala, depending on whether or not the mixture contains tobacco.

In this report, "betel nut" will be treated as synonym to "betel quid," unless otherwise specified.

B. HISTORY OF BETEL NUT CHEWING

There is a wide range of linguistic diversity attributed to the areca nut, betel leaf, and betel quid, which testifies to the long history of these product consumptions in various cultural and geographic realms (Reid 1985). The long-term trends of areca use are complex and variable from region to region. It is understood that ancestors of current Polynesians established a wide-ranging pottery trade, and in doing so effectively colonized much of what is currently inhabited South Western Pacific around 3,600 years ago. During this time of economic expansion, Polynesians likely introduced the consumption of areca nut to many different locales (Ebot et al. 1992).

The earliest scientific evidence indicating the use of areca came from the Spirit Cave in Northwestern Thailand, where petioles of the areca fruit have been carbon dated back to roughly 8,000-9,000 years ago (Gorman 1970). The habit of chewing is thought to have originated in modern day Southeast Asia, most likely Malaysia (Norton 1998). The use of betel quid may have spread around the time of the early Christian period, as during the first two centuries AD, maritime trade was encouraged among denizens of the South Asian subcontinent, the Middle East, and Southeast Asia (McPherson 1993).

Both the practice of consuming betel quid and the mixture itself differ greatly in detail depending on the geographic region, culture, and historical stages. Use of the areca

nut is very common as an act of social graces, goodwill, and camaraderie. In India and much of the rest of the South Asian subcontinent, the betel leaf and areca nut combination is not only used in social settings and Ayurvedic healing remedies, but in religious ceremonies as well (Da Costa 2012).

C. CONSUMPTION GEOGRAPHY

The areca nut is indigenous to the tropical Pacific and areas surrounding the Indian Ocean. As such, it is cultivated and considered to be in the natural makeup of the indigenous plant species in a number of regions and countries, including southern China, India, Nepal Bangladesh, the Maldives, Cambodia, Thailand, Laos, Indonesia, Malaysia, New Guinea, Sri Lanka, and Pakistan (Jones 2001). Not surprisingly, its consumption overlaps geographically with its growth (Fig. 1). The areca nut is commonly used between the geographic boundaries of longitudes 68°-168° and latitudes 20° - 30° north (Reichart et al. 2006). Researchers have had difficulty pinpointing the global number of areca users; a well-accepted estimate places the number near 600 million individuals (Mack 2001; WHO 2012).

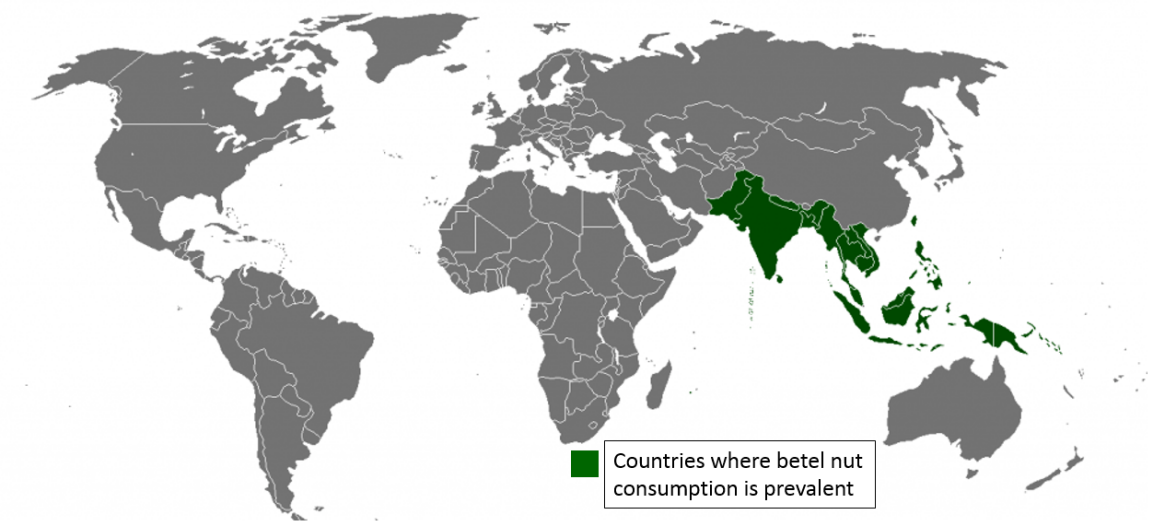

Fig 1. A world map indicating countries with high levels of betel nut consumption.

Betel nut remains widespread in regions the habit originated, ranging from the Western Pacific Islands to Southeast Asia to the Indian subcontinent. In the Western Pacific region, it is most commonly chewed in Papua New Guinea, Solomon Islands, Vanuatu, Federated States of Micronesia, and rural regions of Cambodia and Vietnam. In these regions, the frequency and method of betel nut consumption varies, although the areca nut, betel leaf, and lime are relatively consistent components of popular betel quid mixtures. In some regions of Papua New Guinea, the areca nut is first chewed for a few seconds before lime powder (calcium hydroxide) is added to the mouth with a spatula; in other regions, the nut is cracked, lime added, and the mix is wrapped in betel leaf prior to chewing (Gupta and Ray 2004).

The Indian subcontinent, comprised of India, Pakistan, and Bangladesh, is another region where betel nut chewing is widely pervasive. Pan masala (betel quid which includes areca nut, catechu, lime, and spices), and Gutka or Guṭkha (betel quid with tobacco and proprietary ‘other’ contents) were first introduced as commercial products in

1980 (Dyavanagoudar 2009). Gutkha is one of the most widely used smokeless tobacco products in the Indian subcontinent (World Health Organization 2009-2010; Stanfill et al. 2011). In Hindu culture, there is extensive Ayurvedic description of experience induced by the areca nut, detailing reasons for use of the plant. The areca nut is believed to play roles ranging from a purgative to intoxicant aiding in digestion and removing bad tastes (Strickland 2007; Meulenbeld et al. 1987). In India, the preparation is made with or without tobacco, and referred to as paan, or paan masala (Da Costa 2012). Betel nut chewing is not limited to Hindu regions; it is also extremely pervasive in the Muslim countries of Pakistan and Bangladesh as well.

The use of betel quid has spread quite extensively in recent years, as migration patterns have taken consumers of betel quid to areas not previously exposed to the product. For instance, use has expanded to parts of the United States of America, United Kingdom, Australia, and New Zealand (Bedi 1996; Hirsch 1995; WHO 2012). It appears that the ever-expanding availability of prepackaged areca mixtures has also perpetuated the current spread and increasing popularity of the nut (Bhonsle et al. 1992). Recently, the betel nut chewing practice has become increasingly popular in China, especially in the Hunan province (Zhang et al. 2012). A wide variety of areca nut products are readily available globally for purchase on the internet.

Furthermore, betel quid and other areca nut products are readily available in cities in the United States with large communities of South Asians. In Chicago, such products are prominently displayed on the counter of many shops on Devon Street, a hub for the Indian, Pakistani, and Bangladeshi communities. Bright red spits of betel nut chewers covering sidewalks of the street are not only a local eyesore but also an indication that a

sizable population in Chicago consumes betel nut and is at risk for betel nut-associated diseases including oral cancer.

D. LEGAL CLIMATE OF BETEL NUT SALES

The World Health Organization (WHO) has publicly stated that betel quid products pose a substantial health threat to the populations that consume them. The WHO specifically focused on the Western Pacific region, another geographic area where betel nut chewing is part of the social fabric of society. Betel quid products both with and without tobacco have been associated with increased incidence of oral cancer. In fact, betel nut is classified by the International Agency for Research on Cancer (IARC, 2004) as a Group 1 carcinogen. The WHO concluded that a major effort needs to be undertaken to provide decision-makers with evidence of serious harm caused by betel nut chewing. The WHO further added that public health measures, including policies and legislation, can save lives and prevent unnecessary suffering from oral cancer and other diseases (IARC Monograph 2004).

1. India. In India, the government has required sellers to include a warning label on commercial areca-nut and tobacco products. However, there are no regulations about the size of the letters. Additional regulations, restrictions, and bans vary from state to state within India. For instance, *gutkha* has been banned by Tamil Nadu, Adhra Pradesh and Goa state governments. Several other states are at various stages of considering and passing similar laws, and are fighting against industry figures. The Central Committee on

Food Safety has recommended that *gutkha* be banned nationwide, although such a ban has not of yet been passed nationwide (IARC, 2004).

In 2002, the Commissioner for Food and Drug Administration and Food (Health) Authority of Maharashtra State banned the manufacture, sale and storage of *gutkha* and *pan masala* or any similar product containing or not containing tobacco (Sharma, 2002). The ban, as well as a subsequent one, was struck down by the Supreme Court, but more recent bans have been upheld and proven to be largely effective (Deshpande 2012, Wal 2015).

2. North America. In Canada, areca nuts are listed on the schedule of herbs that are unacceptable as a non-medicinal ingredient in oral use products (Health Canada, 1995). In the United States, however, the Food and Drug Administration maintains an import alert within the USA, although the primary concern cited is the potential adulteration of the product and addition of unsafe food additives, rather than health concerns (Croucher & Islam, 2002). In 1976, the US Government banned interstate traffic of areca nut (Burton-Bradley, 1978). Additional localized restrictions exist in the US. For instance, possession of areca nut in the California public school system is grounds for suspension (Croucher & Islam, 2002).

3. European Union and United Kingdom. There are no specific laws or regulations regarding the sale or consumption of areca products, even when mixed with smokeless tobacco (Council of the European Communities, 2001). In the UK, no law regulates the import or sale of products containing areca nut. As a result, numerous areca preparations and betel quid variations, both with and without tobacco, are commercially available (Bedi, 1996; Vora *et al.*, 2000). Areca derivatives are particularly easy to

acquire in the UK, partly due to its lax laws. In fact, the UK's Department of Trade and Industry classifies such products as sweets (Hogan, 2000). Additionally, the packaging often fails to properly include labels or list ingredients. Minors and children have little difficulty acquiring betel quid in the UK (Shetty & Johnson, 1999; Warnakulasuriya *et al.*, 2002, National Centre for Transcultural Oral Health, 2001). By and large, packages do not include detailed health warnings beyond the statutory display regarding tobacco use. Among 20 commercially processed areca-nut products on sale in the UK, only three carried a health warning related to oral cancer, and none addressed potential addiction (Trivedy, 2001).

4. Other Countries. Certain Southeast Asian countries have instituted outright bans or other restrictions on the consumption of betel quid. In Papua New Guinea in the late 1970s, the Public Services banned the chewing of betel quid in government offices (Burton-Bradley, 1978). In Singapore, spitting in public places is an offense punishable by a fine, thus discouraging individuals from chewing (and spitting) betel quid (Cheong, 1984). Betel quid, as well as its betel and areca components, are banned in the United Arab Emirates (United Arab Emirates).

E. NEUROLOGICAL EFFECT

The biochemistry of betel quid can be best understood by analyzing each main ingredient in turn. Chewing the areca nut itself results in the release of nine alkaloids, the most abundant of which are arecoline, arecaidine, guvacine, and guvacoline (IARC, 2004). The alkaloids undergo a process called nitrosation, giving rise to N-nitrosamine, which is cytotoxic (Hoffman et al. 1994). Betel quid typically contains lime powder

(calcium oxide or quicklime) or slaked lime (calcium hydroxide) in the mixture (MacLennan et al. 1985). The lime acts to enhance the stimulant effect of betel nut as it hydrolyzes arecoline to form arecadine, a central nervous system stimulant (Norton 1998). Part of the betel vine may be added to enhance flavor. The betel leaves contain phenols with a spicy flavor.

The stimulating effects of betel nut can be felt within minutes of the start of chewing. The quid is typically parked against oral mucosa, a tissue highly permeable to alkaloids. The alkaloids activate the sympathetic nerves to induce a wide range of symptoms, including: dizziness and heart palpitations, heightened awareness, hot sensation and sweating, increased respiratory rate, and diminished hunger and thirst (Rooban et al 2005). Effects are typically more pronounced in first time users (Cawte 1985).

F. ORAL CANCER ASSOCIATION

Squamous cell carcinoma accounts for over 90% of all oral cancers worldwide (Pindborg 1980; Jamrozik 1985; Forastiere et al. 2001). Approximately 22,000 Americans are diagnosed with oral cancer each year, of which 90% are squamous cell carcinomas. Despite drastic advances in medicine over the past few decades, the 5 year survival rate has remained 40% over the last 40 years. There are estimated to be 350,000-400,000 new cases diagnosed each year. The incidence of oral cancer has actually risen over the years; this rise in rate cannot be accounted for by the major oral cancer risk factors of tobacco and alcohol use (Parkin 1999, Shiboski 2005, Schmidt 2004).

The association between betel nut chewing and oral cancer has been known since at least the early 20th century (Orr 1933; Eisen 1946). In 1956, Paymaster observed slow growing squamous cell carcinoma in one third of his patients with OSF. In 1972, Pindborg and others established five criteria that confirmed the disease is precancerous: high occurance of OSF in oral cancer patients, increased incidence of squamous cell carcinoma in patients with OSF, histologic diagnosis of cancer despite lack of clinical suspicion, and high frequency of epithelial dysplasia and prevalence of leukoplakia among OSF patients (Pindborg 1966). OSF has traditionally been the most reliable indicator for predicting potential malignant transformation of an oral precancerous lesion (Warnakulasuriya 2001). Malignant transformation of OSF was found to range from 7-13% (Tilkaratne 2006). The International Agency for Research on Cancer has long determined that betel quid with tobacco is carcinogenic to humans, betel quid without tobacco is carcinogenic to humans, and areca nut is carcinogenic to humans (IARC, 2004). These conclusions were drawn from strong evidence of betel nut causing precancerous oral submucosal fibrosis (OSF) conditions, as well as sufficient evidence of carcinogenicity in animals.

In regions where betel nut chewing is common, oral squamous cell carcinomas are most commonly found in the buccal mucosa from the corner of the mouth and posterior, as well as in the lateral boarder of the tongue. In contrast, in western countries where betel nut chewing is not commonplace, the most common sites of oral squamous cell carcinoma are floor of the mouth and tongue (Thomas & MacLennan 1992). This dichotomy lends further credence indicating that betel quid causes squamous cell carcinomas, particularly in areas of the mouth where it is most frequently applied.

G. PRECANCEROUS LESIONS

Betel nut chewers' oral mucosa is often characterized by a brownish-red coloration adjacent to the site the betel quid is parked. This observation was first described in the early 1970s (Mehta et al. 1972). Oral leukoplakia, erythroplakia, lichenoid and OSF are the most common precancerous lesions associated with betel nut chewing.

Leukoplakia is defined as a white patch on oral mucosa that cannot be characterized clinically as another lesion. Malignant conversion rates for leukoplakia are 0.1%-17.5% (Van der Waal et al. 2002). Reported annual global incidences for oral cancer due to leukoplakia are 6.2-29.1 cases per 100,000 people (Downer and Petti 2005). A statistically significant interaction between tobacco and betel quid in the causation of oral leukoplakia has been reported (Lee et al. 2003).

Erythroplakia is defined as a red patch on a mucous membrane that cannot be attributed to any other pathology (Hashibe et al. 2000). The frequency and duration of betel nut chewing have been positively associated with occurrence of leukoplakia and erythroplakia (Lee et al. 2003; Hashibe et al. 2000; Gupta 1984).

In addition to areca nut, lime powder in betel quid is also a direct irritant of the oral mucosa, abrading the tissue and frequently leading to ulceration. Lichenoid lesions have been observed histologically at sites of betel quid usage (Daftary et al. 1980).

Betel nut chewing is the most important etiological factor in the development of OSF (Warnakulasuriya et al. 1997). This etiologic relationship has been demonstrated in both epidemiologic and *in vitro* experimental studies (Caniff & Harvey 1981). OSF is a

characterized by persistent stomatitis leading to progressive sclerosis of muscular and dermal tissues of the oropharynx (Murti et al. 1985; Norton 1998). Over time, the fibrosis advances and eventually may restrict both opening of the oral cavity and protrusion of the tongue. Quality of life diminishes as eating, swallowing, and speech become impeded (Rajendran 1994). Transformation rates to malignancy are 3-7.6% (Murti et al. 1985; Pindborg et al. 1984; Sinor et al. 1990).

Fibrosis and hyalinization of subepithelial tissue yields the clinical features seen in OSF. Recent research has focused on changes in the extracellular matrix; it is logical to hypothesize that there is an increase in collagen synthesis or decrease in collagen degradation contributing to the development of the disease (Tilakartne 2006). There are many mechanisms of action thought to contribute to the progression of OSF, each detailed below.

Arecoline, arecaidine, guvacine, and guvacoline are all alkaloids found in the areca nut; these products have been shown in vivo to have pronounced effects of fibroblasts (Harvey 1986). Arecoline in doses100ug ma^{-1} was shown to be cytotoxic, and cells showed detachment from the culture surface. Separately, a dose-dependent stimulation of collagen synthesis was seen when fibroblasts were exposed to arecadine; investigators believe that arecaidine is the active metabolite in fibroblast stimulation [Harvey 1986). Addition of slaked lime to the areca nut facilitates hydrolysis of arecoline to arecaidine, making the arecaidine compound readily available in the oral cavity (Nieschultz 1968).

In addition to the alkeloids stimulating an increase in collagen production, the tannins found in the areca nut reduce degradation of collagen by forming more stable

collagen suprastructure. Tanins have also been shown been show to inhibit collagenase activity (Meghji 1982). Arecoline as also been shown to reduce MMP-2 secretion and increase TIMP-1 levels, resulting in increased deposition of collagen in the extracellular matrix.

Copper content of areca nut is quite high, and the levels of soluble copper may rise as a result in those who chew betel quid (Trivedy 1997). The association between copper levels and OSF has been made in part due to excess copper found in other fibrotic disorders such as Wilson's disease and primary biliary cirrhosis. Production of the enzyme lysyl oxidase is found to be upregulated in OSF (Trivedy 1999). Lysyl oxidase is copper dependent, and plays a distinct role in collagen synthesis, specifically in collagen cross-linking (Kagen 1991). In vivo, copper added to fibroblast culture has yielded an increase in fibroblast proliferation (Trivedy 2001).

In OSF, constant inflammation at the buccal mucosa causes an increase in cytokines IL-6, TNF-α, and TGF-β. An increase in TGF-β is thought to enhance collagen production and cross-linking (Rajalalitha & Vali 2005). Arecoline has been found to stimulate the connective tissue growth factor production in buccal mucosal fibroblasts, which may cause fibrosis. (Yi-Ting et al. 2009).

H. EFFECT ON ORAL MICROBIOME

The human body is immersed in the microbial world. As a part of the human microbiome, the oral cavity contains billions of resident microorganisms collectively referred to as oral microbiome. It affects health and disease by interacting with the host (Wade, 2013). Therefore, the well-being of the oral microbiome may be important for

betel nut chewing-associated oral diseases including cancer. However, the effects of chewing betel nut on the well-being of oral microbiome have not been studied. Oral microbiome contains hundreds of diverse species of microorganisms, among which most species are not culturable. Its analysis had been limited to the few culturable species. Due to recent advancements in technology, oral microbiomes can be analyzed effectively with a combination of the Polymerase Chain Reaction (PCR) technique and next-generation sequencing hardware and software. PCR is an *in vitro* manipulation of DNA replication, with which specific, targeted gene sequences can be copied in tremendously large quantities. This mass replication allows for specific sequences to be analyzed. Microbial ecologists frequently use PCR amplification to assess nucleic acid strands for overall microbial community composition (Pace 1986, Schmidt 1991, Reysenbach 1994).

Ecologists commonly focus on ribosomal RNA (rRNA) genes as subjects for phylogenetic analysis, as such genes are present in every organism, contain highly conserved bases for critical aspects of functionality, contain variable regions that are strongly lineage specific, and are coded to ribosomes without undergoing translation (Woese 1983). The replicated rRNA gene copies are generally referred to as amplicons. In this study, amplicon sequencing of the 16S rRNA gene was performed. The sequencing produces a robust data set for both ecological and statistical analyses, including α-diversity, β-diversity, Bray-Curtis dissimilarity, and Student's *t*-test. The results revealed that betel nut chewing leads to an increase in oral pathogens by reducing commensal microbial flora.

II. MATERIALS AND METHODS

A. PATIENT RECRUITMENT AND SAMPLE COLLECTION

1. Patient recruitment: The study subjects were recruited at the University of Illinois at Chicago (UIC) College of Dentistry dental clinics. A total of 4 patients reported a history of betel nut chewing, and an additional 4 non-chewing subjects who matched the chewers' age, gender and ethnicity were recruited. A fifth subject in the control group chewed betel nut in his early life but later quit; this subject was later dismissed from the study.

2. Consent: Subjects were provided an information sheet for obtaining informed consent. However, signatures were not collected to protect their privacy.

3. Study subject code number: To protect study subject's privacy, each subject was assigned a study subject code number. For example, the code B1 was for betel nut chewer #1 or C1 for control subject #1. The study samples, their data sheets, and oral examination data sheet were identified with these specially assigned subject code numbers, which cannot be traced back.

4. Data collection: Basic demographic information was collected, including age, gender, ethnicity and country of origin. Information about subject's betel nut usage, such as frequency, duration, preferred side of chewing in the mouth and product type and

brand, was also collected. A data collection sheet (questionnaire) was provided to the study subject with a list of these variables, but no direct identifiers were recorded.

5. Sampling: Among the 4 betel nut chewers and 4 non-betel nut chewers (age- and gender-matched controls), oral swab and saliva samples were taken. All subjects received a non-invasive clinical oral exam, which only used a mirror without probing to detect whether the subject had any betel nut-related oral diseases, such as oral submucous fibrosis, oral leukoplakia and oral cancer, and assess the general health status of the oral cavity, such as the presence or absence of periodontal disease, dental caries and other oral mucosal diseases. Then, a swab sample was obtained from each side of the oral cavity by using a cotton swab to gently touch the tongue, cheek and tooth surfaces. Only a sample was obtained from each subject and only one time. These samples were stored in test tubes with the Tris-EDTA (TE) buffer (pH 7.5) at -80°C until analyzed for pyrosequencing to obtain data for individual oral microbiome.

B. DNA EXTRACTION AND SEQUENCING

1. DNA extraction. The Qiagen DNeasy Blood & Tissue Kit was used to extract DNA from oral microbiological samples. The special protocol for isolating genomic DNA from Gram-positive bacteria provided by the manufacturer was followed. Briefly, oral swab samples were spun in a microcentrifuge tube by centrifuging for 10 min at 5000 x g (7500 rpm) to harvest microbial cells and the supernatant was discarded. The bacterial pellet was resuspend in 180 µl enzymatic lysis buffer which contained 20 mM Tris·Cl, pH 8.0, 2 mM sodium EDTA, 1.2% Triton® X-100, and 20 mg/ml lysozyme, and incubated for at least 30 min at 37°C. After incubation, the sample was heated in a water bath at 56°C and 25 µl proteinase K and 200 µl Buffer AL (without ethanol) were added and mixed by vortexing. The sample was incubated at

56°C for 30 min and 200 µl ethanol (100%) was added to the sample, and mixed thoroughly by vortexing. All of the sample was applied to the DNeasy Mini spin column for DNA isolation.

C. 16S rRNA GENE (V3 – V4 REGION) LIBRARY CONSTRUCTION

This protocol was to generate 250-base paired-end reads on the MiSeq Illumina for subsequent assembly of sequences spanning the V3-V4 variable regions. Primers composed of a 3' region representing the bacterial primer 341F (Muyzer et al. 1993) and the universal primer 806R (Caporaso et al. 2011, PNAS 108:4516–4522). The primers also possess priming sites for Illumina sequencing and indexes on the reverse primer read for combining multiple samples into one sequencing run. In addition, the forward primer contains a high-diversity region downstream of the Illumina forward sequencing primer, allowing for improved cluster identification by the Illumina data analysis pipeline. Paired-end sequencing has the advantage of providing quality control during the assembly process. The library construction involves two steps: PCR and gel purification.

1. PCR reaction. The Taq DNA polymerase was used because this enzyme generates more product than the previously used Phusion polymerase. This is a modification from the Bartram et al. 2011 protocol, as is the cycle number. The primer set used was 341F/806R.

PCR set-up (per reaction):

10 Standard Taq buffer	2.5 µl
100 µM V3_mod_Fw primer	0.05 µl
10 µM V4_Rev primer	0.5 µl
100 mM dNTPs	0.05 µl

Taq DNA polymerase (0.6 units/µl)	0.125 µl
Template (10 ng)	1.0 µl
Nuclease-free water	to 25µl

Three PCR amplifications for each sample was prepared with 1-10 ng of template DNA per reaction. Multiple PCR amplifications are pooled to reduce bias.

PCR conditions: 95˚C/30 sec, 30x (95˚C/15 sec, 50˚C/30 sec, 68˚C/30 sec), 68˚C/5 min. PCR products were carefully quantified on an agarose gel.

2. Purification of target PCR DNA fragments. First, three PCR products from the same sample were pooled in one tube, and then, the PCR product for each sample (unique index) and 2 µl of 10x loading dye were loaded in wells of a 1% agarose gel. Samples were loaded at least one lane away from the ladder on the same gel to eliminate cross contamination of libraries.

3. Excise DNA. Using a Dark Reader (less damaging to DNA than a UV transilluminator), the band corresponding to the V3-V4 region was excised and placed in a pre-weighed 1.5-ml tube.

4. Purification. The excised DNA from gel slice was purified using a Qiagen gel extraction kit and quantified.

5. Quality control. The DNA quality was checked on final library mixture (containing indexed DNA from all samples):

a. Run on gel to verify correct size, purity and concentration.

b. Use NanoDrop spectrophotometer to confirm concentration and purity.

The concentration of DNA was approximately 10 ng/µl. The samples were diluted before being added to the flow cell.

D. NEXT GENERATION SEQUENCING

The emergence of capillary sequencing technology and the use of dye-based detection methods, accompanied by automated analysis, has moved traditional DNA Sequencing into 'Next Generation' DNA Sequencing. Next Generation Sequencing (NGS) is a very high-throughput technique for generating millions of sequences at one time in order to analyze organisms at a genomic level. This allows the researcher to sequence, resequence and compare data at a rate previously not possible.

Next Generation Sequencing of the oral microbiological samples was done at the UIC Research Resources Center with the Illumina NGS machine. The NGS technology has allowed amplicon lengths of 400-500 bases to be sequenced in their entirety on multiple sequencing platforms. Shorter amplicon sequences are desirable as short reads are more likely to be sequenced completely. On occasion, amplicons are accidentally formed from a combination of various DNA strands due to sequencing error. These erroneous amplicons, or chimeras, are more likely to form during the replication of longer amplicon sequences, rather than shorter ones. Naturally, smaller sequences will have a smaller data footprint. In our study, we assessed the PCR amplicon from 16S rRNA genes, which has a mean length of 450 base pairs.

Error correction is a routine part of amplicon sequencing, as there are many points at which error may be introduced. Background contamination may be introduced, PCR processing error and chimera formation may occur, and the sequencing instrumentation

itself may result in error (Quince 2011). To prevent PCR processing error, proofreading enzymes were utilized. Additionally, running fewer cycles of PCR would decrease number of negative control amplification reactions, and decrease rate of chimera formation (Polz 1998, Qiu 2001, Liu 2014). There is no minimum number of samples necessary for any single study; however, a healthy sample would contain 20,000-50,000 reads with the assumption that 20% would be chimeric and removed.

Sequences were clustered into Operational Taxonomic Units (OTUs) for analysis. OTUs generally demonstrate a minimum 97% similarity rate, meaning that the genomic sequences differ by at most 3% (Stackebrandt 1994). OTUs, rather than species, were used to cluster different organic material, as OTU clustering was based purely on genetic sequence of the specific amplicon sequenced. It is important to note that there might be clustering of sequences from organisms of substantially diverse genomes.

After quality control and clustering steps were completed, library sizes of each sample were normalized such that appropriate comparisons could be made and statistical tests performed. Three common techniques for rarifying data sets are: rarifying the data sets such that each has the same number of sequences, converting OTUs to a proportion, and converting libraries to a common scale by statistically estimating relative depths of the sample libraries and standardizing OTU variance using a form of mixture model (Gihring 2012, McMurdie 2014).

E. STATISTICAL ANALYSES

For statistical analyses of microbiomes, the total species diversity in an environment (γ-diversity) is determined by the mean species diversity in local habitats (α-

diversity) and the differentiation in species compositions among these habitats (β-diversity) (Whittaker, 1960). Therefore, as illustrated in Fig. 2, the total species diversity is the product of α-diversity and β-diversity:

$$\gamma = \alpha \times \beta$$

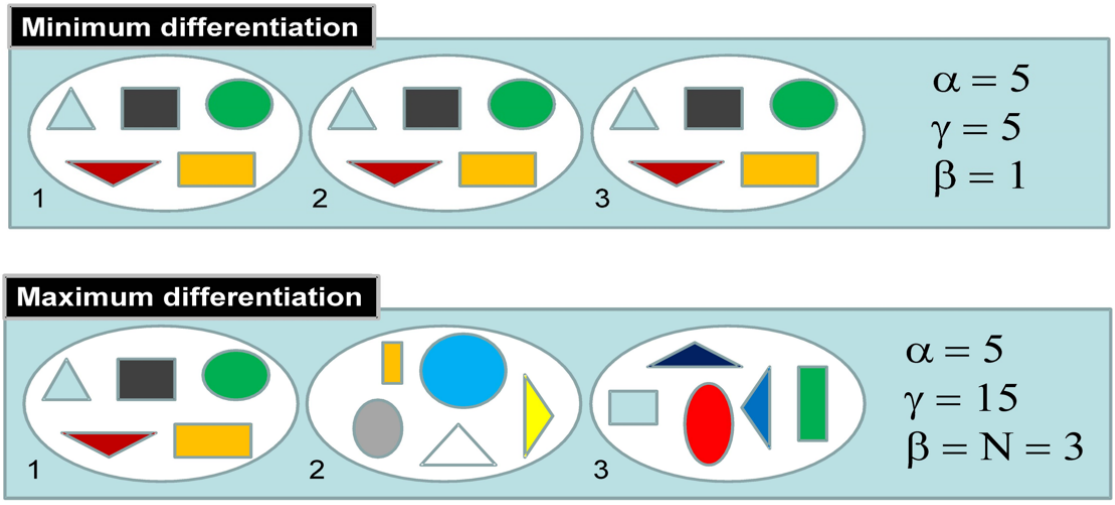

Fig 2. Illustration of the relationship between α, β, and γ diversities.

1. α-diversity

α-diversity is the number of taxonomic units (e.g., species richness) in each sample. To perform α-diversity analysis, we first measured the oral microbial composition of each subject. Based on OTUs collected from cheek swabs, one cannot conclusively determine the total number of different microbial species living in the entire oral cavity due to sample limitation. Nonetheless, information gleaned from a simple swab can clearly illustrate the α-diversity, the overall number of microbial species or level of biologic diversity, within the swab sample.

One statistical measurement of species richness is the Margalef Index (Margalef, 1958). The Margalef diversity index (d) can easily be calculated in a spreadsheet:

$$d = \frac{S - 1}{\ln N}$$

where S is the number of species, and N is the total number of individuals in the sample. Besides species richness, further testing is required to shed light on the proportional dominance of each species. An even distribution of 150 unique species, for instance, could have substantially different implications than a widely uneven spread of the same group of species. Thus, the fewer types of life, the higher the percentage of total lifeforms which is represented by each group.

A method of measuring relative abundance of species consists of mapping the squared proportion of the total that each species represents. This metric is sometimes referred to as the Herfindahl Index in the study of markets or the Simpson diversity index in the study of ecology. This cross-disciplinary method results in a score on a scale from 0 to 1. It is often re-scaled to a 0 to 10,000 scale by using percentages rather than decimals. For example, an index of .25 is the same as 2,500 points. 1 represents total domination by one species, while 0 represents a large number of species in the ecosystem, where the population percentage of each species rounds to 0. Mathematically, this is represented as:

$$H = \sum_{i=1}^{N} s_i^2$$

Where N is the number of species in the ecosystem, and S_i is the population percentage of the *i*th species. Thus, in an ecosystem with two species that each has 50 percent population, the Herfindahl index equals $0.50^2+0.50^2 = ½$.

As is the case with other tests of α-diversity, this study looked only at the 43 most common OTUs to calculate Herfindahl indices, for reasons of statistical relevance.

Another test, the Shannon Diversity Index, takes into account both the proportional number of species in a sample and the distribution of population among these species. It takes the summation of the multiple of each species' proportion of the population and the natural log of said proportion:

$$H' = -\sum_{i=1}^{R} p_i \ln p_i$$

where *pi* is the proportion of individuals belonging to the *i*th species in the dataset of interest.

2. β-diversity

β-diversity is the differentiation in microbiomes between different samples. β-diversity is a crucial component of microbiome research. α-diversity can capture whether test groups have different diversity than control groups; β-diversity dives one step deeper, and better reflects whether the increase or decrease in diversity holds true across multiple species within the samples. To clarify to the lay person: α-diversity scores can tell researchers that treated groups have more diverse microbiomes. β-diversity can tell researchers whether treated groups contain similar microbiomes, or simply similarly diverse microbiomes.

β-diversity can be assessed by first clustering sequence data into an OTU table, then generating sample-by-taxon matrices from OTU data, and lastly building a resemblance matrix by using pair-wise calculations through use of coefficients such as Bray-Curtis dissimilarity (Kuczynski 2010, Buttigieg 2014, Clarke 2006). Bray-Curtis dissimilarity is a comparison of the composition of two different samples, where 0 indicates that the samples are identical and 1 implies that the samples have no species in

common.

A number of statistical analyses can be performed to measure β-diversity. Analysis of Similarity, or ANOSIM, is a commonly used measure of similarity between samples. It is a modified matrix measure of closeness. Results from ANOSIM give a test statistic (R) and associated significance value (*p*-value). A test statistic close to 1 indicates strong correlation, and a *p*-value <0.05 indicates statistical significance at a 95% confidence level. In other words, the samples are ordered by closeness, and compared to each possible permutation of ordering. The number of more extreme orderings is divided by the total number of possible alternate combinations, yielding the appropriate *p*-value. With eight samples, there are 35 alternate combinations.

SIMPER, or Similarity of Percentages, is an analysis used to identify which taxa within samples cause most of the similarity or divergence between the control and treatment groups. SIMPER uses the Bray-Curtis measure of dissimilarity, ultimately comparing each sample in the test group with each sample in the control group, allowing for analysis between two groups of multiple samples each.

To best visualize the closeness of different biomes, non-metric multidimensional scaling (non-metric MDS) plots are often used. Such plots essentially visualize data extracted by applying different taxon population figures to a matrix, and calculating Bray-Curtis dissimilarity scores. The end result is simpler than the method: the smaller the distance between two points on the plot, the more closely related their microbiomes are.

An overlay of vectors can be applied, with each taxon present in the samples represented by a unique vector. Typically representative of Pearson or Spearman's

correlation, such lines indicate the presence of the genomic sequence in question, with the arrows pointing toward greater abundance. Points closer to the lines have stronger correlations.

Another statistical test, the Kruskal-Wallace Test, can be performed across multiple taxonomic levels, to indicate whether the rank of prevalence of particular taxonomic units (phyla, species) differs significantly between test and control samples. It relies on fewer assumptions, but comes to weaker conclusions, than a *student's t*-test.

3. Student's *t*-test

A *t*-test was used to determine if the makeup of oral bacterial species or OTUs was significantly different between the betel nut chewers and non-chewers.

III. RESULTS AND DISCUSSION

A. BETEL NUT CHEWERS SHOW INCREASED SPECIES DIVERSITY

In the samples taken, the total number of taxa found in the oral cavities of betel quid chewers was substantially higher than the number found in control subjects (Fig. 3). In fact, the range of taxa for betel quid chewers was 163-202, while the range for control subjects was found to be 152-162. A simple *t*-test corroborates the significance of this difference at a 0.05 level, with a *p*-value of 0.045 (Table 1).

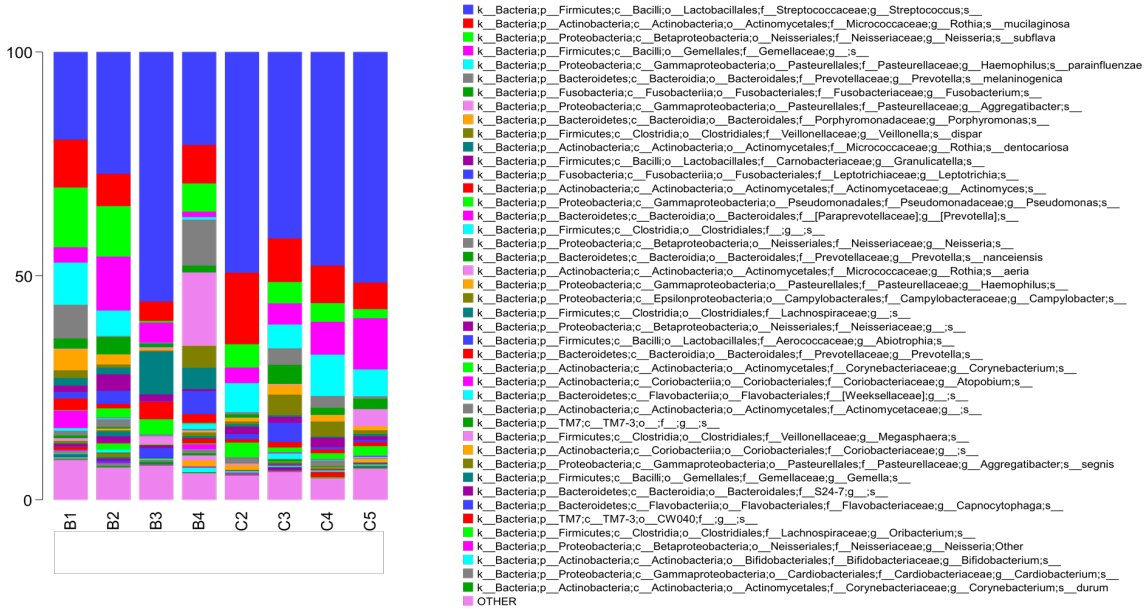

Fig. 3. Bar chart indicating species diversity within each sample.

Just as the total number of oral bacterial species (taxonomic units) found in the samples of betel quid chewers significantly exceeded the number of species found in control samples, the Margalef scores for betel quid samples were significantly higher than their control counterparts. Betel quid Margalef scores ranged from 14.72 to 18.27, while the control samples' Margalef scores ranged from 11.0 to 14.63. The *p*-value was again 0.045.

Herfindahl scores (Fig. 4) of each sample were also calculated, and range from 895.62 to 3279.69. A maximum score of 10,000 indicates complete monopoly of one species (100% biome representation, squared), while a minimum score of 0 indicates perfect equality among hundreds of different species (the sum of hundreds of 0% biome representation, squared). With one exception, B3 (the third betel quid-chewer's sample), the betel quid-affected samples all displayed higher levels of "evenness" than control samples. The *p*-value of a *t*-test of Herfindahl scores was 0.189, indicating that with a sample size so small, no conclusion could be definitively drawn at the 0.05 confidence level (Table 1).

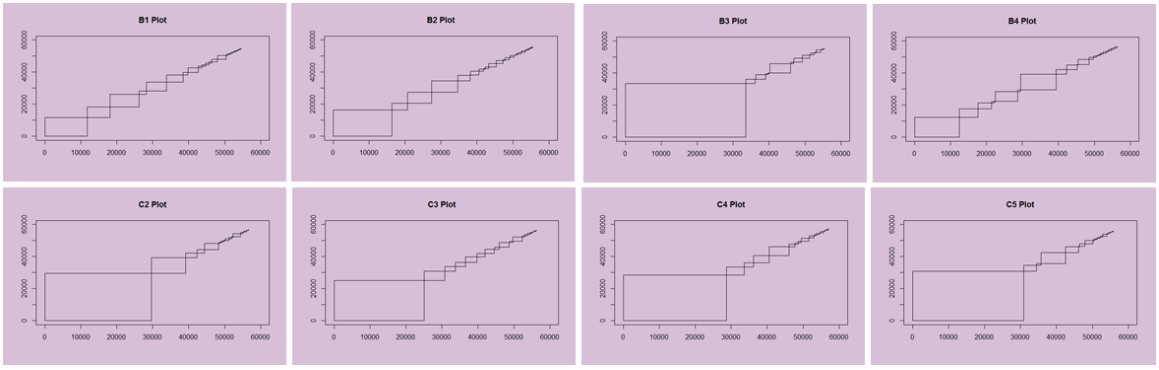

Fig 4. A visual depiction of Herfindahl scores, or sum of squared proportions of the samples. Plots B1-B4 (top row) were taken from betel quid chewers, and C2-C5 (bottom row) are control samples.

As was the case with Herfindahl scores, the Shannon diversity scores for all betel quid users were lower than the lowest control score, with the exception of B3. A *t*-test yielded a *p*-value of 0.097, indicating a correlation between population "evenness" and consumption of betel quid, but with such a small sample size the null cannot be rejected at the 0.05 level.

The initial results of α-diversity tests indicate that chewing betel quid appears to increase in species diversity in the oral microbiome.

B. BETEL NUT CHEWERS SHOW MICROBIOME SHIFT

Several tests for β-diversity were performed to determine differentiation among various microbiomes as opposed to the tests for species diversity within each microbiome as α-diversity measures.

The ANOSIM sample statistic (R) was 0.24, with a significance level of 8.6%, meaning that there were 2 more extreme combinations (3/35 = 0.086). Yet again, the sample size restricted us from determining with 95% certainty that there is a correlation between betel quid chewing and oral microbiome shift, although the results are strong enough to warrant further research with a greater sample size.

As can be viewed visually in Fig. 5, the control samples reflected oral microbiomes that were much more closely related than the samples taken from betel quid chewers. An overlay of correlation vectors is illustrated on the raw plot points, with major taxa clearly labeled (Fig. 6). Each vector points to an increased concentration of the genetic sequence in question. Similar results can be gleaned from Fig. 7, which, much

like an evolutionary tree, displays Bray-Curtis dissimilarity in relatedness between different biomes.

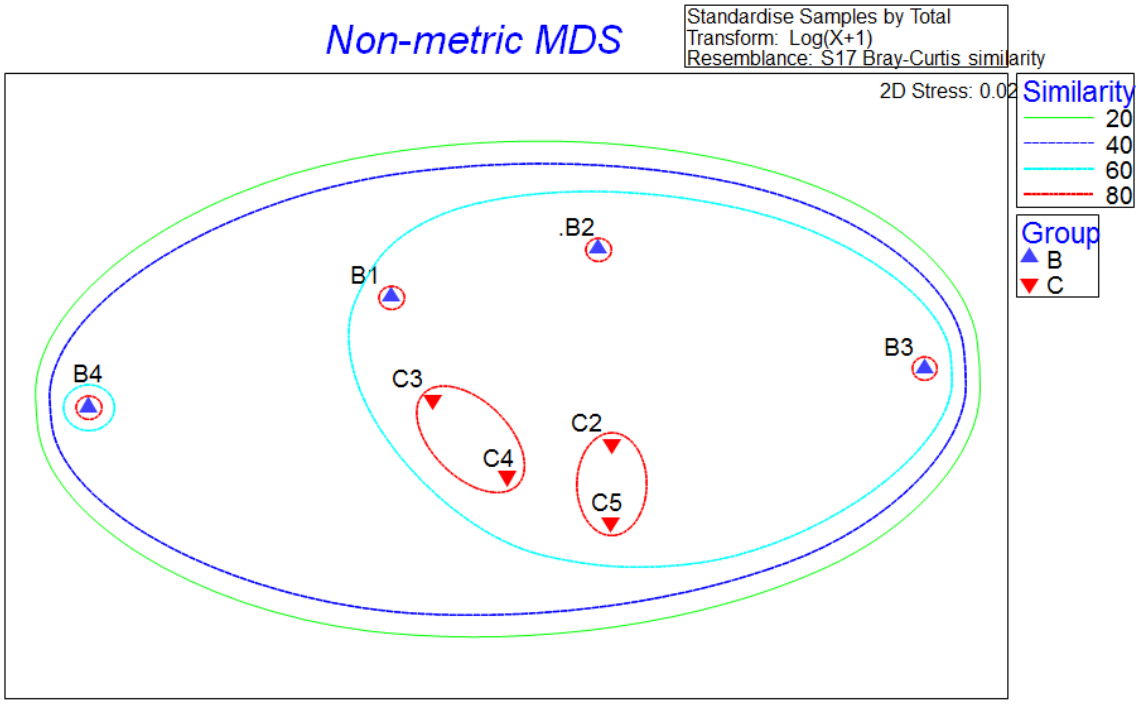

Fig. 5. Non-metric MDS displaying relative closeness of different oral microbiomes.

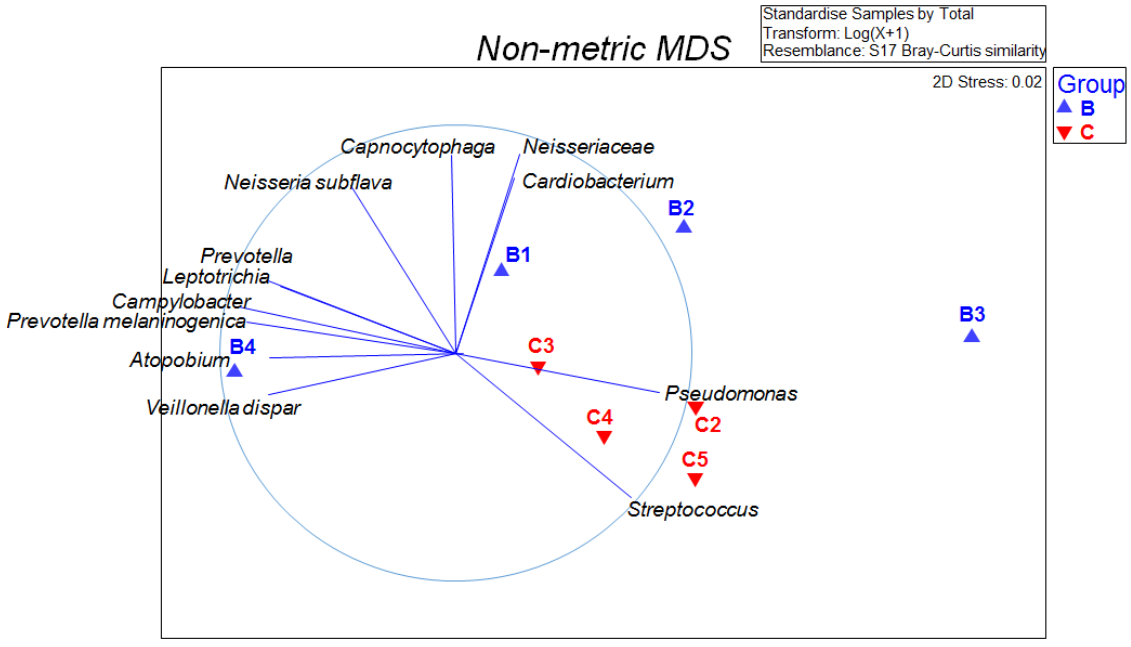

Fig 6. Non-metric MDS plot with an overlay of correlation vectors.

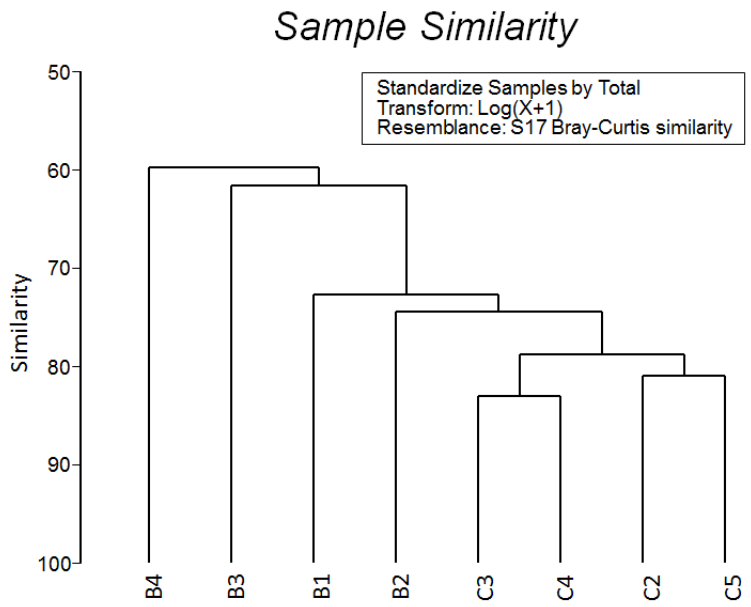

Fig. 7. Tree of Bray-Curtis sample dissimilarity, with more similar samples grouped closer together.

A SIMPER analysis further crystallizes the major points of differentiation between control samples and test samples. Among the control samples, which have an average similarity of 81.25, the majority of the similarity between samples can be attributed to *Streptocococcus* (55.33%), *Rothia mucilaginosa* (9.07%), and *Haemophilus parainfluenzae* (7.06%). Samples taken from betel quid consumers have an average similarity of 54.90%, which largely comes from the same species: *Streptococcus* (38.73%), *R. mucilaginosa* (10.96%), *Neisseria subflava* (7.57%), and the family *Gemellaceae* (4.57%).

The two groups, control and test samples, have an average dissimilarity of 37.13%, which again stems from many of the same taxa, which vary the most in prominence. *Streptococcus* accounts for 28.03% of the dissimilarity, followed by *N.*

subflava (7.58%), *Aggregatibacter* genus (6.16%), *Prevotella melaninogenica* (5.9%), and the family *Gemellaceae* (5.68%). The results are displayed in Table 2.

The Kruskal-Wallace tests performed measured the rank differences of different taxonomic levels between control and test samples. Nothing met the false-discovery rate, yet test results on a number of taxonomic levels revealed significant results at higher *p*-values. More statistically revealing t-tests were performed, indicating significant differences regarding prevalence of certain taxa.

C. BETEL NUT CHEWERS SHOW LESS DOMINANCE OF "GOOD BUGS"

Although many oral bacterial species were identified with the next-generation sequencing, eight genera/species showed apparent differences between the group of betel nut chewers and that of non-chewers. These are *Streptococcus, Neisseria subflava, Rothia dentocariosa, Acticomyces, Prevotella, Neisseriaceae2, Corynebacterium*, and *Bacteria TM7*. In the betel nut-chewing group, only the genus of *Streptococcus*, which is considered as a commensal bacterium or "healthy microbe," has lower dominance than it does in the control group, while the proportion of all other bacterial species are higher in the test group. The Student's *t*-test was used to analyze whether these differences are statistically significant. The results showed that the differences in three species are significant ($p < 0.05$). These are *Rothia dentocariosa, Acticomyces, Neisseriaceae2,* and *Bacteria TM7*. The results are shown in Fig. 8. The relative abundance (percent composition) of these six genera/species is shown in Fig. 9.

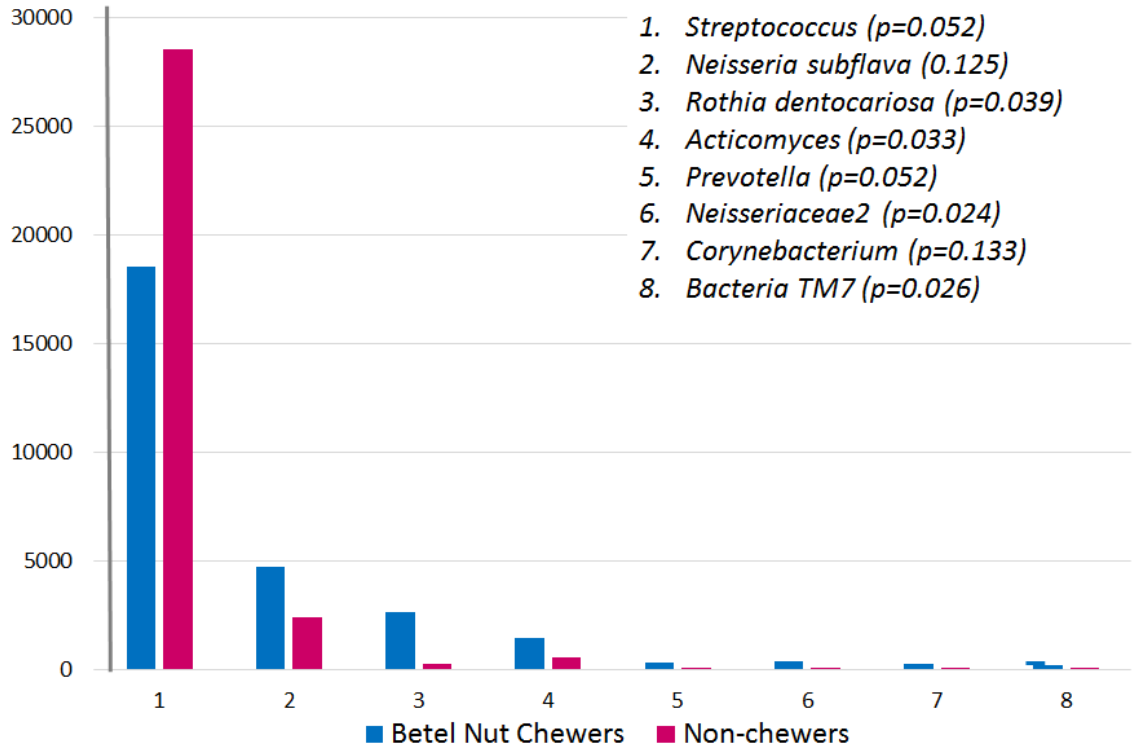

Fig. 8. Species dominance among chewers and non-chewers.

Fig. 9. Difference in relative abundance of bacterial species between betel nut chewers and non-chewers.

D. CLINICAL IMPLICATION

Initial results show that chewing betel quid correlates with increased diversity of the oral microbiome, and greater "evenness" within the microbiome. The microbiomes of betel quid chewers are also less similar to each other than those of controls. While consistently more diverse than their control counterparts, they vary widely in terms of which particular species are more prevalent. In the oral cavity of a healthy subject, the dominant microbial genus is *Streptococcus*, which is considered as a commensal

bacterium or "good bug" that competitively inhibits other potentially pathogenic microbial genera or "bad bugs" by producing antimicrobial substances, such as bacteriocins, hydrogen peroxide, and lactic acid (Jakubovics et al., 2014). The microbiome analysis result shows that the overall *Streptococcus* population is reduced in betel nut chewers. This could be caused by the toxicity associated with chemicals released from chewing the areca nut, betel leaves, and/or lime power (e.g., calcium hydroxide). Reduction of the "good bug," *Streptococcus,* may allow other bacterial species, including "bad bugs," normally inhibited by *Streptococcus* to appear and overgrow. This may result in a richer oral microbial species diversity in betel nut chewers. Some of the increased microbial species (e.g., *Prevotella*) found in betel nut chewers are associated with periodontal disease. This may explain that betel nut chewers often have increased incidences and more severe types of periodontal disease (Hsiao et al., 2014). Whether the betel nut chewing-associated oral microbiome shift contributes to other betel nut-promoted diseases, such as leukoplakia, erythroplakia, lichenoid, OSF and oral cancer, is currently unknown but deserves further investigation.

It should be noted that correlation does not necessarily indicate causation. Although samples were taken from South Asians, this study did not control for other factors which make an individual more predisposed to consume betel quid and have a more diverse oral microbiome. Predisposition to consume various food types, for instance, could therefore explain this study's findings.

Additionally, a major limitation of the study is the sample size. Although in the South Asian community many individuals chew betel nuts, few seek dental care at the UIC College of Dentistry. Nonetheless, we were able to recruit 4 betel quid chewers and

4 age, sex and ethnicity-matched control subjects to participate in the study. Despite the small sample size, their microbiome data yielded strong indications of correlation and statistically significant differences in some data sets, especially dominant species.

E. FURTHER RESEARCH

The results from this project warrant further research into the effect of betel quid chewing on oral microbiomes. Future research projects should include a larger sample size in order to better determine statistical significance at a 95% confidence level across multiple tests, for both α- and β-diversities, and also focus more on other potential correlative factors, such as region of origin and ethnic diet. Future research should also more carefully separate consumption of tobacco-infused betel quid and non-tobacco betel quid, to better isolate the effect of tobacco in conjunction with areca nut, betel leave and slaked lime.

If a more statistically rigorous link between consumption of betel quid and oral microbiome shift is determined, further investigation should be completed to ascertain the effect of such a shift. Does the oral microbiome shift associated with betel quid consumption contribute to increased risk of oral diseases including periodontal disease, leukoplakia, erythroplakis, lichenoid, OSF and cancer? Can other oral diseases be linked to betel quid consumption through this mechanism indirectly? Do other, similar oral microbiomes correlate with health and disease as well?

Table 1 - Measures of α-Diversity

Sample	S	N	d	J'	H'(loge)	1-Lambda'	SS
Tao1.B1	182	60000	16.45	0.5573	2.9	0.9081	895.626447
Tao2.B2	163	60000	14.72	0.5422	2.762	0.882	1158.49605
Tao3.B3	202	60000	18.27	0.3702	1.965	0.6698	3279.68476
Tao4.B4	167	60000	15.09	0.5405	2.766	0.8975	1010.49245
Tao6.C2	152	60000	13.72	0.3923	1.971	0.7193	2788.71743
Tao7.C3	162	60000	14.63	0.4599	2.34	0.7988	1986.78251
Tao8.C4	122	60000	11	0.436	2.095	0.7447	2537.46843
Tao9.C5	153	60000	13.82	0.3952	1.988	0.7077	2890.35017
p-val	0.04539495		0.04537831	0.16223061	0.09656543	0.18849473	0.18900745

S	Number of taxa (species)
N	Number of sequences
d	Species richness (Margalef)
J'	Pielou's evenness
H'(loge)	Shannon Index (log e)
1-Lambda'	Simpson's index
SS	Simple Sum of Squares
p-val	p-value from the t-test

TABLE I.

MEASURES OF α-DIVERSITY.

Table 2 - SIMPER Analyses

Table 2a1 - SIMPER Within Samples from Betel Nut Chewers
Average similarity: 54.90

Species	Av.Abund	Av.Sim	Sim/SD	Contrib%	Cum.%
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus;s__	30.84	21.26	7.13	38.73	38.73
OTHER	7.4	6.63	8.29	12.08	50.81
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Micrococcaceae;g__Rothia;s__mucilaginosa	7.71	6.01	3.17	10.96	61.76
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Neisseriales;f__Neisseriaceae;g__Neisseria;s__subflava	7.82	4.15	0.91	7.57	69.33
k__Bacteria;p__Firmicutes;c__Bacilli;o__Gemellales;f__Gemellaceae;g__;s__	5.31	2.51	1.71	4.57	73.9

Table 2a2 - SIMPER Within Control Samples
Average similarity: 81.25

Species	Av.Abund	Av.Sim	Sim/SD	Contrib%	Cum.%
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus;s__	47.56	44.96	12.31	55.33	55.33
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Micrococcaceae;g__Rothia;s__mucilaginosa	10	7.37	4.46	9.07	64.4
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Haemophilus;s__parainfluenzae	6.75	5.73	12.04	7.06	71.46

Table 2b - SIMPER Between Groups B & C
Average dissimilarity = 37.13

	Group B	Group C				
Species	Av.Abund	Av.Abund	Av.Diss	Diss/SD	Contrib%	Cum.%
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus;s__	30.84	47.56	10.41	2.33	28.03	28.03
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Neisseriales;f__Neisseriaceae;g__Neisseria;s__subflava	7.82	4.03	2.82	1.79	7.58	35.61
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Aggregatibacter;s__	4.22	1	2.29	0.69	6.16	41.78
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__melaninogenica	4.54	1.84	2.19	1.33	5.9	47.67
k__Bacteria;p__Firmicutes;c__Bacilli;o__Gemellales;f__Gemellaceae;g__;s__	5.31	6.76	2.11	1.29	5.68	53.35
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Haemophilus;s__parainfluenzae	3.91	6.75	2.1	1.49	5.67	59.02
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Micrococcaceae;g__Rothia;s__dentocariosa	4.41	0.43	1.99	1.19	5.35	64.37
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Micrococcaceae;g__Rothia;s__mucilaginosa	7.71	10	1.91	1.18	5.14	69.51
k__Bacteria;p__Fusobacteria;c__Fusobacteriia;o__Fusobacteriales;f__Leptotrichiaceae;g__Leptotrichia;s__	2.44	1.59	1.04	1.32	2.81	72.32

TABLE II.
SIMPER ANALYSES.

CITED LITERTURE

Bartram A, Lynch MDJ, Stearns JC, Moreno-Hagelsieb G, Neufeld JD. 2011. Generation of multi-million 16S rRNA gene libraries from complex microbial communities by assembling paired-end Illumina reads. Appl. Environ. Microbiol. 77:3846-3852.

Bedi R. Betel-quid and tobacco chewing among the United Kingdom's Bangladeshi community. Br. J. Cancer. 1996;74 Suppl. 29:S73–S77.

Bhonsle RB, Murti PR, Gupta PC. Tobacco habits in India. In: Gupta PC, Hamner JE, Murti PR, editors. Control of tobacco-related cancers and other diseases. Bombay: Oxford University Press; 1992, pp. 25± 47

BMC Oral Health. 2014 Nov 4;14:133.

Burton-Bradley B.G. Betel chewing in retrospect. Papua New Guinea med. J. 1978;21:236–241.

Buttigieg PL, Ramette A. 2014. A guide to statistical anal- ysis in microbial ecology: a community-focused, living review of multivariate data analyses. FEMS Microbiol Ecol 90:543–550.

Canniff, J. P., and W. Harvey. "The aetiology of oral submucous fibrosis: the stimulation of collagen synthesis by extracts of areca nut." *International journal of oral surgery* 10.Suppl 1 (1980): 163-167.

Cawte J. 1985 'Psychoactive substances of the South Seas: betel, kava and pituri', *Australian New Zealand Journal of Psychiatry*,Vol.19, pp.83-87.

Cheong Y.H. The fading links between tradition and oral health in Singapore. Int. dent. J. 1984;34:253–256.

Clarke KR, Somerfield PJ, Chapman MG. 2006. On resemblance measures for ecological studies, including taxonomic dissimilarities and a zero-adjusted Bray-Curtis coefficient for denuded assemblages. JExp Mar Biol Ecol 330:55–80.

Council of the European Communities, author. Directive 2001/37/EC of the European Parliament and of the Council of 5 June 2001 on the approximation of the laws, regulations and administrative provisions of the Member States concerning the manufacture, presentation and sale of tobacco products. Off. J. Eur. Communities. 2001;L194:26–34.

Croucher R., Islam S. Socio-economic aspects of areca nut use. Addict. Biol. 2002;7:139–146.

Daftary, D.K., Bhonsle, R.B., Murti, R.B., Pindborg, J.J., Mehta, F.S., 1980 'An oral lichen planus-like lesion in Indian betel-toabcco chewers', *Scandinavian Journal of Dental Research*, Vol.88, pp. 244-9.

Deng, Yi-Ting, et al. "Arecoline-stimulated connective tissue growth factor production in human buccal mucosal fibroblasts: Modulation by curcumin." *Oral oncology* 45.9 (2009): e99-e105.

Deshpande, Swati. "'Current Ban on Stronger Footing than in 2002' - The Times of India." The Times of India. Times of India, 14 July 2012. Web. 06 Apr. 2016.

Downer, M.C., Petti, S. (2005) 'Leukoplakia prevalence estimate lower than expected. What is the global prevalence of leukoplakia?' *Evidence-Based Dentistry*. Vol.6, pp.12–14.

Ebot V, Merlin M, Lindstrom L. Kava: the Pacific drug. New Haven: Yale University Press; 1992, pp. 25± 6, 54± 5, 141.

Eisen MJ. 1946 'Betel chewing among natives of the southwest Pacifi c islands. Lack of carcinogenic action', *Cancer Research*, Vol.6, pp. 39 141.

Forastiere, A., Koch, W., Trotti, A., Sidransky, D. 2001 'Head and k cancer', *New England Journal of Medicine*, Vol.345, pp.1890-1900

Gihring TM, Green SJ, Schadt CW. 2012. Massively par- allel rRNA gene sequencing exacerbates the potential for biased community diversity comparisons due to variable library sizes. Environ Microbiol 14:285–290.

Gorman, Chester F. "Excavations at spirit cave, North Thailand: some interim interpretations." *Asian Perspectives* 13 (1970): 79-107.

Gupta, Bhawna, and Newell W. Johnson. "Systematic review and meta-analysis of association of smokeless tobacco and of betel quid without tobacco with incidence of oral cancer in South Asia and the Pacific." *PloS one* 9.11 (2014): e113385.

Gupta PC. 1984 'A study of dose-response relationship between tobacco habits and oral leukoplakia' *British Journal of Cancer*, Vol.50, pp.527-531.

Gupta, P.C., Ray, C.S. 2004 'Epidemiology of betel quid usage', *Annals Academy of Medicine Singapore*, Vol.33 (suppl), No.4, pp. S31- S36

Harvey, W., et al. "Stimulation of human buccal mucosa fibroblasts in vitro by betel-nut alkaloids." *Archives of oral biology* 31.1 (1986): 45-49.

Hashibe, Mia, et al. "Chewing tobacco, alcohol, and the risk of erythroplakia." *Cancer*

Epidemiology Biomarkers & Prevention 9.7 (2000): 639-645.

Health Canada (1995). Herbs Used as Non-medicinal Ingredients in Nonprescription Drugs for Human Use.

Hirsch E. Efficacy and concentration: analogies in betel use among the Fuyuge (Papua New Guinea). In: Goodman J, Lovejoy PE, Sherratt A, editors. Consuming habits: drugs in history and anthro- pology. London: Routledge; 1995, pp. 88± 102.

Hoffman, Neil W., Jean-Pierre Wuarin, and F. Edward Dudek. "Whole-cell recordings of spontaneous synaptic currents in medial preoptic neurons from rat hypothalamic slices: mediation by amino acid neurotransmitters." *Brain research* 660.2 (1994): 349-352.

Hogan J. Tobacco products require redefining. Br. dent. J. 2000;188:61.

Hsiao CN, Ting CC, Shieh TY, Ko EC. Relationship between betel quid chewing and radiographic alveolar bone loss among Taiwanese aboriginals: a retrospective study.

IARC Monograph, Betel-quid and Areca-nut Chewing and Some Areca-nut-derived Nitrosamines; 2004:Volume 85

IARC Working Group on the Evaluation of Carcinogenic Risk to Humans. Betel-quid and Areca-nut Chewing and Some Areca-nut-derived Nitrosamines. Lyon (FR): International Agency for Research on Cancer; 2004. (IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, No. 85.) Betel-quid and Areca-nut Chewing. Available from: http://www.ncbi.nlm.nih.gov/books/NBK316574/

Jakubovics NS, Yassin SA, Rickard AH. Community interactions of oral streptococci. Adv Appl Microbiol. 2014;87:43-110.

Jamrozik, K. 1985 'Regional variation of oral cancer in Papua New Guinea'. *Papua New Guinea Medical Journal*, Vol. 28, pp. 9-13

Jones, D. (2001), *Palms Throughout The World*, Reed New Holland, Australia
Li WC Molecular and cellular cues of diet-associated oral carcinogenesis—with an emphasis on areca-nut-induced oral cancer development, Journal of Oral Pathology and Medicine, in press. 2014.

Kagen hM, Trackenman PC. Properties and function of lysyl oxidase. Am J Respir Cell Mol Biol 1991; 5:206-10

Kuczynski J, Liu Z, Lozupone C, Mcdonald D, Fierer N, Knight R. 2010. Microbial community resemblance meth- ods differ in their ability to detect biologically relevant patterns. Nat Meth 7:813–819.

Lee, C.H., Ko, Y.C., Huang, H.L., Chao, Y.Y., Tsai,C.C., Sheih, T.Y., Lin, L.M.2003 'The precancer risk of betel quid chewing, tobacco use and alcohol consumption in oral leukoplakia and oral submucous fibrosis in southern Taiwan', *British Journal of Cancer*, Vol.88, pp.366-372.

Liu J, Song H, Liu D, Zuo T, Lu F, Zhuang H, Gao F. 2014. Extensive recombination due to heteroduplexes gen- erates large amounts of artificial gene fragments during PCR. PLoS One 9:e106658.

Mack TM (2001). The new pan-Asian paan problem. Lancet 357: 1638–1639

MacLennan R, Paissat D, Ring A, Thomas S.,1985 'Possible aetiology of oral cancer in Papua New Guinea,' *Papua New Guinea Medical Journal*, Vol. 28, pp.3-8.

Matts, Robert L., et al. "Toxic heavy metal ions activate the heme-regulated eukaryotic initiation factor-2 alpha kinase by inhibiting the capacity of hemin-supplemented reticulocyte lysates to reduce disulfide bonds." *Journal of Biological Chemistry* 266.19 (1991): 12695-12702.

McMurdie PJ, Holmes S. 2014. Waste not, want not: why rarefying microbiome data is inadmissible. PLoS Comput Biol 10:e1003531.

McPherson K. The Indian Ocean. A history of people and the sea. Oxford: Oxford University Press; 1993, pp. 82± 91.

Meghji, S., et al. "INHIBITION OF COLLAGENASE ACTIVITY BY ARECA NUT TANNINS-A MECHANISM OF COLLAGEN ACCUMULATION IN ORAL SUBMUCOUS FIBROSIS (OSF)." *J Dent Res* 61.4 (1982): 545-545.

Mehta, F.S., Gupta, P.C., Daftary, D.K., Pindborg, J.J., Choksi, S.K., 1972 'An epidemiologic study of oral cancer and precancerous conditions among 101, 761 villagers in Maharashtra, India', *International Journal of Cancer*, Vol.10, pps. 134-141

Meulenbeld, Gerrit Jan. "Reflections on the basic concepts of Indian pharmacology." *Studies on* (1987).

Murti, P.R., Bhonsole R.B, Pindborg J.J, Daftary, D.K., Gupta PC, Mehta, F.S., 1985 'Malignant transformation rate in oral submucous fi brosis over a 17-year period', *Community Dentistry and Oral Epidemiology*, Vol. 13, pp. 340-341

National Centre for Transcultural Oral Health (2001) Educational Programme for Trading Standards and Customs and Excise Officers, London, Resolutions, p. 17.

Nieschultz V, Schmersahl P Zur Phamakologie der Wirkstoffe des Betels Umwenlung

des arecoline in arecadin. Arzneimittel-Forsch 1968; 18:222-5.

Norton, S.A. 1998 'Betel: Consumption and consequences' *Journal of the American Academy of Dermatology*, Vol. 38, No. 1 pp 81- 87

Orr, I.M. 1933 'Oral cancer in betel nut chewers in Travancore; its aetiology, pathology and treatment', *Lancet*; 2:575 580.

Pace NR, Stahl DA, Lane DJ, Olsen GJ. 1986. The anal- ysis of natural microbial populations by ribosomal RNA sequences, Adv Microb Ecol 9:1–55.

Parkin, D. Max, Paola Pisani, and J. Ferlay. "Global cancer statistics." *CA: a cancer journal for clinicians* 49.1 (1999): 33-64.

Pereira, Andresa Costa, et al. "Matrix metalloproteinase gene polymorphisms and oral cancer." *Journal of clinical and experimental dentistry* 4.5 (2012): 297-301.

Pindborg, Jens J., and Satyavati M. Sirsat. "Oral submucous fibrosis." *Oral Surgery, Oral Medicine, Oral Pathology* 22.6 (1966): 764-779.

Pindborg, J. J., et al. "Incidence and early forms of oral submucous fibrosis." *Oral Surgery, Oral Medicine, Oral Pathology* 50.1 (1980): 40-44.

Pindborg JJ, Zheng Kui-Hua, Kong Ching-Ren, Lin Fa-Xing. Pilot survey of oral mucosa in areca (betel) nut chewers on Hainan Island of the People's Republic of China. Community Dent Oral Epidemiol 1984;12:195± 6.

Polz MF, Cavanaugh CM. 1998. Bias in template-to-prod- uct ratios in multitemplate PCR. Appl Environ Microbiol 64: 3724–3730.

Qiu X, Wu L, Huang H, McDonel PE, Palumbo AV, Tiedje JM, Zhou J. 2001. Evaluation of PCR-generated chimeras, mutations, and heteroduplexes with 16S rRNA gene-based cloning. Appl Environ Microbiol 67:880–887.

Quince C, Lanzen A, Davenport RJ, Turnbaugh PJ. 2011. Removing noise from pyrosequenced amplicons. BMC Bioinform 12:38.

Rajalalitha, P., and S. Vali. "Molecular pathogenesis of oral submucous fibrosis–a collagen metabolic disorder." *Journal of oral pathology & medicine* 34.6 (2005): 321-328.

Rajendran, R., 1994 'Oral submucous fi brosis: etiology, pathogenesis, and future research' Bulletin of the World Health Organization, Vol. 72 no.6 pp. 831-1016

Reichart, Peter A., Ulrich Creutz, and Christian Scheifele. "The 'Skull from Bangkok': a skull of a betel quid chewer in the anthropological collection of Rudolf Virchow

(Berlin)." *Journal of oral pathology & medicine* 35.7 (2006): 410-412.

Reid A. From betel-chewing to tobacco-smoking in Indonesia. J Asian Studies 1985; XLIV: 529± 47.

Reysenbach A-L, Wickham GS, Pace NR. 1994. Phyloge- netic analysis of the hyperthermophilic pink filament com- munity in Octopus Spring, Yellowstone National Park. Appl Environ Microbiol 60:2113–2119.

Rooban, T., et al. "Health hazards of chewing arecanut and products containing arecanut." *Calicut Medical Journal* 3.2 (2005).

Schmidt, Brian L., et al. "Tobacco smoking history and presentation of oral squamous cell carcinoma." *Journal of oral and maxillofacial surgery* 62.9 (2004): 1055-1058.

Schmidt TM, DeLong E, Pace N. 1991. Analysis of a marine picoplankton community by 16S rRNA gene clon- ing and sequencing. J Bacteriol 173:4371–4378.

Sharma, B.B. (2002) Prevention of Food Adulteration Act, 1954 (Notification No. PFA/574-2002/7, dated 23 July), Mumbai, Commissioner for Food and Drug Administration and Food (Health) Authority, Maharashtra State.

Shetty K.V., Johnson N.W. Knowledge, attitudes and beliefs of adult South Asians living in London regarding risk factors and signs for oral cancer. Community Dent. Health. 1999;16:227–231.

Shiboski, Caroline H., Brian L. Schmidt, and Richard CK Jordan. "Tongue and tonsil carcinoma." *Cancer* 103.9 (2005): 1843-1849.

Sinor, P.N., Gupta, P.c., Murti, P.R., Bhonsle, R.B., Daftary, D.K., Mehta, F.S., Pindborg, J.J., 1990 'A case controlled study of oral submucous fi brosis with special reference to the etiologic role of areca nut', *Journal of Oral Pathology and Medicine*, Vol.19, No.2, pp.94-98

Stackebrandt E, Goebel B. 1994. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. Int J Syst Bacteriol 44:846–849.

Stanfill SB, Connolly GN, Zhang L, Jia LT, Henningfield JE, et al. (2011) Global surveillance of oral tobacco products: total nicotine, unionised nicotine and tobacco-specific N-nitrosamines. Tob Control 20:2.

Sudhakar, A. "Oral Submucous Fibrosis: Review on Etiopathogenesis." *J Canc Sci Ther* 1 (2009): 1-4.

Thomas, S.J., MacLennan, R. 1992 ‘Slaked lime and betel nut cancer in Papua New Guinea’, *Lancet*, Vol.340, pp 577

Tilakaratne, W. M., et al. "Oral submucous fibrosis: review on aetiology and pathogenesis." *Oral oncology* 42.6 (2006): 561-568.

Trivedy C.R. The legislative issues related to the sale and consumption of areca (betel) nut products in the UK: Current status and objectives for the future. J. Indian Med. Assoc. 2001;1:10–16.

Trivedy, C. R., G. Craig, and S. Warnakulasuriya. "The oral health consequences of chewing areca nut." *Addiction biology* 7.1 (2002): 115-125.

Trivedy, Chetan, et al. "Copper content in Areca catechu (betel nut) products and oral submucous fibrosis." *Lancet* 349.9063 (1997): 1447.

Trivedy, C., et al. "The upregulation of lysyl oxidase in oral submucous fibrosis and squamous cell carcinoma." *Journal of oral pathology & medicine* 28.6 (1999): 246-251.

Trivedy, C. R., et al. "Copper stimulates human oral fibroblasts in vitro: a role in the pathogenesis of oral submucous fibrosis." *Journal of oral pathology & medicine* 30.8 (2001): 465-470.

United Arab Emirates. Travel and Transportation. Dubai. Dubai.ae. N.p., 08 July 2014. Web. 3 Apr. 2016.

van der Waal, Isaäc, and T. Axell. "Oral leukoplakia: a proposal for uniform reporting." *Oral oncology* 38.6 (2002): 521-526.

Vora A.R., Yeoman C.M., Hayter J.P. Alcohol, tobacco and paan use and understanding of oral cancer risk among Asian males in Leicester. Br. dent. J. 2000;188:444–451.

Wade WG. The oral microbiome in health and disease. Pharmacol Res. 2013. 69:137-43.

Wal, Aradhna. "Strong Support for Gutka Ban in 8 States, Finds Study." DNAIndia.com. DNA India, 17 Oct. 2015. Web. 5 Apr. 2016.

Warnakulasuriya, K. A. A. S., et al. "Localised oral histoplasmosis lesions associated with HIV infection." *Journal of oral pathology & medicine* 26.6 (1997): 294-296.

Warnakulasuriya, Saman. "Histological grading of oral epithelial dysplasia: revisited." *The Journal of pathology* 194.3 (2001): 294-297.

Warnakulasuriya S., Trivedy C., Peters T.J. Areca nut use: An independent risk factor for

oral cancer. Br. med. J. 2002;324:799–800.

Whittaker, R. H. (1960) Vegetation of the Siskiyou Mountains, Oregon and California. Ecological Monographs, 30, 279–338.

Woese CR, Gutell R, Gupta R, Noller H. 1983. Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. Microbiol Rev 47:621.

World Health Organization 2009-2010.

World Health Organization (2012). Review of areca (betel) nut and tobacco use in the Pacific: a technical report. Regional Office for the Western Pacific.

Zhang SS, Li WH, Gao YJ, Liu ZW, Liu L, Tang JQ, Ling TY. Betel-quid and oral submucous fibrosis: a cross-sectional study in Hunan province, China. J Oral Pathol Med. 2012; 41:748-54.