Appl Microbiol Biotechnol (2017) 101:5391–5403
DOI 10.1007/s00253-017-8282-9

APPLIED MICROBIAL AND CELL PHYSIOLOGY

Smokeless tobacco products harbor diverse bacterial microbiota that differ across products and brands

Eoghan M. Smyth[1,2] · **Prachi Kulkarni**[2] · **Emma Claye**[2] · **Stephen Stanfill**[3] · **Robert Tyx**[3] · **Cynthia Maddox**[1] · **Emmanuel F. Mongodin**[1] · **Amy R. Sapkota**[2]

Received: 9 December 2016 / Revised: 10 March 2017 / Accepted: 31 March 2017 / Published online: 22 April 2017

Abstract Smokeless tobacco products contain numerous chemical compounds, including known human carcinogens. Other smokeless tobacco constituents, including bacteria, may also contribute to adverse health effects among smokeless tobacco users. However, there is a lack of data regarding the microbial constituents of smokeless tobacco. Our goal was to characterize the bacterial microbiota of different smokeless tobacco products and evaluate differences across product types and brands. DNA was extracted from 15 brands of smokeless tobacco products (including dry snuff, moist snuff, snus, and Swedish snus) and 6 handmade products (e.g., toombak) using an enzymatic and mechanical lysis approach. Bacterial community profiling was performed using PCR amplification of the V1–V2 hypervariable region of the 16S rRNA gene, followed by 454 pyrosequencing of the resulting amplicons and sequence analysis using the QIIME package. Total viable counts were also determined to estimate the number of viable bacteria present in each product. Average total viable counts ranged from 0 to 9.35×10^7 CFU g^{-1}. Analysis of the 16S rRNA gene sequences revealed high bacterial diversity across the majority of products tested: dry snuff products where characterized by the highest diversity indices compared to other products. The most dominant bacterial phyla across all products were *Firmicutes*, *Proteobacteria*, *Actinobacteria*, and *Bacteroidetes*. Significant differences in both bacterial community composition and in silico predicted gene content were observed between smokeless tobacco product types and between brands of specific smokeless tobacco products. These data are useful in order to comprehensively address potential health risks associated with the use of smokeless tobacco products.

Keywords Smokeless tobacco · Snus · Snuff · Microbiota · Bacteria · 16S rRNA

Introduction

Smokeless tobacco products are used either orally or nasally, and include products such as dry snuff, moist snuff, snus, and toombak. Snuff is a general term for finely cut or powdered, flavored tobacco, which can be prepared as dry snuff (fire-cured, fermented tobacco powder that may contain aroma and flavor additives) or as moist snuff (air-cured and fire-cured tobacco, flavored and powdered into fine particles, containing 20–55% moisture by weight) (Boffetta et al. 2008). Snus is the Swedish variation of moist snuff and consists of ground tobacco mixed with water, salt, sodium bicarbonate, sodium chloride, humectants, and flavoring; the processing of this product includes heat treatment (pasteurization) to eliminate microorganisms (Juarez and Merlo 2013). Toombak, also called Shammah, is a type of snuff found in parts of North and East Africa and the Middle-East (Idris et al. 1998; Zhang et al. 2001). Toombak is a mixture of fermented ground powder from tobacco leaves and is mixed with an aqueous solution

Electronic supplementary material The online version of this article (doi:10.1007/s00253-017-8282-9) contains supplementary material, which is available to authorized users.

✉ Amy R. Sapkota
ars@umd.edu

[1] Institute for Genome Sciences, University of Maryland School of Medicine, 801 West Baltimore St., Baltimore, MD 21201, USA

[2] Maryland Institute for Applied Environmental Health, University of Maryland School of Public Health, 4200 Valley Drive, Bldg no. 255, Room 2234P, College Park, MD 20742, USA

[3] Division of Laboratory Sciences, Centers for Disease Control and Prevention, 1600 Clifton Road, Atlanta, GA 30329, USA

of sodium bicarbonate (i.e., an alkalinizing agent) (Idris et al. 1998).

Smokeless tobacco products are widely used in North America, Sweden, Africa, and throughout the world (Ayo-Yusuf and Burns 2012). In the USA, 3.6% of adults 18 years and older use some form of smokeless tobacco, and rates of usage among males have been slowly increasing since 2000 (USDHHS 2014). Use of smokeless tobacco products varies widely depending on sex, age, ethnic origin, and socio-economic status (Boffetta et al. 2008). In terms of toxicity, smokeless tobacco products are commonly thought to be less toxic than smoked products despite the fact that they are associated with many adverse health effects including nicotine addiction, oral lesions, oral and pancreatic cancer, and cardiovascular disease (Stepanov et al. 2008).

Many of these adverse health outcomes are attributed to chemical carcinogens present in smokeless tobacco products including tobacco-specific nitrosamines (TSNAs), nitrosamino acids, and polycyclic aromatic hydrocarbons (Stepanov et al. 2008). TSNAs are thought to be among the most important tobacco-associated carcinogens due to their toxicity and high abundance in smokeless tobacco products (Stepanov et al. 2012). Different TSNAs are formed from the reaction of alkaloids with nitrite (Wei et al. 2014), and levels of available nitrite are influenced by nitrite-reducing bacteria that are known to be part of the consortium of bacteria, or microbiota, associated with tobacco products (Fisher et al. 2012). However, little research has been carried out on the bacterial microbiota of tobacco products and their expressed activities.

Therefore, the aim of this study was to further characterize the bacterial microbiota of multiple smokeless tobacco products using a culture-independent, next-generation sequencing approach. In addition, we used traditional culture-based methods to provide evidence that bacteria detected in smokeless tobacco products can be viable, thereby highlighting the potentially critical role that these tobacco-associated communities may play in the metabolism of chemical carcinogens, such as TSNAs, and the resulting human health impacts among smokeless tobacco users.

Materials and methods

Samples

A total of 21 smokeless tobacco products were characterized in this study: 15 commercially available product brands and 5 handmade products that included 5 toombak samples and 1 alkalinizing agent used to make the toombak. For each commercially available product brand, we included three lots and tested four samples per lot ($n = 12$ samples per product brand). Product brands, types, and manufacturers (when known) are summarized in Table 1. All samples were stored at room temperature prior to DNA extraction. All handling, sampling, and experimental procedures involving the smokeless tobacco products were carried out in a biosafety level II cabinet that had been sterilized using both UV-light for 20 min and 70% ethanol. In addition, all laboratory tools that were used during experiments were sterilized by autoclaving and flame sterilizing immediately prior to use.

Total viable counts

For each commercially available smokeless tobacco product, four cans per product were pooled into one sterile stomacher bag. The toombak products that were delivered in aggregate form were crushed into a powder and transferred into a sterile 50-ml Falcon tube (Fisher Scientific, Pittsburg, PA, USA). Then, 1 g of pooled smokeless tobacco (per product) was weighed and added to 9 ml of phosphate-buffered saline (PBS) in a 50-ml Falcon tube and vortexed. Serial dilutions (1:10; to a 10^{-5} dilution) were made and 100 μl of each dilution was plated onto a tryptic soy agar (TSA) plate amended with 250 μg ml^{-1} actidione to suppress fungal growth. Four to five glass beads were added to the plate followed by rotation in a circular motion until all of the liquid was evenly spread over the plate surface and the plate surface was dry. This was repeated for all dilutions, and the plates were incubated at 37 °C for 24 ± 2 h. After incubation, colonies were counted and colony forming units (CFU) per gram of tobacco product were calculated. This was completed in triplicate for each product.

DNA extraction

Each smokeless tobacco product (0.2 g) was transferred under sterile conditions to a Lysing Matrix B tube (MP Biomedicals, Solon, OH), and enzymatic lysis was initiated by adding to each tube: 1 ml ice cold 1× PBS buffer (molecular grade; Gibco, Life Technologies, NY), 5 μl of 10 mg ml^{-1} lysozyme (Sigma-Aldrich, MO), 5 μl of 5 mg ml^{-1} lysostaphin (Sigma-Aldrich, MO), and 15 μl of 1 mg ml^{-1} mutanolysin (Sigma-Aldrich, MO). Tubes were then incubated at 37 °C for 30 min, after which 10 μl of 20 mg ml^{-1} Proteinase K (Invitrogen by Life Technologies, NY) and 50 μl of 10% *w/v* SDS (BioRad) were added to each tube, followed by incubation at 55 °C for 45 min. Samples were then mechanically lysed using a FastPrep Instrument FP-24 (MP Biomedicals, CA) at 6.0 m s^{-1} for 40 s. The resulting lysate was then briefly centrifuged, and DNA was purified using the QIAmp DSP DNA Mini Kit 50, v2 (Qiagen, CA), according to the manufacturer's protocol. In order to remove any potential PCR inhibitors, the extracted DNA was further purified by precipitation using 3 M sodium acetate (one tenth volume) and 100% ethanol (2.5 volumes), followed by overnight incubation at

Table 1 Smokeless tobacco products tested in this study

Product brand	Abbreviation	Product type	Manufacturer
Camel Snus Frost	CSF	Snus	RJ Reynolds Tobacco Company
Marlboro Original Snus	MOS	Snus	Philip Morris USA, Inc
Navy Sweet Scotch	NSS	Dry snuff	Swisher International, Inc
Tops Dry Sweet Snuff	TDSS	Dry snuff	Swisher International, Inc
General Original	GOSS	Swedish snus	Swedish Match
Ettan Loose	ELS	Swedish snus	Swedish Match
Hawken Wintergreen Chewing Tobacco	HWCT	Moist snuff	American Snuff Company, LLC
Kodiak Premium Wintergreen	KPW	Moist snuff	American Snuff Company, LLC
Copenhagen Snuff Original Fine Cut	CSOFC	Moist snuff	U.S. Smokeless Tobacco Company
Skoal Bandits Mint Pouches	SBMP	Moist snuff	U.S. Smokeless Tobacco Company
Timber Wolf Wintergreen Fine Cut	TWWFC	Moist snuff	Swedish Match North America, Inc.
Kodiak Premium Straight	KPS	Moist snuff	American Snuff Company, LLC
Silver Creek Wintergreen Fine Cut	SCWFC	Moist snuff	Swisher International, Inc
Longhorn Long Cut Wintergreen	LLCW	Moist snuff	Swedish Match North America, Inc.
Renegades Wintergreen Pouches	RWP	Moist snuff	Swedish Match North America, Inc.
Alkalinizing agent	A agent	Alkalinizing agent	Handmade
Brown Toombak	BT	Toombak	Handmade
Grey Toombak	GT	Toombak	Handmade
Toombak tube	TT	Toombak	Handmade
Toombak leaves	TL	Toombak	Handmade

−20 °C. The DNA was then pelleted, washed with 80% ethanol, and resuspended in Tris–EDTA buffer.

16S rRNA gene PCR amplification and sequencing

The universal primers 27F and 338R were used for PCR amplification of the V1–V2 hypervariable regions of 16S ribosomal RNA (rRNA) gene. The 338R primer included a unique sequence tag to barcode each sample, as published previously (Zupancic et al. 2012). Using 96 barcoded 338R primers, the V1–V2 regions of the 16S rRNA gene were amplified in 96-well microtiter plates using Phusion High-Fidelity PCR MasterMix with HF buffer (NE Biolabs, Ipswich, MA), with 0.5 μM of each primer, an additional 0.375 μl of BSA (20 mg ml^{-1}), and 50 ng of template DNA in a total reaction volume of 25 μl, using the cycling conditions described previously (Zupancic et al. 2012). Negative controls without a template were included for each barcoded primer pair. PCR products were quantified using the Quant-iT PicoGreen dsDNA assay, and equimolar amounts (100 ng) of the 16S PCR amplicons were mixed in a single tube. The purified amplicon mixture was then sequenced by 454 FLX Titanium pyrosequencing using 454 Life Sciences primer A by the Genomics Resource Center at the Institute for Genome Sciences, University of Maryland School of Medicine, using protocols recommended by the manufacturer as amended by the Center.

Analysis of the 16S rRNA gene sequences

Raw sequences were processed and analyzed using the Quantitative Insights Into Microbial Ecology (QIIME) package (QIIME release v. 1.9; Caporaso et al. 2010b). Briefly, raw sequencing reads were demultiplexed using 5′barcodes, trimmed of forward and reverse primer sequences, filtered for length and quality, and corrected for homopolymer errors (Zupancic et al. 2012). Sequences were then clustered into operational taxonomic units (OTUs) using USEARCH (Edgar 2010) with a 97% identity threshold. Representative sequences of each cluster were aligned using PyNAST (Caporaso et al. 2010a) and used to construct a neighbor-joining phylogenetic tree with RAxML (Stamatakis 2006). Taxonomic assignment was performed using the RDP classifier (Wang et al. 2007) trained by the Silva16S database (Quast et al. 2013) release 111 with a minimum confidence threshold of 0.97. The resulting OTU table was then analyzed

using tools implemented within the R statistical software package (R Core Team 2016). Select OTUs were also identified by checking their representative sequences against the NCBI nucleotide database with BLASTn (Altschul et al. 1990).

In silico prediction of smokeless tobacco functional metagenomic profiles

The in silico prediction of functional metagenomic profiles was performed using the PICRUSt software package (Langille et al. 2013). First, OTUs were picked using a closed reference OTU picking strategy using QIIME, against the GreenGenes v.13_5 (McDonald et al. 2012) reference database at 97% similarity. The resulting Biological Observation Matrix (BIOM)-formatted OTU file was then used as input in the online Galaxy version of the PICRUSt software (http://huttenhower.sph.harvard.edu/galaxy/). OTUs were initially normalized by dividing their abundances by known or predicted 16S rRNA gene copy number abundances, before metagenomic functional predictions were created using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. The resulting PICRUSt-generated BIOM output file was then analyzed within the R framework, as described below.

Visualization and statistical analysis

Visualization and analysis were performed with the R statistical software v. 3.2 (R Developmental Core Team 2010) using the R packages ggplot2 (Wickham 2009) and Phyloseq v. 1.12.2 (McMurdie and Holmes 2013). Alpha- and beta-diversity metrics were computed using the Phyloseq package. Significant differences between product brand, type, and manufacturer were tested using analysis of variance (ANOVA), with Tukey–Kramer post-hoc test and Bonferroni corrections. These tests were performed using Statistical Analysis of Metagenomic Profiles (STAMP) software (Parks et al. 2014). Smokeless tobacco products represented by less than one sample after removal due to low sequencing counts were not included in the ANOVA tests (i.e., alkalinizing agent, toombak, toombak leaves and toombak tube). In all cases, p values of ≤ 0.05 were defined as statistically significant.

Availability of supporting data

Data concerning the samples included in this study have been deposited in the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/) under BioProject accession number PRJNA311747. All sequence files were linked to BioProject PRJNA311747 and deposited to Genbank's Short Read Archive (SRA).

Results

Total viable counts

Cultivation-independent techniques using phylogenetic microarrays and DNA sequencing-based strategies have dramatically increased our knowledge of the microbial diversity present in various environmental and human niches. However, a major limitation of these DNA-based strategies is that they cannot assess viability of microbial cells in complex communities. In order to circumvent this limitation, we performed cultivation experiments on tryptic soy agar plates for each smokeless tobacco product to provide evidence that the tested products contain viable bacteria. For all smokeless tobacco products (Table 1), the average total viable counts ranged from 0 to 9.35×10^7 CFU g^{-1} (Fig. 1). All of the snus and Swedish snus products, as well as Hawken Wintergreen Chewing Tobacco, were characterized by no bacterial viable counts. Toombak leaves had the highest average total viable counts (9.35×10^7 CFU g^{-1}). Dry snuff products had higher average total viable counts compared to moist snuff products. Average total viable counts for Navy Sweet Scotch Dry Snuff and Tops Dry Sweet Snuff were 1.45×10^7 and 6.63×10^4 CFU g^{-1}, respectively. Moist snuff products had average total viable counts ranging from 4.27×10^4 to 2.73×10^6 CFU g^{-1}. The alkalinizing agent had the lowest total viable count (6.5×10^3 CFU g^{-1}) among products for which bacteria could be cultured (Fig. 1).

16S rRNA gene sequencing dataset

Overall, bacterial community profiling using 16S rRNA gene sequencing was performed on 202 smokeless tobacco samples (Table S1). A total of 1,372,703 16S rRNA gene sequences were obtained, representing 2635 unique OTUs at a 97% similarity cutoff across all samples. On average, 6795 sequences were obtained per sample, which could be clustered to 102 OTUs (min 11–max 844) on average per sample. After OTU clustering and taxonomic assignments, OTUs assigned to the phylum *Cyanobacteria* were removed from further downstream analysis, as these likely represent sequences amplified from tobacco chloroplast DNA. Forty-seven samples had less than 300 sequencing reads and were removed from further downstream analyses (Table S1). This sequence cutoff was chosen based on the Good's coverage index calculated for each sample (Fig. S1). On average, the Good's coverage for samples with >300 sequencing reads was 99.25% (min 93.6%; max 100%), indicating that only an additional 7.5 phylotypes would be expected on average for every 1000 additional sequences in these samples. This sequence coverage level also indicates that, independently of the number of sequencing reads obtained for each sample, the 16S rRNA gene sequences identified in these samples represent the vast

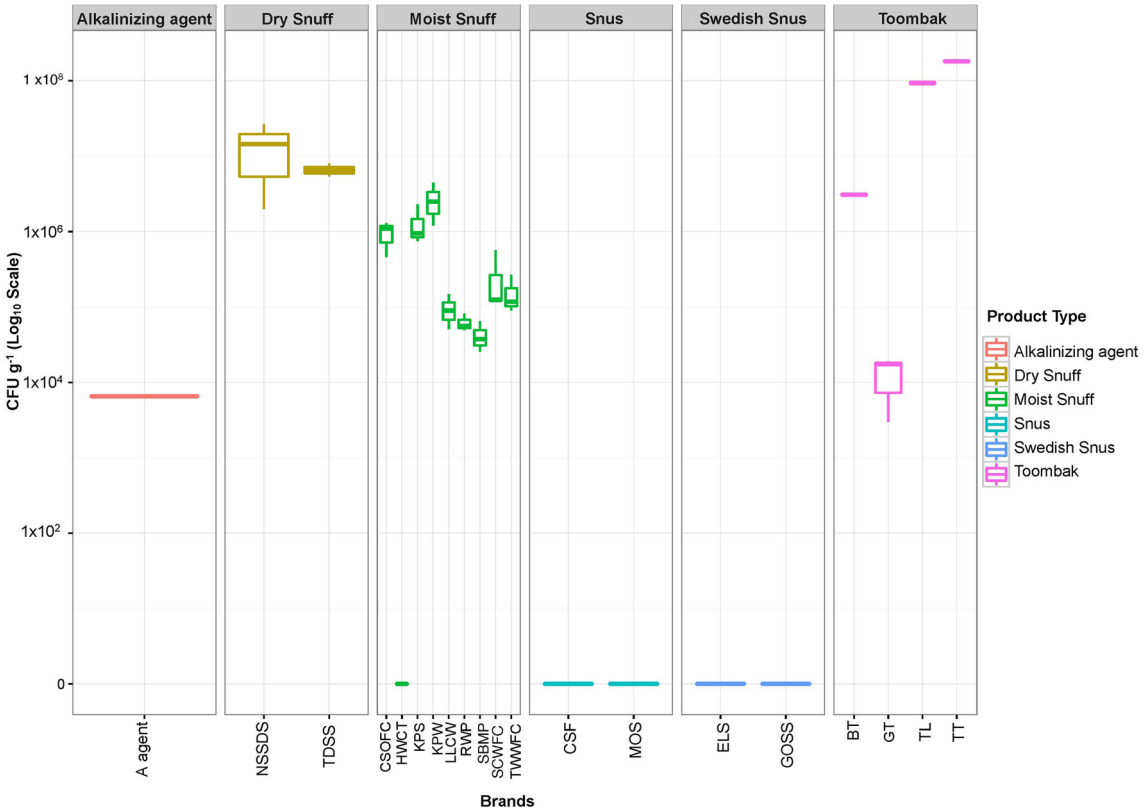

Fig. 1 Total viable counts (TVCs) of smokeless tobacco products

majority of the bacterial taxa present in each smokeless tobacco product characterized in our dataset. After using this cutoff, 155 samples remained: 13 out of 24 Snus samples; 18 out of 24 Swedish Snus samples; 7 out of 12 Hawken Wintergreen Chewing Tobacco moist snuff samples; and 9 out of 16 Toombak samples had less than 300 sequences and were removed from further analysis (Table S1).

Besides some of the Toombak samples (e.g., G3-2 and B1-1 that had low sequence reads but high CFU counts), the samples with very low sequence yields corresponded with total viable counts of zero (Fig. 1; Table S1).

Bacterial 16S rRNA diversity between smokeless tobacco products

The Phyloseq package in R was used to calculate alpha diversity as a measure of the biodiversity within each sample. This biodiversity is defined in microbial ecology as an attribute that has two components: richness (total number of species) and evenness (proportions—or relative abundance—of species within each sample). In this study, we reported alpha-diversity using the Observed (total number of OTUs), Chao1, and Shannon diversity indices (Fig. 2). Overall, dry snuff products had higher diversity indices than moist snuff, Swedish snus, and snus products. The alkalinizing agent had the highest diversity indices, followed by dry snuff products, moist snuff products (except for Hawken Wintergreen Chewing Tobacco), and toombak products. Snus and Swedish snus had the lowest diversity indices (Fig. 2). No correlation was observed between the diversity measures and the total viable counts.

Beta diversity (differences in composition between samples) was measured using Bray–Curtis distance and plotted using nonmetric multidimensional scaling (NMDS). Product type, brand (Fig. 3a), and manufacturer (Fig. 3b) were the variables that explained the most variance in bacterial microbiota between samples.

Characterization of the bacterial microbiota of smokeless tobacco products

The most dominant bacterial phyla across all smokeless tobacco products were *Firmicutes* (0.0–100%; median = 95.2% across all samples), *Proteobacteria* (0.0–98.9%; median = 0.1% across all samples), *Actinobacteria* (0.0–80.1%; median = 2.5% across all samples), and *Bacteroidetes* (0.0–11.1%; median = 0.0% across all samples). There were a large variability between product brands, types, and manufacturers and some variabilities between different lots of the same brand (Figs. 3 and 4, Fig. S2). On average, relative abundance of *Firmicutes* was higher in moist snuff products, whereas levels of *Proteobacteria* were higher in Swedish snus, snus, and dry

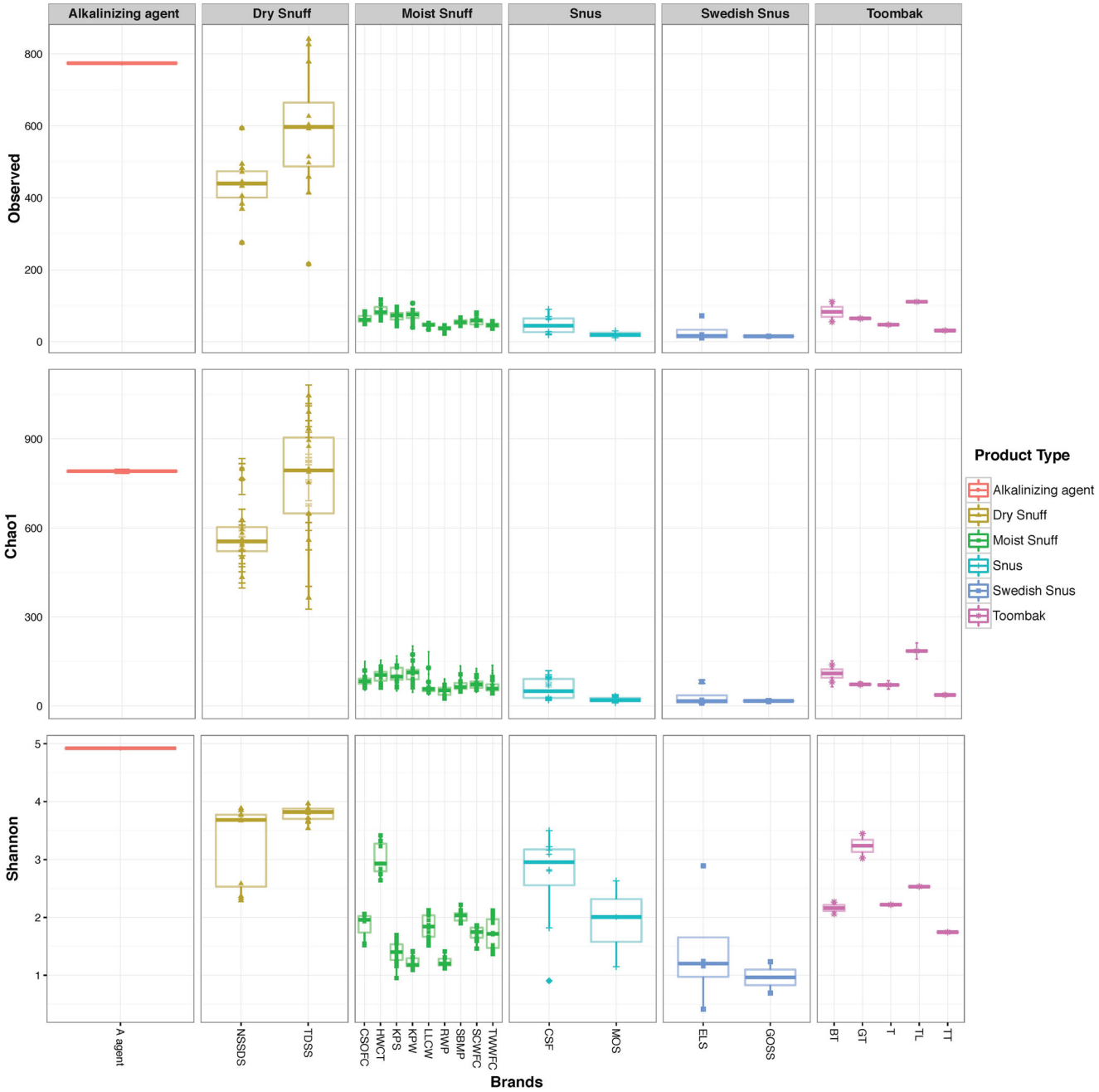

Fig. 2 Bacterial diversity of smokeless tobacco products as measured by the *Observed*, *Chao1*, and *Shannon* diversity indices

snuff products and *Actinobacteria* was highest in toombak products. There were significant differences between product brand ($p < 0.001$; Tukey–Kramer), product type ($p < 0.001$; Tukey–Kramer) and product manufacturer ($p < 0.001$; Tukey–Kramer) for all these phyla (Fig. 4). The relative abundance of *Firmicutes* ranged from 35.2 to 63.2% in dry snuff products, 56.4 to 99.9% in moist snuff products, 0.7 to 21.0% in Swedish snus products, 36.8 to 52.1% in snus products, 21.2 to 56.6% in toombak products, and 4.6% in the alkalinizing agent (Fig. 4).

Proteobacteria relative abundance ranged from 12.8 to 44.9% in dry snuff products, 0.0 to 15.8% in moist snuff products, 73.3 to 98.4% in Swedish snus products, 33.6 to 47.9% in the snus product, 0.0 to 49.6% in toombak products, and 8.0% in the alkalinizing agent (Fig. 4).

The relative abundance of *Actinobacteria* ranged from 11.6 to 13.7% in dry snuff products, 0.0 to 19.9% in moist snuff products, 0.0 to 3.7% in Swedish snus products, 9.5 to 13.7% in the snus product, 11.2 to 78.8% in toombak products, and 39.5% in the alkalinizing agent (Fig. 4).

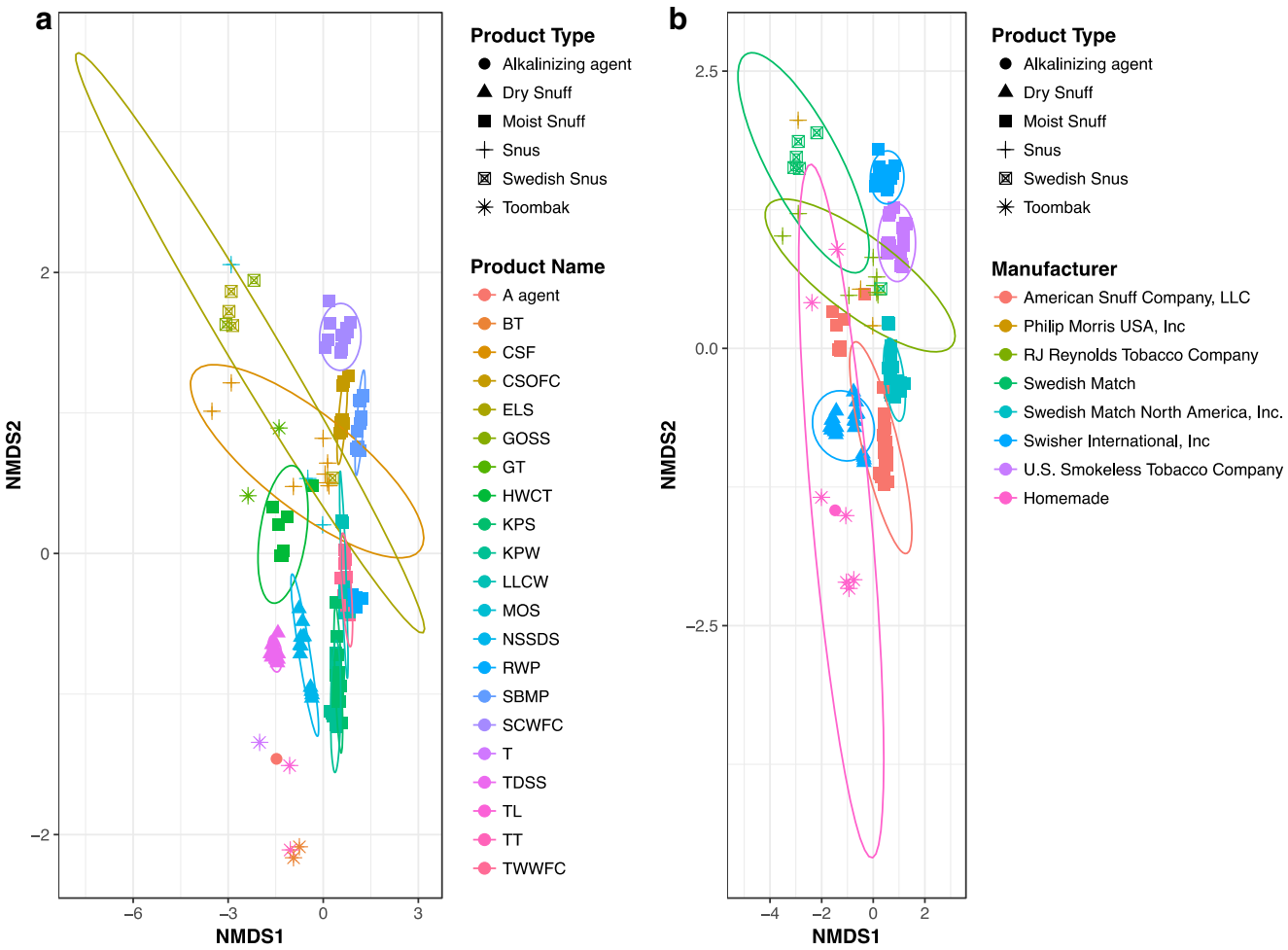

Fig. 3 Ordination plot derived from nonmetric multidimensional scaling (NMDS) using Bray–Curtis distance of bacterial community composition. **a** *Colors* represent product brands, *shapes* represent product types, and *ellipses* are drawn at 95% confidence intervals for product brand. **b** *Colors* represent product manufacturers, *shapes* represent product types, and *ellipses* are drawn at 95% confidence intervals for product manufacturer

Bacteroidetes relative abundance ranged from 0.8 to 2.0% in dry snuff products, 0.0 to 2.0% in moist snuff products, 0.4 to 1.2% in Swedish snus products, 0.6 to 3.9% in the snus product, 0.0 to 5.1% in toombak products, and 2.7% in the alkalinizing agent (Fig. 4).

In the top 20 OTUs for each product, only one OTU (OTU 0) was common for all product types, and this was identified as *Staphylococcus cohnii* (Table S2). Of the 2628 OTUs, 721 were significantly different between product brands ($p < 0.05$), 578 were significantly different between product types ($p < 0.05$; Tukey–Kramer), and 317 were significantly different between product manufacturers ($p < 0.05$; Tukey–Kramer).

From the top 100 OTUs, OTUs taxonomically assigned within the *Acetobacter* (OTUs 6, 16, 22, 1515), *Kurthia* (OTU 44), and *Lactobacillus* (OTUs 7, 10, 31, 48) genera were significantly more abundant in dry snuff compared to all other product types ($p < 0.001$; Tukey–Kramer) (Fig. 5; Table S2). OTUs from the genera *Ralstonia* (OTU 15) and *Burkholderia* (OTU 21) were significantly higher in Swedish snus compared to all other product types ($p < 0.001$) (Fig. 5; Table S2). OTUs from the genera *Enteractinococcus* (OTU 11) and *Corynebacterium* (OTU 18) were significantly higher in toombak products compared to all other product types ($p < 0.001$; Tukey–Kramer) (Table S2).

Predicted bacterial community functions

PICRUSt was used to compare the (in silico) predicted functional potential of the bacterial microbiota associated with different smokeless tobacco products using the 16S rRNA gene sequence data. In total, 6909 genes were predicted by PICRUSt. Of these, 5115 were significantly different between smokeless tobacco brands ($p < 0.05$; Tukey–Kramer), 4232 were significantly different between smokeless tobacco types

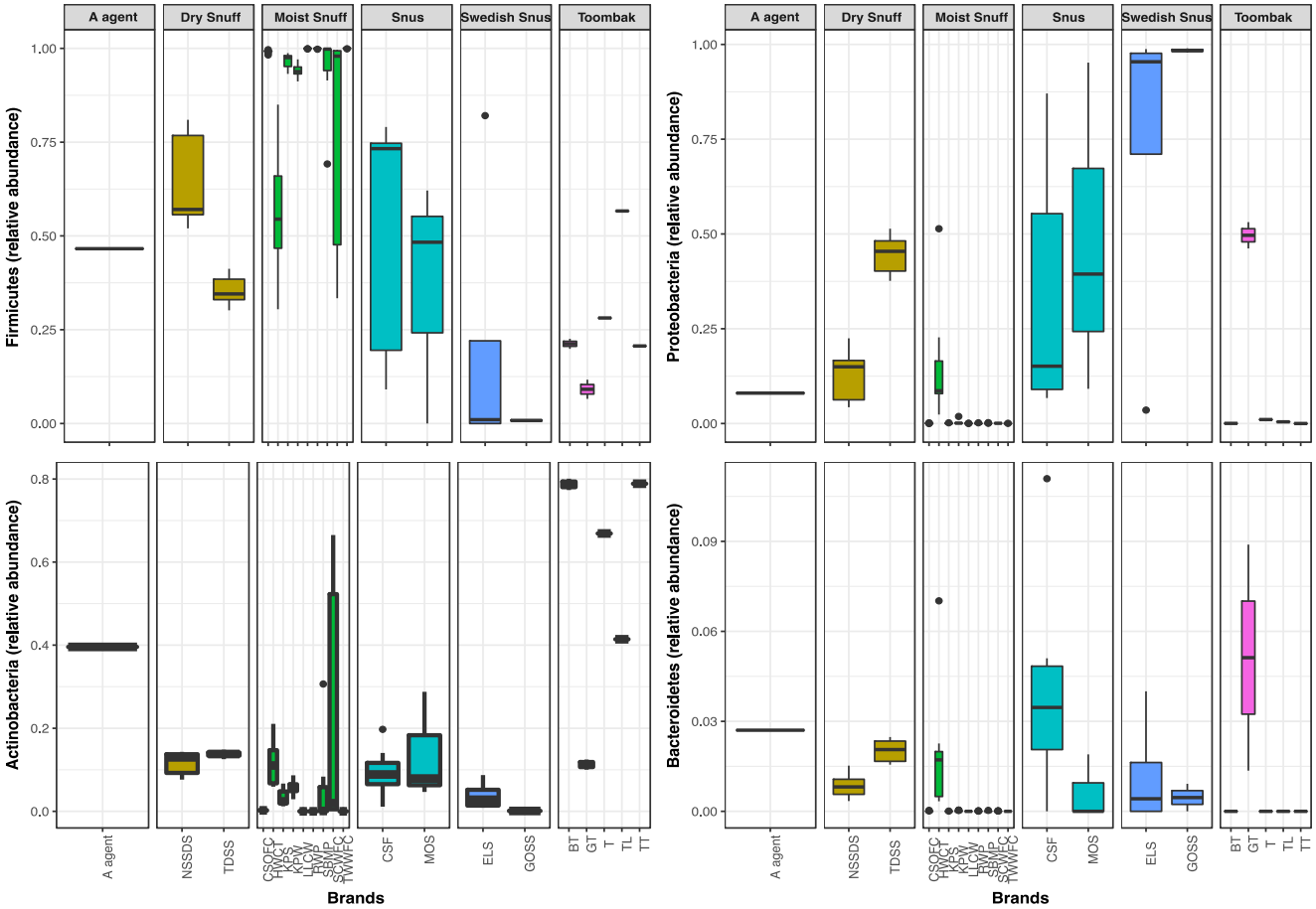

Fig. 4 Boxplots of relative abundance of the following phylum across smokeless tobacco brands: *Actinobacteria*, *Bacteroidetes*, *Firmicutes*, and *Proteobacteria*. *Colors of bars* designate the product type of each brand

($p < 0.05$; Tukey–Kramer), and 4468 were significantly different between smokeless tobacco manufacturers ($p < 0.05$; Tukey–Kramer).

The 6909 predicted genes were assigned to 328 KEGG level 3 modules (a KEGG module is a group of predicted functional units); of these, modules 272 were significantly different between smokeless tobacco product brands ($p < 0.05$; Tukey–Kramer), 234 were significantly different between product smokeless tobacco types ($p < 0.05$; Tukey–Kramer), and 261 were significantly different between smokeless tobacco product manufacturers ($p < 0.05$; Tukey–Kramer) (Table S3).

For KEGG hierarchal level 1 functional modules, "Metabolism" was the most abundant module ranging from 47.54 to 55.13%, then "Genetic Information Processing" ranging from 15.0 to 22.9%, "Environmental Information Processing" ranging from 12.8 to 19.9%, "Cellular Processes and Signaling" ranging from 2.7 to 4.6%, and "Cellular Processes" ranging from 0.9 to 4.6% (Fig. S2). Higher level functional modules were significantly different for smokeless tobacco product brands ($p < 0.001$; Tukey–Kramer), types ($p < 0.001$; Tukey–Kramer; except for Cellular Processes and Signaling), and manufacturers ($p < 0.001$; Tukey–Kramer).

Significantly different KEGG hierarchal level 3 modules included plant–pathogen interactions, *Staphylococcus aureus* infections, hydrocarbon degradation, metabolism of xenobiotics, and the biosynthesis of antibiotics (Table S3). The metabolism of xenobiotics and degradation pathways of hydrocarbon compounds (toluene, nitrotoluene, and naphthalene) was predicted to be significantly enriched in Swedish snus and snus compared to other product types ($p < 0.05$; Tukey–Kramer) (Table S3; Figs. S3 and S4). Pathways involved in *S. aureus* infections were predicted to be significantly greater in moist snuff compared to all other product types ($p < 0.001$; Tukey–Kramer) (Table S3), a finding consistent with the significantly higher abundance of OTUs from the *Staphylococcus* genus in moist snuff products ($p < 0.001$; Tukey–Kramer) (Fig. 5). Several predicted antibiotic biosynthesis pathways were also significantly different between smokeless tobacco product types; these

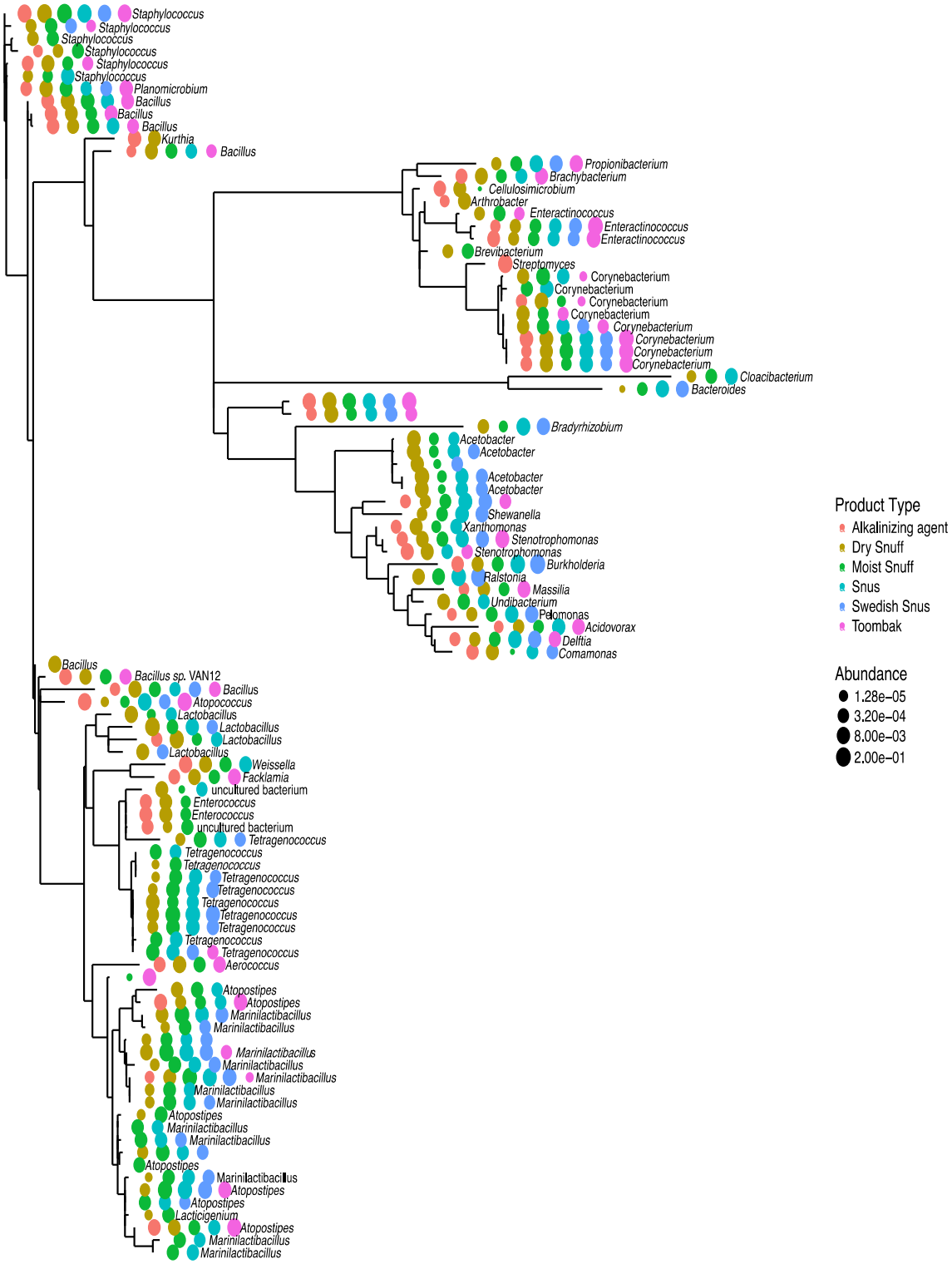

Fig. 5 Phylogenetic tree of the 100 most abundant OTUs with associated genera labeled at the tips. *Color of circle* designates product type, and *size of circle* designates mean relative abundance of OTU per product type on a log scale

included clavulanic acid, penicillin, cephalosporin, novobiocin, streptomycin, tetracycline, and vancomycin biosynthesis pathways ($p < 0.001$; Tukey–Kramer) (Table S3; Fig. S5).

We looked at predicted functions assigned to "Nitrogen metabolism" modules at KEGG hierarchal level 3, and these were present across all product brands, ranging from 0.0052 to 0.0092% of the total number of functions predicted by

PICRUSTt. There was no significant difference between product types ($p > 0.05$; Tukey–Kramer). We also looked for genes involved in nitrite reduction, which has been shown to play a role in TSNA formation by increasing the levels of nitrite in the tobacco. Several nitrite reduction genes (KEGG orthology accessions: K00366, K00362, K00363, K03385, K00368; Table S4) were predicted to be associated with the bacterial metagenome predicted in all the smokeless tobacco products, but there was no significant difference between product types ($p > 0.05$; Tukey–Kramer).

Discussion

This study examined the bacterial diversity of commercially available and handmade smokeless tobacco products and revealed that the majority of tested products harbor viable bacteria, as well as highly diverse bacterial microbiotas and predicted metagenomes. The bacterial microbiota and predicted metagenomes were significantly different when compared by smokeless tobacco product brand, type, and manufacturer. Depending on the product brand and type, these products were dominated by bacteria from the phyla *Firmicutes*, *Proteobacteria*, and *Actinobacteria*. Within these phyla, many of the detected OTUs were identified in genera with known human pathogens, such as *Acetobacter*, *Bacillus*, *Burkholderia*, *Corynebacterium*, *Staphylococcus*, and *Streptomyces*.

Several previous studies have shown that cigarette tobacco harbors diverse bacterial microbiota. These studies identified several specific bacterial species present in tobacco products including *Actinomycetes* spp. (Di Giacomo et al. 2007), *Bacteriovox* sp. (Zhao et al. 2007), *Kurthia* spp., *Bacillus* spp. (Di Giacomo et al. 2007; Huang et al. 2010; Larsson et al. 2008; Rooney et al. 2005; Rubinstein and Pedersen 2002; Zhao et al. 2007), *Pseudomonas* spp. (Huang et al. 2010), *Lactobacillus* spp., and *Staphylococcus* spp. (Di Giacomo et al. 2007). However, these studies focused on cigarette tobacco products not smokeless tobacco products and relied on culture-dependent techniques and/or older sequencing technologies that limited their ability to capture the total bacterial diversity of these products. Members of our group employed a microarray-based approach to broaden our knowledge of the bacterial diversity of tobacco products and showed that commercially available cigarettes harbored a diverse microbiota including bacteria from the phyla *Actinobacteria*, *Bacteroidetes*, *Chloroflexi*, *Cyanobacteria*, *Firmicutes*, and *Proteobacteria* (Sapkota et al. 2010). This previous study also provided evidence that cigarette tobacco harbors potential human pathogens (*Acinetobacter* spp., *Bacillus* spp., *Clostridium* spp., *Klebsiella* spp., *Pseudomonas aeruginosa* spp., and *Serratia* spp.), and we hypothesized that smoking cigarettes could introduce potential pathogens into the oral cavity and lungs of cigarette tobacco users (Sapkota et al. 2010).

In our present study, we focused on smokeless tobacco products which are placed directly in the mouths of users: a direct potential way of introducing bacterial pathogens into the oral cavity. We identified several OTUs within smokeless tobacco products from genera including *Acetobacter*, *Bacillus*, *Burkholderia*, *Corynebacterium*, *Staphylococcus*, and *Streptomyces*, all of which contain species that can include human pathogens. A recent study by Tyx et al. (2016) also demonstrated the presence of bacteria in smokeless tobacco products using culture-independent methods. These authors showed that bacteria from the families *Staphylococcaceae*, *Aerococcaceae*, and *Enterococcaceae* were most abundant in moist snuff; bacteria from the families *Bacillaceae*, *Staphylococcaceae*, and *Lactobacillaceae* were most abundant in dry snuff; and bacteria from the families *Staphylococcaceae* and *Aerococcaceae* were abundant in Toombak. Similarly, Han et al. (2016) recently showed that bacteria from the genera *Tetragenococcus*, *Carnobacterium*, *Lactobacillus*, *Geobacillus*, *Bacillus*, and *Staphylococcus* were dominant in smokeless tobacco products tested using a 16S rRNA sequencing method. In addition, Han et al. (2016) completed culture-based detection methods on the smokeless tobacco products and showed that the following species were most commonly isolated from the tested products: *Bacillus pumilus*, *Bacillus licheniformis*, *Bacillus safensis*, and *Bacillus subtilis.*

In this study, we also identified several predicted antibiotic biosynthesis pathways among bacterial communities within the smokeless tobacco products and these functional attributes might play a role in helping the smokeless tobacco microbiota colonize the oral cavity of users. In addition, we predicted the presence of hydrocarbon degradation pathways in smokeless tobacco products. Smokeless tobacco products have been shown to contain a variety of hydrocarbons such as naphthalene (McAdam et al. 2013), which may provide an important source of carbon to some bacterial groups associated with smokeless tobacco products, potentially providing them with a competitive growth advantage in the tobacco microenvironment.

Another group of chemical consitutents found in smokeless tobacco are TSNAs, which are thought to be among the most important tobacco-associated carcinogens (Stepanov et al. 2012). Smokeless tobacco products have been shown to contain TSNAs in the range of 5218–90,024 ng g^{-1} of smokeless tobacco product with significant variations between brands (Richter et al. 2008). The formation of TSNAs is controlled by the levels of available nitrite, which are influenced by nitrate-reducing bacteria. We detected taxa that are predicted to encode nitrate reduction genes in all smokeless tobacco products. This is in line with a recent

study that also showed the presence of predicted nitrite reduction genes in smokeless tobacco products (Tyx et al. 2016). Specifically, Tyx et al. (2016) showed that bacteria from the families *Corynebacteriaceae*, *Staphylococcaceae*, *Enterobacteriaceae*, and *Bacillaceae*, which were detected in smokeless tobacco products, had predicted nitrite reductase genes which are involved in TSNA formation. Neveretheless, in both the present study and the Tyx et al. (2016) study, these results were based on an in silico prediction tool that has limitations depending on the input data and the environment sampled. However, in a study by Di Giacomo et al. (2007), which utilized both culture methods and molecular techniques to identify isolated strains from dark fire-cured tobacco, the authors showed that bacteria from the genus *Corynebacterium* likely played a role in TSNA formation during tobacco fermentation. Several bacterial OTUs from the *Corynebacterium* genus were also identified in all smokeless tobacco products in our study; however, further study is needed to evaluate whether specific *Corynebacterium* species, as well as other bacterial groups, found in smokeless tobacco play a role in TSNA generation.

Our results also showed that varying smokeless tobacco product types differed significantly with regard to their bacterial microbiotas and the levels of bacteria that could be cultured from each product type. These differences could be partially attributed to the variable physical attributes of the different smokeless tobacco products that were tested. Moist snuff is finely ground or shredded tobacco, dry snuff is fine powdered tobacco, and chewing tobacco comes as a twist, plug, or loose leaf. The fine powdered nature of dry snuff, which equates to a great amount of surface area, compared to other products, could help to explain the higher total viable counts of culturable bacteria observed in dry snuff products in this study. Other smokeless tobacco products such as toombak are handmade where tobacco is mixed with other substances (i.e., alkalinizing agents) that are usually a form of sodium bicarbonate (Hatsukami et al. 2007). This practice might play a role in modifying smokeless tobacco bacterial communities by raising the pH (Idris et al. 1998). Moreover, since sodium bicarbonate has known antimicrobial properties (Gawande et al. 2008), the addition of this substance may serve to suppress the growth of some members of the bacterial communities of smokeless tobacco products. Nevertheless, we found that the alkalinizing agent that we tested harbored a diverse bacterial community and at least some members of this community were viable. This may be due to the fact that this agent is a handmade product and, therefore, more susceptible to sample-to-sample and lot-to-lot variation, as well as environmental contamination during manufacturing.

In addition, the specific manufacturer of the smokeless tobacco products had a significant impact on bacterial microbiota. During the processing of Swedish snus and snus, the products are heat-treated or pasteurized, rather than fermented (Hatsukami et al. 2007). This may explain why cultivation of the Swedish snus and snus products yielded no total viable counts. In addition, it is also very likely that this processing resulted in partial inactivation and potential degradation of the DNA in killed bacterial cells, leading to lower PCR/sequencing yields and lower sequencing counts compared to the other product types.

One OTU (OTU 0) which was highly abundant in all product types was identified by BLAST as *S. cohnii*. This bacterium is associated with human skin (Schleifer and Kloos 1975), and is known to carry antibiotic resistance elements (Zong and Lu 2010). Interestingly, Di Giacomo et al. (2007) also noted a high prevalence of this bacterial species in Italian Toscano cigars that had been processed using dark fire-cured tobacco fermentation. The presence of this OTU in all smokeless tobacco product types tested in this study could potentially be due to shedding of skin bacteria by workers during the harvesting, processing, and factory production of these products.

As noted above, total viable counts of smokeless tobacco were also assessed in this study and our findings demonstrate that viable bacteria are present in the majority of tested products. However, because we (1) used a DNA-based sequencing approach to assess total bacterial diversity across the tested smokeless products and (2) did not perform bacterial idenfication methods on our total viable count plates, we are presently not able to delineate which proportion of bacteria detected with 16S sequencing represents the active and viable bacterial communities in smokeless tobacco products. Recently, Carini et al. (2016) employed a method using a photoreactive DNA intercalating dye, propidium monoazide, to estimate the levels of "relic DNA" in estimates of soil microbial diversity. They showed that up to 40% of DNA was relic DNA, representing inactive bacteria, and this could in some cases affect the estimation of true soil microbial diversity. Further studies of smokeless tobacco that combine sequencing methods with these types of approaches that tease out the active/viable portion of the bacterial microbiota would be of interest. Another limitation of our study is that we focused on the bacterial microbiota and not the fungal microbiota of these products. However, the fungal microbiota of these products may have important health implications for smokeless tobacco product users and deserves further attention.

In summary, our findings show that the majority of smokeless tobacco products tested in this study harbor diverse bacterial microbiotas, and the bacterial microbiotas, as well as levels of culturable bacteria, differ significantly depending on the brand, type, and manufacturer of the product. The public health implications of these results are not fully understood at this time, and further research is needed to examine the role

that smokeless tobacco bacterial microbiota may play in introducing potentially harmful bacteria into users' oral cavities, as well as the effects that this might have on the oral microbiome and users' health.

Acknowledgements We would like to thank Ghazi Zaatari, M.D. (Department of Pathology and Laboratory Medicine, American University of Beirut; Beirut, Lebanon) for providing the toombak samples and alkalinizing agent and Dr. Angel J. Rivera for his helpful review and input on this manuscript.

Compliance with ethical standards

Funding This study was funded by a contract (200-2012-M-51035) from the Centers for Disease Control and Prevention (CDC) and by grant number P50CA180523 from the National Cancer Institute and FDA Center for Tobacco Products (CTP) awarded to the University of Maryland. The content is solely the responsibility of the authors and does not necessarily represent the official views of CDC, NIH, or the Food and Drug Administration.

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215(3):403–410. doi:10.1016/S0022-2836(05)80360-2

Ayo-Yusuf OA, Burns DM (2012) The complexity of 'harm reduction' with smokeless tobacco as an approach to tobacco control in low-income and middle-income countries. Tob Control 21(2):245–251. doi:10.1136/tobaccocontrol-2011-050367

Boffetta P, Hecht S, Gray N, Gupta P, Straif K (2008) Smokeless tobacco and cancer. Lancet Oncol 9(7):667–675. doi:10.1016/S1470-2045(08)70173-6

Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R (2010a) PyNAST: a flexible tool for aligning sequences to a template alignment. Bioinformatics 26(2):266–267. doi:10.1093/bioinformatics/btp636

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich J, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010b) QIIME allows analysis of highthroughput community sequencing data. Nat Methods 7(5):335–336. doi:10.1038/nmeth0510-335

Carini P, Marsden PJ, Leff JW, Morgan EE, Strickland MS, Fierer N (2016) Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. Nat Microbiol 2:16242. doi:10.1038/nmicrobiol.2016.242

Di Giacomo M, Paolino M, Silvestro D, Vigliotta G, Imperi F, Visca P, Alifano P, Parente D (2007) Microbial community structure and dynamics of dark fire-cured tobacco fermentation. Appl Environ Microbiol 73(3):825–837. doi:10.1128/AEM.02378-06

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26(19):2460–2461. doi:10.1093/bioinformatics/btq461

Fisher MT, Bennett CB, Hayes A, Kargalioglu Y, Knox BL, Xu D, Muhammad-Kah R, Gaworski CL (2012) Sources of and technical approaches for the abatement of tobacco specific nitrosamine formation in moist smokeless tobacco products. Food Chem Toxicol 50(3–4):942–948. doi:10.1016/j.fct.2011.11.035

Gawande PV, LoVetri K, Yakandawala N, Romeo T, Zhanel GG, Cvitkovitch DG, Madhyastha S (2008) Antibiofilm activity of sodium bicarbonate, sodium metaperiodate and SDS combination against dental unit waterline-associated bacteria and yeast. J Appl Microbiol 105(4):986–992. doi:10.1111/j.1365-2672.2008.03823.x

Han J, Sanad YM, Deck J, Sutherland JB, Li Z, Walters MJ, Duran N, Holman MR, Foley SL (2016) Bacterial populations associated with smokeless tobacco products. Appl Environ Microbiol 82(20):6273–6283. doi:10.1128/AEM.01612-16

Hatsukami DK, Ebbert JO, Feuer RM, Stepanov I, Hecht SS (2007) Changing smokeless tobacco products new tobacco-delivery systems. Am J Prev Med 33(6 Suppl):S368–S378. doi:10.1016/j.amepre.2007.09.005

Huang J, Yang J, Duan Y, Gu W, Gong X, Zhe W, Su C, Zhang KQ (2010) Bacterial diversities on unaged and aging flue-cured tobacco leaves estimated by 16S rRNA sequence analysis. Appl Microbiol Biotechnol 88(2):553–562. doi:10.1007/s00253-010-2763-4

Idris AM, Ibrahim SO, Vasstrand EN, Johannessen AC, Lillehaug JR, Magnusson B, Wallstrom M, Hirsch JM, Nilsen R (1998) The Swedish snus and the Sudanese toombak: are they different? Oral Oncol 34(6):558–566

Juarez SP, Merlo J (2013) The effect of Swedish snuff (Snus) on offspring birthweight: a sibling analysis. PLoS One 8(6):e65611. doi:10.1371/journal.pone.0065611

Langille MG, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, Clemente JC, Burkepile DE, Vega Thurber RL, Knight R, Beiko RG, Huttenhower C (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol 31(9):814–821. doi:10.1038/nbt.2676

Larsson L, Szponar B, Ridha B, Pehrson C, Dutkiewicz J, Krysinska-Traczyk E, Sitkowska J (2008) Identification of bacterial and fungal components in tobacco and tobacco smoke. Tob Induc Dis 4:4. doi:10.1186/1617-9625-4-4

McAdam KG, Faizi A, Kimpton H, Porter A, Rodu B (2013) Polycyclic aromatic hydrocarbons in US and Swedish smokeless tobacco products. Chem Cent J 7(151). doi:10.1186/1752-153X-7-151

McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J 6(3):610–618. doi:10.1038/ismej.2011.139

McMurdie PJ, Holmes S (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 8(4):e61217. doi:10.1371/journal.pone.0061217

Parks DH, Tyson GW, Hugenholtz P, Beiko RG (2014) STAMP: statistical analysis of taxonomic and functional profiles. Bioinformatics 30(21):3123–3124. doi:10.1093/bioinformatics/btu494

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41(Database issue):D590–D596. doi:10.1093/nar/gks1219

R Core Team (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

R Developmental Core Team (2010) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Richter P, Hodge K, Stanfill S, Zhang L, Watson C (2008) Surveillance of moist snuff: total nicotine, moisture, pH, un-ionized nicotine, and tobacco-specific nitrosamines. Nicotine Tob Res 10(11):1645–1652. doi:10.1080/14622200802412937

Rooney AP, Swezey JL, Wicklow DT, McAtee MJ (2005) Bacterial species diversity in cigarettes linked to an investigation of severe pneumonitis in U.S. Military personnel deployed in operation iraqi freedom. Curr Microbiol 51(1):46–52. doi:10.1007/s00284-005-4491-z

Rubinstein I, Pedersen GW (2002) *Bacillus* species are present in chewing tobacco sold in the United States and evoke plasma exudation from the oral mucosa. Clin Vaccine Immunol 9(5):1057–1060. doi:10.1128/cdli.9.5.1057-1060.2002

Sapkota AR, Berger S, Vogel TM (2010) Human pathogens abundant in the bacterial metagenome of cigarettes. Environ Health Perspect 118(3):351–356. doi:10.1289/ehp.0901201

Schleifer KH, Kloos WE (1975) Isolation and characterization of Staphylococci from human skin. International Journal of Systemic Bacteriology 25(1):50–61

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22(21):2688–2690. doi:10.1093/bioinformatics/btl446

Stepanov I, Jensen J, Hatsukami D, Hecht SS (2008) New and traditional smokeless tobacco: comparison of toxicant and carcinogen levels. Nicotine Tob Res 10(12):1773–1782. doi:10.1080/14622200802443544

Stepanov I, Knezevich A, Zhang L, Watson CH, Hatsukami DK, Hecht SS (2012) Carcinogenic tobacco-specific N-nitrosamines in US cigarettes: three decades of remarkable neglect by the tobacco industry. Tob Control 21(1):44–48. doi:10.1136/tc.2010.042192

Tyx RE, Stanfill SB, Keong LM, Rivera AJ, Satten GA, Watson CH (2016) Characterization of bacterial communities in selected smokeless tobacco products using 16S rDNA analysis. PLoS One 11(1):e0146939. doi:10.1371/journal.pone.0146939

USDHHS (2014) The health consequences of smoking—50 years of progress: a report of the surgeon general. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health., Atlanta, Georgia

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 73(16):5261–5267. doi:10.1128/AEM.00062-07

Wei X, Deng X, Cai D, Ji Z, Wang C, Yu J, Li J, Chen S (2014) Decreased tobacco-specific nitrosamines by microbial treatment with *Bacillus amyloliquefaciens* DA9 during the air-curing process of burley tobacco. J Agric Food Chem 62(52):12701–12706. doi:10.1021/jf504084z

Wickham H (2009) ggplot2: elegant graphics for data analysis. Springer Publishing Company, Incorporated, New York

Zhang X, Schmitz W, Gelderblom HR, Reichart PA (2001) Shammah-induced oral leukoplakia-like lesions. Oral Oncol 37:609–612

Zhao M, Wang B, Li F, Qiu L, Li F, Wang S, Cui J (2007) Analysis of bacterial communities on aging flue-cured tobacco leaves by 16S rDNA PCR-DGGE technology. Appl Microbiol Biotechnol 73(6):1435–1440. doi:10.1007/s00253-006-0625-x

Zong Z, Lu X (2010) Characterization of a new SCCmec element in *Staphylococcus cohnii*. PLoS One 5(11):e14016. doi:10.1371/journal.pone.0014016

Zupancic ML, Cantarel BL, Liu Z, Drabek EF, Ryan KA, Cirimotich S, Jones C, Knight R, Walters WA, Knights D, Mongodin EF, Horenstein RB, Mitchell BD, Steinle N, Snitker S, Shuldiner AR, Fraser CM (2012) Analysis of the gut microbiota in the old order Amish and its relation to the metabolic syndrome. PLoS One 7(8):e43052. doi:10.1371/journal.pone.0043052