Applied Microbiology and Biotechnology (2020) 104:10613–10629
https://doi.org/10.1007/s00253-020-10999-w

GENOMICS, TRANSCRIPTOMICS, PROTEOMICS

Microbial communities and gene contributions in smokeless tobacco products

A. J. Rivera[1] · R. E. Tyx[1] · L. M. Keong[2,3] · S. B. Stanfill[1] · C. H. Watson[1]

Received: 7 May 2020 / Revised: 29 October 2020 / Accepted: 3 November 2020 / Published online: 12 November 2020

Abstract

Smokeless tobacco products (STP) contain bacteria, mold, and fungi due to exposure from surrounding environments and tobacco processing. This has been a cause for concern since the presence of microorganisms has been linked to the formation of highly carcinogenic tobacco-specific nitrosamines. These communities have also been reported to produce toxins and other pro-inflammatory molecules that can cause mouth lesions and elicit inflammatory responses in STP users. Moreover, microbial species in these products could transfer to the mouth and gastrointestinal tract, potentially altering the established respective microbiotas of the consumer. Here, we present the first metagenomic analysis of select smokeless tobacco products, specifically US domestic moist and dry snuff. Bacterial, eukaryotic, and viral species were found in all tobacco products where 68% of the total species was comprised of *Bacteria* with 3 dominant phyla but also included 32% *Eukarya* and 1% share abundance for *Archaea* and *Viruses*. Furthermore, 693,318 genes were found to be present and included nitrate and nitrite reduction and transport enzymes, antibiotic resistance genes associated with resistance to vancomycin, β-lactamases, their derivatives, and other antibiotics, as well as genes encoding multi-drug transporters and efflux pumps. Additional analyses showed the presence of endo- and exotoxin genes in addition to other molecules associated with inflammatory responses. Our results present a novel aspect of the smokeless tobacco microbiome and provide a better understanding of these products' microbiology.

Key points

- *The findings presented will help understand microbial contributions to overall STP chemistries.*
- *Gene function categorization reveals harmful constituents outside canonical forms.*
- *Pathway genes for TSNA precursor activity may occur at early stages of production.*
- *Bacteria in STPs carry antibiotic resistance genes and gene transfer mechanisms.*

Keywords Smokeless tobacco · Microbial communities · Metagenome · Metagenomics

Introduction

Significant concern exists over the potential health risks associated with diverse microbial elements known to be present in smokeless tobacco products (STPs) in the domestic market. STPs are cured tobacco leaves sweetened and flavored to yield various products in different forms. These products are characterized for being non-combustible contrary to cigarettes, electronic cigarettes, or similar nicotine delivery systems. Common consumption practices include chewing, sniffing, or placing the product between the gums and cheeks where it releases nicotine and other substances. These products can be viewed as a dynamic microbial system in which bacterial communities proliferate, metabolize, and excrete compounds. Consequently, it can be hypothesized that even at a small scale within a single tin, in a single production lot, discrete ecosystems can exist with unique microbiotas and chemical composition. In it, a myriad of biochemical reactions produces a series of conditions and by-products whose effects could potentially present significant challenges to the long-term health of STP users.

✉ A. J. Rivera
yic7@cdc.gov

[1] Division of Laboratory Sciences, Centers for Disease Control and Prevention, 4770 Buford Highway, NE MS 110-03, Atlanta, GA 30341-3717, USA

[2] Battelle Analytical Services, Atlanta, GA, USA

[3] Influenza Division, Centers for Disease Control and Prevention, Atlanta, GA, USA

Past studies on health effects associated with STPs have focused primarily on chemical constituents and associated exposure to select toxic and carcinogenic compounds absorbed during use (Critchley and Unal 2003). However, scarce information on microbial presence, identification, and functionality has been reported and many information gaps need to be addressed. Previous biological research on tobacco or tobacco products relied heavily on older culture-based approaches that could bias the data output and misrepresent community diversity estimations. Only a few studies have used more modern, culture-independent methods, mainly 16S rDNA analyses of STPs aimed at characterizing bacterial communities' composition (Monika et al. 2020; Sami et al. 2020; Tyx et al. 2016). Although 16S rDNA-based taxonomical classification is now becoming an important part of species identification in tobacco research, it does not provide the evidence to fully address the genetic contributions by the microbial occurrence in these products.

Only recently, a comprehensive molecular survey was reported affirming that microorganisms were present in varying diversities, some of which were predicted to have the potential to generate toxins, pro-inflammatory biomolecules, and possibly precursors for the formation of tobacco-specific N'-nitrosamines (TSNAs) (Tyx et al. 2016). In a succeeding report, similar taxonomic classifications were obtained, and the authors also suggested that potential opportunistic pathogens could be present in the STP microbiota (Han et al. 2016). Although these studies shed light on the bacterial composition of STPs and further our understanding, questions stand about the genetic potential of these organisms and the impact of their metabolic capabilities.

The generation of potentially harmful by-products may originate primarily from curing and aging steps during the production of moist and dry snuffs (Law et al. 2016). Supporting arguments for this are found in processing and storage conditions modifications (e.g., pasteurization) for products such as Scandinavian snus where lower microbial quantities lead to a decreased levels of harmful TSNA chemical constituents (Law et al. 2016). Thus, microbial metabolism could play a role in the resulting chemistry of the untreated products. Consequently, characterizing STP's microbial taxa and their genetic contexts may be an important step towards a more robust understanding of the overall product microbiology and its impact on harm.

Metagenomic analyses provide a new dimension in the study of community organization and metabolism in natural microbial communities (Law et al. 2016). This report presents the first approach characterizing STP microbiota and their genetic context using shotgun metagenomics analysis. The findings will help understand microbial communities' contributions to the overall chemistry of this type of tobacco products.

Materials and methods

Tobacco sample collection and DNA extraction Two US domestic moist and two dry snuff tobacco samples were obtained by an outside vendor at undisclosed retail sites. All samples received were kept sealed, barcoded, and stored at − 80 °C. Smokeless tobacco samples were thawed to room temperature and 200 mg of ST material used for total genomic DNA extraction. An enzyme cocktail was prepared by mixing 5-μl lysostaphin (5 μg/μl), 5-μl lysozyme (10 μg/μl), and 15-μl mutanolysin (1 μg/μl). Separately, the measured smokeless tobacco samples and 1000-μl molecular grade 1X phosphate-buffered saline (PBS) were added to Lysing Matrix J bead-beating tubes (MP Biomedicals, Santa Ana, CA, USA). The 25-μl aliquot of the enzyme cocktail was added to the bead-beating tubes and the resulting mixture incubated at 37 °C for 30 min. Immediately upon incubation completion, 10 μl of proteinase K (20 μg/μl) and 50 μl of 10% SDS were added to the tubes and incubated at 55 °C for 30 min. Following incubation, tubes were agitated at 4800 RPM for 2 min in a Mini-BeadBeater (BioSpec, Inc.; Bartlesville, OK, USA). Sample tubes were centrifuged at 10,000*g* for 5 min, and the supernatant was transferred to a new tube. Each sample was then diluted 2:1 with 100% ethanol. Binding, washing, and elution of DNA was accomplished using QIAamp Mini Spin columns (Qiagen Sciences Inc.; Germantown, MD, USA) as specified by the manufacturer. Best results were obtained when DNA was passed through a second QIAamp Mini Spin column as described previously (Tyx et al. 2016). All aforementioned reagents were obtained from Sigma-Aldrich (St. Louis, MO, USA).

DNA libraries and shotgun sequencing Library preparation was performed by loading extracted genomic DNA into a screw-cap 6 × 16 microTube and optimally sheared with a Covaris M220 focused-ultrasonicator (Covaris, Inc., Woburn, MA, USA), to an approximate size of 550 bp. DNA Input concentrations of 200 ng per sample were used in accordance with Illumina® TrueSeq Nano DNA LS Samples Preparation (Part # 15041110 Rev. A, May 2013). Each sample DNA was subjected to end repair, adenylation, and Illumina adaptor ligation steps as indicated by manufacturer, then immediately followed by a DNA fragment enrichment step using Illumina reagents (Part #15041757). Quantification was performed using a fluorometer (Qubit 2.0, Life Tech. Carlsbad, CA, USA), while insert size and quality were determined by running DNA in a high-sensitivity DNA assay chip and reading in a Bioanalyzer 2100 (Agilent Technologies, Waldbornn, Germany). To address possible low diversity, normalization of DNA libraries was achieved by adding 5% Control PhiX1 V3 genomic DNA.

Shotgun DNA sequencing runs were performed in-house using a MiSeq 2 X 300 V3 sequencing Kit and sequence data obtained from an Illumina MiSeq sequencer (Illumina, Inc.; San Diego, CA, USA).

Sequence data processing and assembly Raw sequences were checked for quality with FastQC v0.11.3 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Read processing was performed using the bbduk.sh script (part of BBMAP v34.72, Ref: BBMap - Bushnell B. - sourceforge.net/projects/bbmap/) to remove sequence adapters and phiX control sequences. Quality filtering was adjusted for a minimum insert size of 60 base pairs, and otherwise, default parameters were performed using SICKLE v1.33 (Joshi and Fass 2011). To improve assembly, reads were normalized using the bbnorm.sh script (also from BBMAP v34.72) applying default parameters. SPADES meta-assembler v3.10 was used for all metagenomes (Nurk et al. 2017). Assemblies based on size of the output assembly and L50 (given by BBMAP stats.sh script) were chosen for further analysis (Table S10, Figure S1).

Function categorization Final assemblies for each STP were uploaded for functional annotation by submitting datasets to the integrated microbial genomes and metagenomes (IGM/M) pipelines (MGAP v.4) (Chen et al. 2019). Functional annotation was achieved by feature prediction tools that included the identification of protein-coding genes and other genetic elements. Structural annotation followed by assignment of protein product names and functions under parameters is described by Huntemann et al. (2015). Function categorization is then performed via comparative analysis executed by the "Function Profile" tool within the IMG/MER package. Each metagenomic datasets was then interrogated against multiple public datasets, and relative abundance of protein families was identified (Chen et al. 2017).

The IMG Genome Taxon ID numbers are as follows: D1 3300020227, D2 3300020212, M1 3300020216, M2 3300020217. Raw sequence files were filed at NCBI SRA, accession SRP158032.

Results

Microbial taxonomic composition

We selected four smokeless tobacco products of the snuff type, two dry (D1 and D2) and two moist (M1 and M2). Sequence assemblies produced a total of 847,251 scaffolds for which 693,318 genes were found in all metagenomes. The domain *Bacteria* was established as the most abundant across all STPs, with 68% of the sequence hits identified as *Bacteria*, whereas *Eukaryotes* accounted for 32% of genes identified. *Archaea* and *Viruses* were much less abundant and accounted for less than 1% of the taxonomic diversity in these metagenomes, as determined by gene hits for these domains. Although our shotgun metagenomic approach allowed the recognition of organisms across all domains of life, microbial species were the focus of these studies. Thus, it is important to note that a number of reads in these products mapped to tobacco plant genes in the *Streptophyta* phylum as has been previously observed (Behzad et al. 2015; Chang et al. 2017; Uyaguari-Diaz et al. 2016). However, because our focus resides on microbial communities, these were considered superfluous sequences from plant material in STPs and were not analyzed further.

As the most abundant domain across all STPs, *Bacteria* showed taxonomic sub-classifications mainly within 3 dominant phyla (Fig. 1). *Firmicutes* are the most abundant contributor, comprising 35.5% of the taxa observed and found in all snuff products. Interestingly, moist products contained higher percentages (87.5–98.0%) than dry ones (47.10–49.0%). *Proteobacteria*, the second most abundant phylum with 25.5% of the total bacterial domain, showed a higher number of sequences in dry snuff (36.1–49.0%) over the moist products (≤ 0.96%). Finally, *Actinobacteria* the third most dominant group covered 5.7% of the total bacterial community with nominal abundances across all products.

In the phylum *Firmicutes*, the highest abundances were found for 5 families in the metagenomes of all snuff products (Fig. 1, circles 5–9). The highest sequence hits were found for *Bacillaceae* (≤ 79.1%) and *Staphylococcaceae* (≤ 76.4%) in all products but specifically higher in dry snuffs (Fig. 2a). *Carnobactericeae*, the third most abundant group, is also present in all products. Three products (D1, D2, and M1), ranged from ~ 21 to 29% of the total sequence hits in *Firmicutes*, with a markedly dominant 81.5% found for product M2 (Fig. 2a). The remaining two families differed in their highest abundances across the all products, starting with *Lactobacillaceae*, which was found dominant mainly in dry snuff with the highest abundance (43.1%) in product D2. Conversely, *Enterococcaceae* appeared dominant in moist snuff with high sequence hits (64.7%) in product M1 (Fig. 2a).

Proteobacteria ranked second in abundance within the domain *Bacteria* and exhibited dominant species in dry snuff by only six families of the phylum representing a 6% of the total community composition of all the products (Fig. 1, circles 1–4). The highest sequence hits were found for *Erwinaceae*, *Burkholderiaceae*, *Enterobactericeae*, and *Pseudomonadaceae* all predominantly in snuff D1, while *Lysobacteraceae* and *Halomonadaceae* were unique to snuff D2 (Fig. 2b). Conversely, considerably lower number of sequence hits was found in both moist products (≤ 10 hits/taxon) suggesting that a nominal number of *Proteobacteria* species were present in moist snuffs at the time of sampling (Fig. 2b).

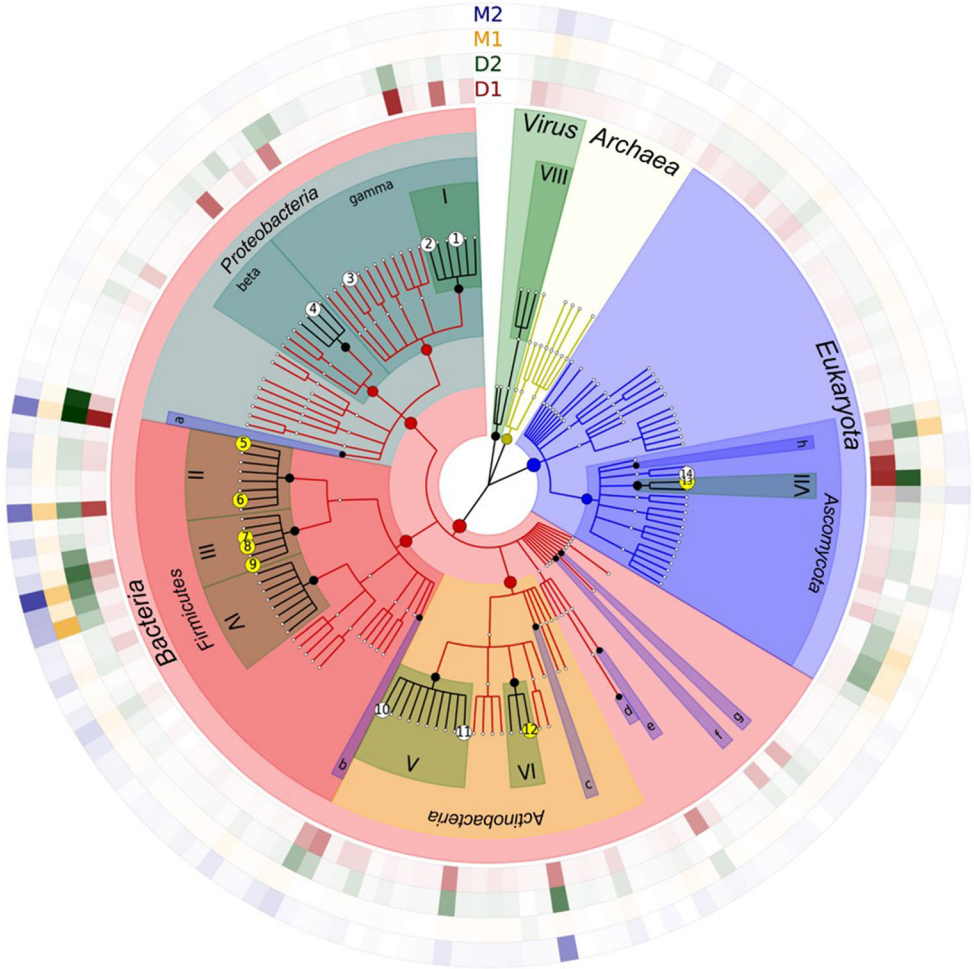

Fig. 1 Taxa distribution in smokeless tobacco products is highly complex. Species relationships are presented in a taxonomical hierarchy diagram progressing to the family level from sequences found in all STPs. At the center of the tree, domains *Bacteria* (red), *Eukaryota* (blue), *Archaea* (yellow), and *Virus* (black) are the ancestral points for each group. Similarly, red-shaded backgrounds denote the bacterial domain species, blue-shaded backgrounds illustrate the observed eukaryotic species in these products, whereas viral shown in green and archaeal in yellow are also included. Branching extends through nodes denoting phylum, class, order, and family by each circle along the branches. As these branches progress outward, black color nodes denote taxa for which relevant gene contributions were found as number of sequences in the metagenomes. The roman numerals represent orders for which abundant taxa were found ((I) *Enterobacterales*, (II) *Bacillales*, (III) *Lactobacillales*, (IV) *Clostridiales*, (V) *Micrococales*, (VI) *Corynebacteriales*, (VII) *Eurotiales*, (VIII) *Caudovirales*). Circled numbers indicate specific families within those orders to which these abundant taxa were assigned ((1) *Enterobactericeae*, (2) *Erwiniaceae*, (3) *Pseudomonadaceae*, (4) *Burkholderiaceae*, (5) *Bacillaceae*, (6) *Staphylococcaceae*, (7) *Lactobacillaceae*, (8) *Carnobacteriaceae*, (9) *Enterococaceae*, (10) *Brevibacteriaceae*, (11) *Micrococcaceae*, (12) *Corynebacteriaceae*, (13) *Aspergillaceae*, (14) *Pleosporaceae*). Lower case letters indicate taxonomic groups showing substantial number of sequences but classified as unclassified or moderate relevant sequence hits outside the main abundant phyla ((a) unclassified *Proteobacteria*, (b) unclassified *Actinobacteria*, (c) unclassified *Bacteroidetes*, (d) *Bacteroidaceae*, (e) *Cytophagales*, (f) *Chloroflexi*, (g) *Cyanobacteria*, (h) unclassified *Ascomycota*). The concentric rings bordering the radial tree denote heatmaps showing that STP-associated species relative abundances for each product by measure of sequence read hits against the Integrated Microbial Genomes - Expert Review (IMG/ER) database (Chen et al. 2019). Innermost rings represent dry snuff products D1 and D2, while the outermost rings represent moist snuff products M1 and M2

Actinobacteria was the third most dominant phylum found for all metagenomes (Fig. 1). The highest sequence hits classified within this phylogenetic group were limited to families (Fig. 1, circles 10–12) *Corynebactericeae* (54%) in this phylum and the highest abundance in dry snuffs, principally in snuff D2. This was followed by *Micrococcaceae* (23%), found in dry snuff products with significantly higher and comparable sequence hits in both dry snuffs but negligible sequence number in the moist products. Third in standing (14%) was *Brevibactericeae*, found with significant number of sequences only in dry snuffs while both moist products showed nominal abundances (Fig. 2c).

Analyses identified other major taxonomical classifications outside the bacterial domain in these STPs. Our data suggests the existence of abundant microbial eukaryotic species for which 15.4% of all sequence hits were classified to the

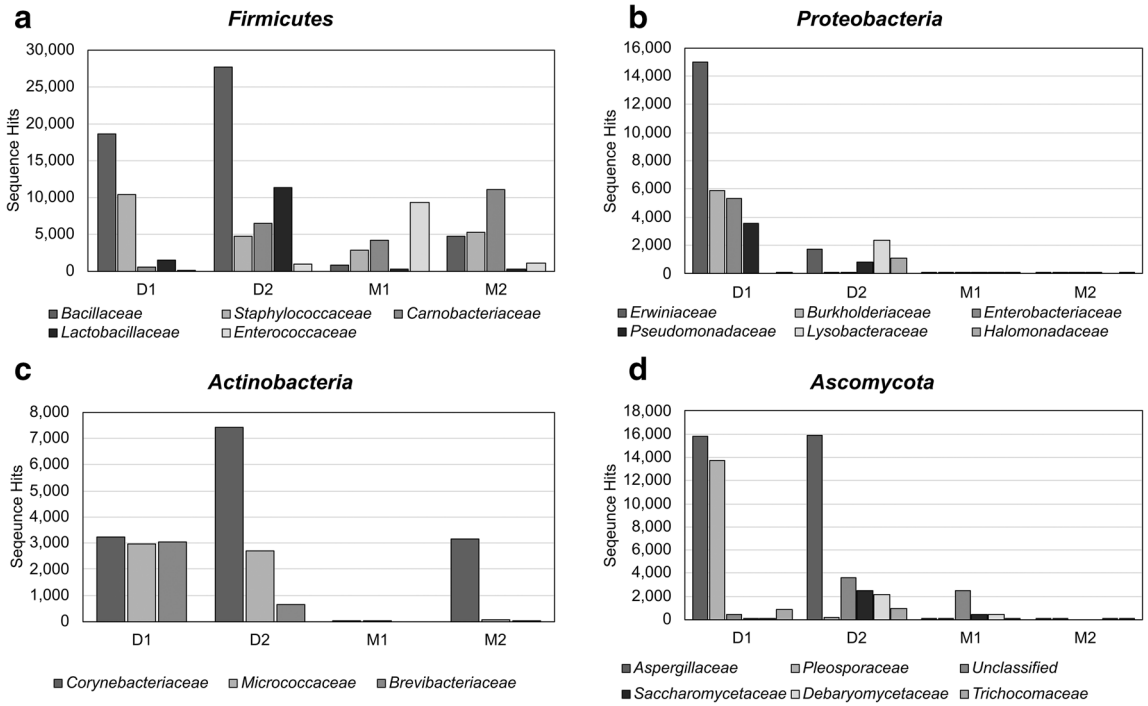

Fig. 2 Dominat taxa in smokeless tobacco. The most abundant found families across all STPs by sequence hits for each snuff product

phylum *Ascomycota* (Fig. 1). We found these taxa abundant in both dry and moist metagenomes, yet product-specific taxonomic distributions for sub-classifications of the group were observed to be abundant for dry but not in moist snuffs (Fig. 1, circles 13–14). The family of *Aspergillaceae,* the larger group (45%) in *Ascomycota* showed a significant number of hits for dry snuffs D1 and D2. The remaining moist snuff products showed comparatively negligible sequence hits (Fig. 2d). The second most abundant family (20%) *Pleosporaceae* constituted a large portion of the *Ascomycota* in product D1 but smaller sequence hit numbers in the remaining products (Fig. 2d). It is noteworthy to that the families *Saccharomycetaceae*, *Debaryomycetaceae*, and *Trichocomaceae* were also of moderate abundance in this phylum for dry snuff products (Fig. 2d). Moreover, unclassified *Ascomycota*-related species were found in both dry snuffs and moist snuff M1 to a lesser degree.

The domains *Archaea* and *Virus* are two additional taxonomic groups in the metagenomes studied and found to constitute approximately 0.5% of the total community as measured by sequence counts. The small number of sequences related to *Archaea* was low when compared to all other major taxonomic groups found in these products, and their genetic and metabolic processes not fully elucidated. Thus, we did not attempt further characterization of the *Archaeal* members of the community. While viral sequences were also identified in moderately lower numbers, we obtained enough sequence hits to determine a limited phylogenetic context. It is important to note our approach only allowed for the collection of total genomic DNA, but not RNA; therefore, the sequences found only belong to the dsDNA *Virus* classification within the *Virus* domain. The totality of the groups identified within this domain was represented by the order *Caudovirales*, where most species clustered to the families *Myoviridae*, *Podoviridae*, and *Siphoviridae*. However, *Myoviridae* reached higher sequence hits with a 64% of the total virus-associated sequences. The remaining two groups were present in less than 16% of the total viral community. Sequences belonging to other major eukaryotic taxonomic groups were identified, albeit in significantly lower abundances (< 0.5% of total abundance). Inherent limitations in IMG's reference sequence databases (being focused on microbial life) for the *Eukaryotes* prevented further characterization of this data. Consequently, we have regarded these sequences as possible remnants of transient contaminants (i.e., single and multi-cell protozoans, diatoms, worms, insects, from soil or water) that may have accumulated along the production process, and fall outside the scope of this report for further investigation.

NO_3^- reduction and NO_2^- transport are believed to be precursor steps in the formation of carcinogens TSNAs (Shi et al. 2013). Function category analysis of assemblies suggests the presence of nitrate and nitrite reduction and transport genes with prominent abundances in dry snuff products. Conversely, identified gene sequences in moist snuff samples were limited to specific nitrate/nitrite transporters and reductases (Fig. 3). Functional analysis showed estimated gene copy numbers to an average of approximately 500 sequence hits for each bacterial nitrate reductase encoding genes *narGHJI* (Table S1). Protein sequence alignments suggest that these genes may originate from

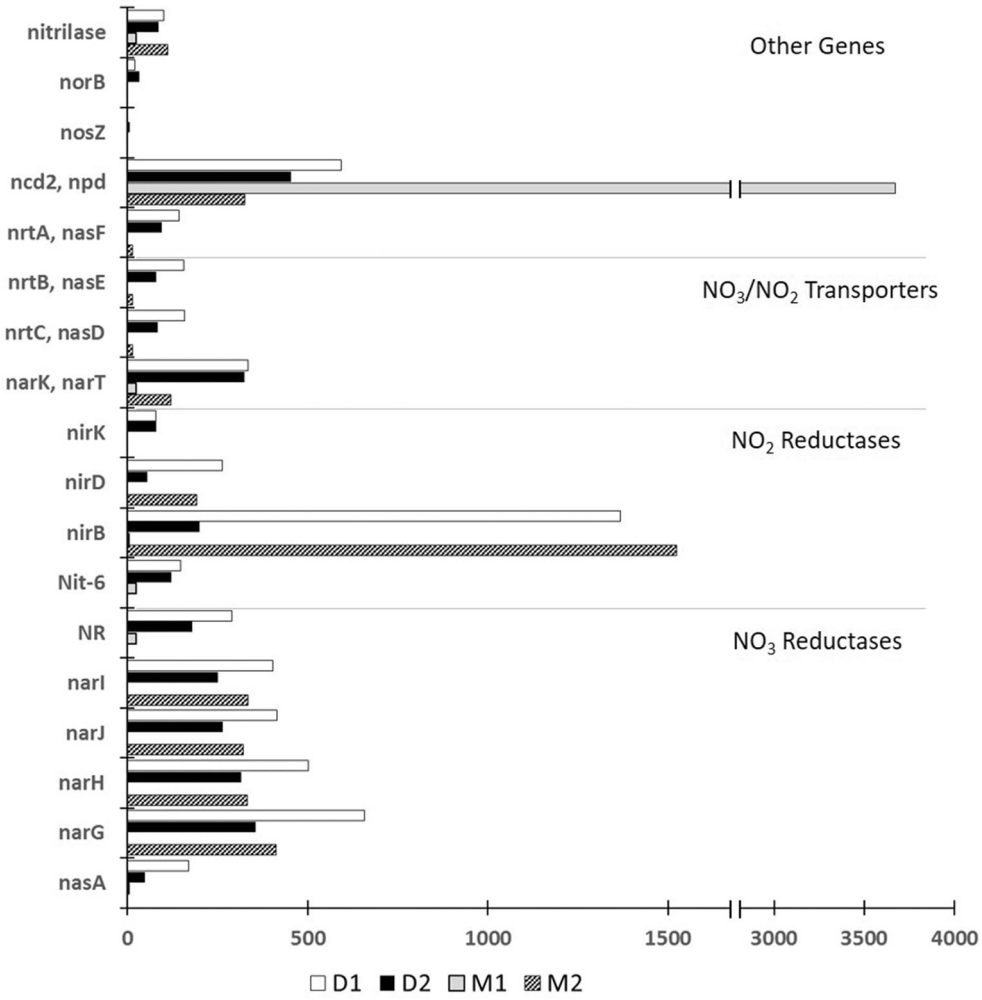

Fig. 3 Nitrate and nitrite reduction and transport are important in the formation of TSNAs. Sequence hits for nitrogen metabolism-associated pathways subsets to nitrate reduction (*nar*), nitrate reduction (*nir*), and nitrate/nitrite (*nas*, *nrt*) transport. Sequences found in significant numbers for NMO (formerly *npd*) nitronate monooxygenase

various microbial species in these products where the more abundant taxa were found to be *Staphylococcus* species in D1 and M2 and *Corynebacterium* species in sample D2 (Table S2a). Gene sequences for each nitrate reductase subunit were consistent across all samples with the distinct exception of moist sample M1 where no *nar* genes were found. However, assimilatory nitrate reductase encoded by *nasA* was identified in products D1, D2, and M1 with remarkably low estimated gene copy numbers, while none was observed in M2 (Table S1). Fungal species associated nitrate reductase (NR) was observed in three out of the four products, though with markedly low gene copy numbers in M1 and completely absent in M2. Sequence hits to nitrite reductase genes were found in lower numbers than nitrate reductase (Fig. 3). However, nitrite reductase gene *nirB* was found to have the highest sequence hits for all reductases, primarily in products D1 with 74% and M2 with 89% of the total nitrite reductase gene sequences in each product (Fig. 3). Like nitrate reductases, sequence alignments suggest that these genes may originate from *Staphylococcal* species (Table S2c).

Estimated gene copies for NO_3/NO_2 transporter genes in products D1, D2, and M2 were abundantly distributed but not for M1, where most of the transporter genes were absent (Table S1). Only two transporter systems encoded by *nar* and *nrt/nas* genes were present in these data sets. Nevertheless, transporter genes *narK* and *narT* appeared to be the dominant transport mechanisms in those species across all products analyzed. Additional genes linked to nitrogen metabolism were found but in low gene copy numbers compared to those for nitrate/nitrite reduction and transport. These were observed only in dry products and identified as enzymes that catalyze nitric oxide to nitrous oxide (*norB*) and subsequently to nitrogen (*nosZ*). Additionally, the enzyme nitrilase that generates ammonium from nitrile was identified in all samples (Table S1).

A single gene was found in significantly higher estimated gene copies than any other sequence in the nitrogen metabolism genes group (Table S1). These sequences mapped to the enzyme nitronate monooxygenase (NMO) were formerly referred to as 2-nitropropane dioxygenase (NPD) (Gadda and Francis 2010). The gene was identified in all STPs, but a

significant number of gene copies were found in product M1. Sequence alignments suggested that this gene is likely associated to species of *Bacilli*, *Ralstonia*, *Staphylococcus*, and fungal genera *Aspergillus* and *Alternaria* in product D1 (Table S2d). Product D2 showed similar species with highest abundances observed for *Corynebacterium ammoniagenes*. Conversely, moist products were marked by an extremely low diversity of taxon and only two species, *Staphylococcus cohnii* in M1and *Corynebacterium ammoniagenes* in M2, were found to be the organisms potentially contributing with NMO gene sequences. Additionally, *S. cohnii* had a significantly high number of estimated gene copies (3645) in M1, with abundance reaching 73% over any other gene found in the products included in this study. We underscore that this gene may represent an additional mechanism to generate nitrite and consequently TSNAs. However, we also note that annotations for NMO have been identified as inaccurate for some sequences in the database and in need of revision before actual metabolic activity can be definitively assigned (Ball et al. 2016).

Bacterial associated toxins' functional category identified based on KEGG orthology suggested that bacterial species carrying toxin genes were present in these products at the point of sampling. The highest and more significant sequence hits were observed for genes *hlyIII* which encodes the membrane-damaging toxin hemolysin III (Table 1). The gene sequence was abundant in all products; however, microbial species linked to this gene varied from product to product. The highest gene copy numbers were found for species of *Staphylococcus* in M1, while dry snuffs showed higher diversity. Interestingly, product D2 showed the highest number of sequences annotated as membrane protein associated to *Lactobacillus pobuzihii*. Moist snuff products appeared to have low species diversity, but estimated gene copies indicated high abundance of close relatives of *Staphylococcus* and *Tetragenococcus* species (Table S3a). A closely related gene annotated as putative hemolysin *tlyC* also had comparable gene copy numbers to *hlyIII* for all products except D1. Amino acid sequence alignment data showed only D2 and M1 to have sequence identities higher than 90% (Table S3b). As observed before, D2 was dominated by *Lactobacillus* species, while M1 showed only *Tetragenococcus* species as the taxa contributing these genes. Sequence alignments returned genes annotated as putative transporters of the HlyC/CorC family. Exfoliative toxins' (eta) sequences linked to a type I category of bacterial toxins were found (Table 1). Sequence orthologs were found in all products where product M1 showed the highest gene copy number and the great majority of sequence hits belonging to the genus *Staphylococcus* and *Tetragenococcus halophilus* (Table S3d). It is noteworthy that all sequence aligned returned gene functions annotated as hypothetical or permease-like proteins across all the products studied. While moist snuff products showed the highest abundance for most of the aforementioned toxins, dry snuff appeared to have much lower gene copies but a broader diversity of sequences encoding additional toxin genes (Table 1). An exception to this was found in product D1 where sequences for gene *fhaB*, which encodes for a filamentous hemagglutinin, showed the highest estimated gene copies in D1. Amino acid sequence alignments suggest that these genes are largely contributed by several species of *Pantoea* and a single *Ralstonia* species (Table S3c). Sequences for the toxin gene *hya*, which encodes a hyaluronoglucosaminidase, were found to be in significant numbers in only moist product M1 while in lower copies or absent in all other products.

Table 1 Bacterial toxins in STP. A specific set of genes closely related to functions associated with toxins generation was found, suggesting that microbial species in STPs potentially carry these genes and consequently the potential for toxin production. Numerical values represent the number of sequence hits per gene found in each smokeless tobacco product metagenome

Gene	Function	D1	D2	M1	M2
plcC	Membrane damaging toxin	38	13	0	0
NEU1	Toxins that damage the extracellular matrix	0	0	0	180
hya		7	33	2505	0
tlyC	Pore-forming toxins	11	5197	5876	9737
RAC1	Ras-related C3 botulinum toxin substrate 1	61	0	0	0
hlyC, *cyaC*, *rtxC*	Cytolysin-activating lysine-acyltransferase	19	0	0	0
hlyD, *cyaD*	Type I secretion system, alpha-Hemolysin/cyclolysin transporter	23	0	0	0
hlyB, *cyaB*	Type I secretion system alpha-hemolysin/cyclolysin transporter	23	0	0	0
tccC	Type III toxins: intracellular toxins	27	0	0	0
eta	Type I toxins: toxins that act from the cell surface, superantigens	482	110	8035	458
hlyIII	Type II toxins: membrane damaging toxins pore-forming toxins	339	6366	11,067	6280
tolC	alpha-Hemolysin/cyclolysin transporter	211	65	0	0
fhaB	adherence to host cells	1159	84	0	0

However, sequence alignment found orthologs annotated as hypothetical proteins from the bacteria *Tetragenococcus halophilus* (Table S3e).

Antibiotic resistance genetic elements Genetic elements with sequence homology to genes coding for antibiotic resistance mechanisms were present in all STPs. A greater number of orthologs appeared to populate dry snuff products suggesting high species diversity, while a higher number of sequence hits for specific genes cluster to moist products suggesting higher abundances in these products (Fig. 4). Additionally, sequence mapping suggests a total of five distinct mechanisms of antibiotic resistance to be present in these STPs.

Beta-lactam resistance The most abundant sequences in this category were identified as *penP* (K17836) and *blaZ* (K18766), which code for class A β-lactamases (Fig. 4). Sequence hits for *PenP* were found in dry snuff only, from species of *Oceanobacillus* species in D1, whereas D2 gene sequence contribution appeared to originate from species of *Lactobacillus* (Table S4a). Sequences for *blaZ* however were present in all STPs with the highest abundance in M1 all from *Staphylococcus* species (Table S4b).

Penicillin-binding proteins (PBP) within the β-lactam resistance showed a D-alanyl-D-alanine carboxypeptidase (K07258) (genes *dacA*, *dacC*, and *dacD*) also known as PBP 5/6 (Amanuma and Strominger 1984) as the most abundant in all products. These genes were observed in high sequence hits for products D2 and M1 specifically from *Lactobacillus* and *Staphylococcus* species (Table S4c). PBP gene *pbp2A* (K12555) was found in all samples with gene sequence contributions from *Lactobacillus pobuzihii* in product D1 and *Tetragenococcus* species in moist snuff product M1 (Table S4e). Additional PBP sequences were distributed and localized mainly to dry products with a few exceptions (K08724, K03693, K12552, K12553, K00687) found in all snuff products (Fig. 4). These genes were found to be contributed by *Tetragenococcus* (Tables S4h-i) and *Staphylococcus* species (Tables S4f-g). These taxa contributed with homologs of penicillin-binding proteins in all samples apart from gene penicillin-binding protein 1 or *pbpA* (K12552) where a serine/threonine kinase associated (PASTA) domain-containing protein was found to higher sequence hits (Table S4g).

Macrolides gene sequences were identified in both snuff products with more genes identified in dry products yet higher

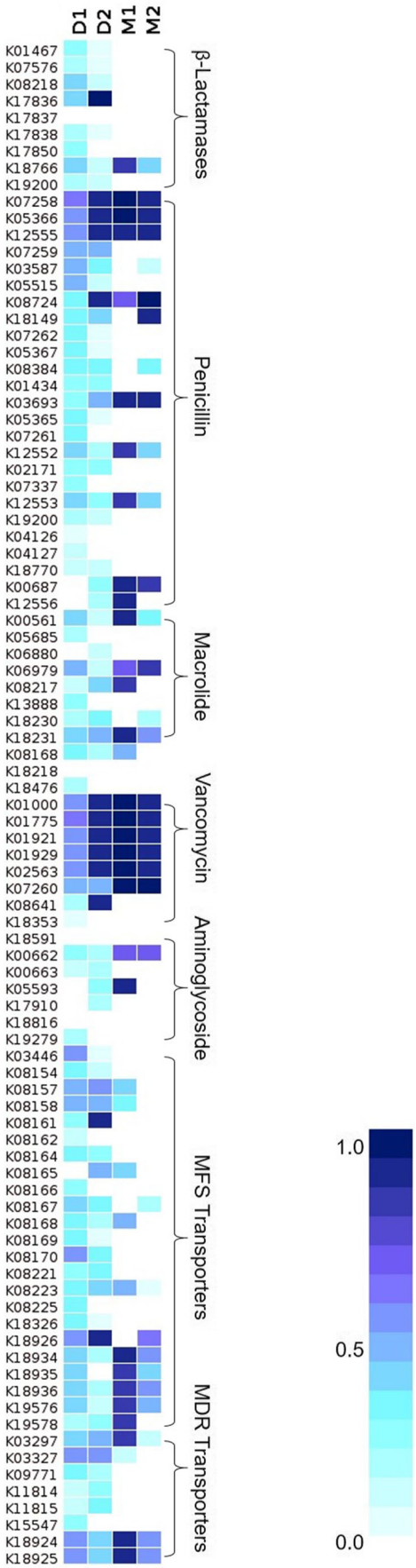

Fig. 4 Antibiotic resistance elements are present in STPs. Antibiotic resistance sequence hits observed in all STPs. The heatmap illustrates antibiotic resistance KEGG orthologs function categories with high number of sequence hits (dark blue) to the less abundant sequences (light blue) in each metagenome. Side labels represent groups of genes associated to specific antibiotic resistance mechanisms. Transporters and efflux pumps were divided in categories of major facilitator super family (MFS) and multidrug resistance (MDR)

number of sequence hits for those found in moist snuffs (Fig. 4). A 23S rRNA (adenine-N6)-dimethyltransferase coded by genes *ermA* and *ermC* (K00561) was found to be from relatives of *Staphylococcus* species in D1 and *Tetragenococcus hamophilus* in M1 (Table S5a). Macrolide transferase, *mph* genes (K06979), was found in all four STPs and associated to *Staphylococcus* species. However, only a small number of this sequences return at higher than 95% identity for this genus (Table S5b). Finally, a noteworthy gene in this group is *msrA* (K18231), an ATP-binding permease known to facilitate macrolide transport. This gene was observed in all STPs with notable higher abundances in sample M1, showing genetic contributions from *Tetragenococcus halophilus*. Other products showed sequence hits in markedly small numbers for *Lactobacillus* and *Staphylococcus species* (Table S5c).

Vancomycin resistance proteins linked to peptidoglycan biogenesis genes *mraY*, *alr*, *ddl*, *murF*, and *vanY* (K01000, K01775, K01921, K01929, K07260, respectively) were found as the highest sequence hits of all genes in the antibiotic resistance category (Fig. 4). Phospho-N-acetylmuramoyl-pentapeptide-transferase (*mraY*) was in all products but more abundant in D2 and M1, where the main sequence contributors were *L. pobuzihii* and *Staphylococcus spp.*, respectively (Table S6b). Similar sequence hit patterns and species associations were observed for alanine racemase (*alr*), and D-alanine-D-alanine ligase (*ddl*), a UDP-N-acetylmuramoyl-tripeptide–D-alanyl-D-alanine ligase (*murF*) and D-alanyl-D-alanine carboxypeptidase (*vanY*); however, *ddl* gene sequences in M1 showed abundance from *Tetragenococcus halophilus* (Table S6c-f). In contrast, D-alanyl-D-alanine dipeptidase (*vanX*) was mainly present in D2 with gene sequences associated to *Lactobacillus pobuzihii* (Table S6g).

Aminoglycoside resistance genes had low gene copy numbers in dry snuff products while in moist snuff, two orthologs were found in significant levels. These two orthologs were *aacC* (K00662) and *aadK* (K05593), for which the highest abundances were found for *aadK* (aminoglycoside 6-adenylyltransferase) in snuff product M1 (Table S7b). Protein sequence alignments returned species of genus *Tetragenococcus* as the main contributors of these resistance genes.

Multidrug transporters We found two types of transporters associated with antibiotic resistance. Gene sequences related to membrane transporters of the major facilitator super family (MFS) as well as multidrug resistance (MDR) transporter proteins were found in all products. A series of antiporters and pumps with the most abundant sequence orthologs in the MFS category was found to be *mdtG* (K08161), *nor* (K08170), *imr* (K18926 & K18934), and *mdeA* (K18936) (Fig. 4). These genes showed significant abundances with origins in *Staphylococcus* species for the majority of the STPs studied. However, two exceptions were observed, genes *mdtG* and lincomycin resistance gene *imrB* (K18926) for which *L. pobuzihii* was observed as the dominant contributing species in product D2 (Table S8a & S8c). Like MFS transporters, the MDR-associated species were most diverse in the dry snuff products (Fig. 4). Abundant genes with up to 8018 sequence hits per gene were observed almost exclusively in moist M1. Consequently, we looked only at those genes considered relevant based on estimated gene copy numbers for this category. We found a subset of orthologs linked to small multidrug resistance pumps, annotated as *ermE* (K03297), *ykkC* (K18924), and *ykkD* (K18925). These genes originated from species of *Staphylococcus*, in product M1, while the remaining products showed other much less abundant taxa (Table S9a-d). Additional sequences classified as multidrug resistance genes were identified with multiple annotations assigned to these orthologs (*norM*, *mdtK*, *dinF*, and SLC47A). These will be referred to as multi-antimicrobial extrusion protein (MATE) family of efflux transporters. They were observed in products D1 and D2 with protein sequence identities at 90% or higher. Though homologies suggested considerable bacterial diversity for which each taxa seemed to have low gene copy numbers, we underscore that for these particular orthologs, species provenance includes fungal species for which *Aspergillus* and *Alternaria* were found with gene copy numbers comparable to those of other bacterial species (Table S9b).

Potential for genetic transfer

Horizontal gene transfer (HGT) is likely a key occurrence in prokaryotic communities' evolution, as it allows the acquisition of new genes that can accelerate evolution and adaptation to new environments. To date, no study of smokeless tobacco has identified transfer of microbes from product to oral microbiota, but it remains highly plausible that such transfer could occur, even if only transiently. Given that genetic transfer occurs abundantly in biofilms such as in the oral cavity, genetic transfer from product to oral microbiome remains possible and warrants further investigation. In this study, we investigated potential mechanisms via HGT-categorized genetic content. Canonical genes associated with transference mechanisms were identified in all smokeless tobacco samples, more so in moist snuff than in its dry counterpart. A group of these genes, seemingly associated with integration and transposition mechanisms, had dominant sequence orthologs in these smokeless tobacco products, particularly protein sequences related to transposase proteins and derivatives. Integration-associated proteins, however, were significantly less abundance, though orthologs assigned in this function had a high number of sequence hits (Fig. 5). Bacteriophage replication and regulation genes, also present in these metagenomes, showed noticeable sequence abundance in moist snuffs, mainly in product M2 (Fig. 5). Among these genes, a high

◀ **Fig. 5** Mobile elements in STPs could be of concern. Horizontal gene transfer (HGT) mechanisms are presumably present in microbial species found on smokeless tobacco. Illustrated in a gradient of purple tones are distributions of sequence hits to clusters of orthologous group (COG) functions associated to HGT mechanisms. The great majority of sequence hits (dark purple) were observed for transposition, integration, and other phage-related gene elements. Heatmaps constructed using HemI illustrator 1.0 (Deng et al. 2014)

sequence abundance was observed for phage terminase or terminase-like proteins responsible for DNA recognition and initiation of DNA packaging in phage. A second group showing evident abundance was comprised of a series of antirepressor proteins and transcriptional regulatory genes primarily associated with replication and translation repression mechanisms (Fig. 5). Assembly and structural component-associated genes, though present, were not observed in high abundance. Nevertheless, key phage components including tail, capsid, and assembly proteins were identified in all STPs with a notable higher abundance in moist snuffs. Interestingly, these three components were represented by protein orthologs related to bacterial cell lysis mechanisms (COG4678, Muramidase) observed in phage (Fig. 5). Plasmids have been recognized as a vehicle of gene transfer between bacteria and species of bacteria (Ma et al. 2012). The presence of plasmid replication and regulatory genes in these metagenomes could suggest plasmids as an alternative for transferring genes. For instance, replication initiation gene sequences hits were found to reach over 200,000 sequence hits (Fig. 5). However, transfer-associated genes corresponding to mechanisms like conjugation were not identified in our data set.

Discussion

As tobacco is planted, grown, and harvested, it is reasonable to argue that a great deal of soil-to-plant microbial transfer and the establishment of endophytic communities occurs during these periods (Kandel et al. 2017). Furthermore, these communities can be altered by post-harvest processing, such as curing and fermentation, bringing about a series of metabolic activities with the potential to generate harmful chemical constituents (Di Giacomo et al. 2007; Zhao et al. 2007). Samples of moist and dry snuff tobacco products were collected and metagenomic analysis performed to describe microbial compositions concomitant to genetic function contributions from the microbial communities present.

Findings showed that representatives for all domains of life were present in smokeless tobacco products to variable relative abundances. In these metagenomes, dry snuff showed a higher diversity than the moist type (Fig. 1). From these major taxonomic groups, *Bacteria* ranked as the most abundant taxa in STPs for which three major phyla, *Firmicutes*,

Proteobacteria, and *Actinobacteria*, showed substantial abundances over any other phylum identified. Similar species abundance trends were observed in previous STP studies, and these three phyla are commonly found in several different rhizospheres for other crops and tree species (Bulgarelli et al. 2013; Han et al. 2016; Smyth et al. 2017; Tyx et al. 2016). This supports the concept that tobacco plants may sample from the adjacent soil environment in addition to carrying contaminants from the field and the manufacturing facilities.

Several important families in the *Firmicutes* group had the highest abundances in these metagenomes. Particularly, members associated to *Bacillaceae*, which we found relevant since the most distinguishing feature of the family is their ability to form endospores affording these species high resistance to heat, chemicals, and radiation, allowing them to survive hostile conditions for prolonged periods of time (Mandic-Mulec et al. 2015). Thus, we infer these taxa may be able to persist through the manufacturing process and storage of tobacco products (Di Giacomo et al. 2007; Han et al. 2016). Han and others reported *Bacillus pumilus* and *B. subtilis* to be among the dominant bacilli in moist STPs, underscoring their potential pathogenicity as a cause of concern. In STPs studied here, these *Bacillus* species were not found to be dominant. Instead, other more distant relatives within *Bacillaceae* were found, specifically in dry STPs (Fig. 1). These dominant taxa, although not known to be relevant human pathogens, were found as large contributors of genes associated with nitrate and nitrite transport and antibiotic resistance mechanisms (Tables S1-S5) (Lagier et al. 2015; Nithya and Babu 2017). Based on these findings, we surmise bacilli presence and richness may fluctuate for specific taxonomic groups along the manufacturing process in response to different environmental conditions along production.

The *Firmicutes* subcategory *Staphylococcaceae* bears potential harmful factors since some of its members are medically significant pathogens. However, the majority of the taxonomic relatives found in these STPs are prominent species in the manufacturing of consumables (i.e., food industry) for their flavoring capabilities and product-enhancing biochemistry (Caplice and Fitzgerald 1999; Talon et al. 1999). Some of these capabilities rest mainly in the reduction of nitrate which could have unintended consequences when considering nitrate reduction plays a central role in the generation of TSNAs during smokeless tobacco production. A noteworthy set of species including *Staphylococcus xylosus*, *S. nepalensis*, and *S. sciuri* was found to contribute several genes associated with toxins and antibiotic resistance. Although not assessed in this study, it is noteworthy that previous reports suggest that these *Staphylococci* have been associated with biofilm formation, endotoxin production, and are believed to be responsible for cases of endocarditis (Hollenberg 2017; Khan and Hollenberg 2017; Planchon et al. 2006). The contributions of *Staphylococcaceae* were dominant across the gene functions analyzed here, suggesting that these species not only persist but are dominant at the point of product sampling. We note that differentiation between genetic traits of the aforementioned *Staphylococci* is challenging and conclude that the genes identified in these metagenomes may apply to these species interchangeably (Novakova et al. 2006a; Novakova et al. 2006b). Thus, we considered all genetic contributions leading to potential harmful constituents to originate from *Staphylococcaceae* as a group rather than from any specific species.

Moderate additional contributions were found in the families *Carnobactericeae* and *Lactobacillaceae*. These two close relatives were comprised of a smaller number of species representatives in the metagenomes analyzed. *Lactobacillaceae*, for instance, rivaled *Staphylococcaceae* species in abundances particularly in product D2 yet the main representatives seemed to be limited mostly to *Lactobacillus* species (Tables S2-S8). Furthermore, *L. plantarum* and *L. pobuzihii* appeared to be the origin of genetic contributions associated with nitrate reduction and transport as well as antibiotic resistance elements examined in this study. For *Carnobactericeae*, only two representatives were found (Tables S1–S7), *Carnobacterium spp*. with limited observed contributions and *Atopococcus tabaci*, a novel species isolated from snus tobacco for which characterizations are not yet available (Cocolin et al. 2011). Interestingly, these species form part of the larger lactic acid bacteria (LAB) group for which metabolic potential is commonly use in fermented products manufacturing (Carr et al. 2002). Taking this into consideration could possibly expound their presence and abundance in these STPs. Additionally, LAB possess the ability to produce bacteriocins which aids to displace other bacteria and are known to be used for its potential to prevent pathogenicity (Parada et al. 2007). In contrast, their presence in the STPs examined here underscores the argument that they may also carry undesirable genetic traits since we have found sequences associated to antibiotic resistance and toxins (Tables S2–S8).

Species of *Proteobacteria* were found in lower abundances for all metagenomes with nominal sequence hits in dry snuff products (Fig. 1). Despite this abundance dearth, results showed a small number of species with potential genetic contributions relevant to the functions emphasized in this study. Interestingly, species of this genus particularly *P. vagans* or *P. agglomerans* stand as proposed biological controls against plant infections (i.e., fire blight) cause by relatives of the species. However, *P. agglomerans* is categorized as a BSL2 organism and has been identified in clinical reports as an opportunistic human pathogen (Rezzonico et al. 2009). Another relevant *Proteobacteria* was *Burkholderiaceae*, a family from which notable genetic contributions were identified as belonging to close relatives of *Ralstonia solanacearum*. The genus *Ralstonia* is becoming more prevalent in cases of nosocomial infections; however, the species within the genus reported as

opportunistic pathogens are relatives of those observed in this study and not the species per se (Ryan et al. 2014). Interestingly, genes associated with pathogenicity in *R. solanacearum* have been reported as identical to those of the well-known pathogen *Pseudomonas aeruginosa* (Meng et al. 2015).

Actinobacteria, the third most abundant bacterial group, showed species in STPs for which *Corynebactericeae* was the most prominent. This family was represented mainly by the genus *Corynebacterium* for which gene sequences associated to TSNA precursor generation, toxin, or toxin-like functions, and several antibiotic resistance mechanisms, were found distributed across all products.

Although *Corynebactericeae* species are known to be used in industrial applications like food production, other species within the family are significant pathogens of humans or domestic animals (Brennan et al. 2002; Schroder et al. 2011; Shi et al. 2013). For instance, *C. ammoniagenes*, a species observed in these STPs, was originally isolated from infant feces and believed to be the culprit for diaper rash. While it lacks virulence in animal studies and raised considerable doubts about its potential role as a pathogen, it is worth noting that *Corynebacteria* virulence factor homologs are widely distributed including non-pathogenic members of the genus such as *C. ammoniagenes* (Antunes et al. 2015; Coyle and Lipsky 1990). Pathogenicity potential notwithstanding, our findings suggest that additional harmful constituents may not depend on pathogenicity to present harm.

Di Giacomo and others found eukaryotic taxa represented by yeast species in cultures from cigar tobacco (Di Giacomo et al. 2007). They argue that these species may have metabolic roles at the beginning of the fermentation process that eventually is overtaken by bacteria in cigars manufacturing. In smokeless tobacco, there have been limited number of reports for fungal species; however, this could be mainly because those approaches have been centered in chemical analyses, 16S rDNA sequencing, or culture-based identification (Larsson et al. 2008; Saleem et al. 2018). Here, we have identified a series of fungal taxa whose genetic contributions were associated primarily to the families *Aspergillaceae* and *Pleosporaceae* as the dominant eukaryotic microbial taxa for dry snuff STPs. Within the former, sequences associated to nitrate/nitrite transport and oxidation of (anionic) alkyl nitronates were identified from the genus *Apergillus*. In the latter family, we identified a lower abundance of these genes associated to *Alternaria* species. The presence of these gene contributions underscores the possibility of an additional group of microbial species outside the bacterial domain that could also be responsible for the generation of TSNA precursors. *Aspergillus* species have been reported to produce harmful toxins and are associated to pathogenicity in immunocompromised individuals. Correspondingly, *Alternaria* species are powerful allergens with several proteins that elicit IgE immune responses and in some cases have resulted in lethal asthma cases (Gabriel et al. 2016). Pathogenicity aside, the *Aspergillus* species identified here have been reported as fermenters for consumable products (i.e., tea) and their characterization has yet to link them to infection (Ge et al. 2016). Conversely, *Alternaria* species present in comparable or slightly lower levels in dry snuff do present a potential for harm to the STP user. More importantly, considering our findings suggesting the presence of TSNA precursor associated metabolism from fungal species and previous reports of fungal activity early in the manufacturing of tobacco products, we can infer that this taxonomic groups may be active at earlier stages of STP production. Moreover, these species and related taxa may also persist throughout the manufacturing process.

In addition to eukaryotic species, we ascertained the presence of major taxonomic groups like *Archaea*. Although *Archaeans* are quite prevalent in inhospitable dwellings like hydrothermal vents, they are also found in more common niches like the gastrointestinal tract. As genomics advance, identification and incorporation of new taxa into *archaeal* databases have reveal not only their uniqueness but their ubiquity. While we observed archaeal sequences for dry and moist snuff, we did not observe a preponderance of sequences associated to *Archaea* nor any associated metabolic activity of interest. Thus, we can only report the presence of these taxa but cannot extrapolate other functions and roles in relation to STPs.

In the case of viruses, we identify a series of taxonomic subcategories that are present in the STPs analyzed. Not surprisingly, most of the viral taxa were bacteriophage classified as *Myovirideae*-, *Podovirideae*-, and *Siphovirideae*-related gene sequences. We speculate the contributions of these taxa come in the form of information exchange between species and suggest that undesirable potential genetic transference could occur in the mouth microbiome of STP users. Moreover, the possibility of microbial species introducing genetic elements into the gastrointestinal (GI) tract microbiome of the user through bacteriophage-mediated jumps seems more tangible. Although we do not offer evidence to support this, we are compelled to consider other clinical and environmental scenarios where this mechanism introduced impactful genetic factors. A classic example is seen in *Vibrio cholera* and cholera toxin genes, *Staphylococcus aureus* pathogenicity, and acquired antibiotic resistance genes (Faruque and Mekalanos 2012; Juhas 2015). In the context of these metagenomes, *Staphylococci*-associated phage sequences were found in considerable numbers. The abundance of these sequences does hint at the prospect of harm introduction to the user's microbiome via transduction. We are also compelled to consider that phage host-range limitations reduce the chances of many events occurring where undesirable genetic elements may be readily imported. However, we argue that although the possibility for transduction events is reduced, it is not eliminated.

The species diversity observed in smokeless tobacco products is not comparable to those observed in soil and similar environments, yet it is rich and diverse when compared to other taxonomic distributions found in commonly consumed victuals (Cocolin et al. 2011; Jung et al. 2011; Knight et al. 2015; Marsh et al. 2013; Meersman et al. 2013; Plengvidhya et al. 2007). Within this diversity, there are specific groups that appeared dominant in the final form of the smokeless tobacco product. Among these, *Firmicutes* have been consistently found as the most dominant group of bacterial species with genus like *Bacillus* and *Staphylococcus* as dominant members of the phyla (Han et al. 2016; Monika et al. 2020; Tyx et al. 2016). Bacilli species, well known for their persistence and resilience, have been reported as dominant in similar types of tobacco products. These reports found *Bacillus* as one of their main constituents and underscored the possibility of pathogenicity or harm from some these taxa (Al-Hebshi et al. 2017; Han et al. 2016). Here, we report similar findings for both moist and dry snuff products, where *Lactobacilli* appeared as one of the most abundant constituents in STPs with considerable genetic contributions. Thus, we argue that pathogens may have been out competed by more metabolically fit species and may present harm by a different mechanism. Additionally, we found fungal species as part of this microbiota that although not dominant, still ranked in high abundance for dry snuff products. Di Giacomo and others highlighted the possibility of important metabolic roles by fungal species early in the STP production process (Di Giacomo et al. 2007). We reason that this could be the case based on fungal sequences found at the time of sampling. Yet, the question remains as to whether STP community dynamics begins with fungal constituents and associated activities later succeeded by bacteria and their metabolism to form the final outlook found in this study.

The addition of nitrogen rich fertilizers to soils where tobacco crops are cultivated is a common practice (MacKown et al. 1999). Such abundance of nutrients plays an important role in crop yields as microbial metabolism linked to nitrogen cycling brings additional benefits to any given crop (Horwath et al. 2015). Consequently, an overabundance of NO_3^- and its reduction to NO_2^- further contributes with TSNA formation throughout the manufacturing process. Canonical genes involved in the transport and reduction of nitrate were present in all products' data sets. For instance, genes for the *nar* operon were all found in moderate number of sequence hits while transport-associated genes were substantially low (Fig. 2). All pathway components for the reduction of NO_3^- to NO_2^- and subsequent transport were present. Nevertheless, gene copy numbers suggest that these functions may not have been necessary in the microbial community at the time of sampling. Thus, we surmise that the low sequence hit numbers may reflect metabolic changes resulting from shifts in the community structure throughout the manufacturing process. The presence of genes from efficient denitrifying and other presumptive nitrate-reducing species may have been elevated mostly before or during the fermentation processes. In contrast, a significant number of sequence orthologs for an enzyme annotated as nitronate monooxygenase (K00459; *npd*, *ncd2*) were identified. Assigned functions are associated with the oxidation of nitroalkanes (i.e., nitromethane, 2-nitropropane, and nitrobenzene) in their ionic form (Fig. 2); however, only 2-nitropropane has been empirically proven as a substrate (Francis et al. 2005). These compounds have been reported in cigarette smoke as a result of high nitrate concentrations in tobacco plants (Hoffmann et al. 2001). Unfortunately, levels of these carcinogens in STPs remain unknown.

Bearing in mind the potential for nitroalkanes availability in the product, a post-production oxygen-rich environment, and the high sequence numbers associated to this function, we propose the possibility that this enzyme could play a role in the generation of nitrite at the aerobic stages of STP manufacturing. Conversely, we only found *npd/ncd2* genes sequence overabundance in one moist sample presumptively contributed only by *Staphylococcus cohnii* (Table S1d). Moreover, a recent study found large numbers of hypothetical proteins in the database misannotated as nitronate monooxygenase (NMO) yet characterized as a ubiquitous quinone reductase (Ball et al. 2016). These two points present a confounding outlook and suggest that some of our sequences may not reflect potential NMO activity despite the observed sequence homologies.

Because the data do not reflect high abundance of genes for TSNA precursors' generation, or the microbial species associated to them, we infer that the activity may have been there in earlier manufacturing steps (Law et al. 2016). Therefore, temporal research approaches including metagenomics and transcriptomics may provide a clearer picture on TSNA formation precursor's mechanisms and its association to microbial metabolic activities in STPs.

Toxins of microbial origin may play a role in health risk exposure to tobacco product users (Han et al. 2016; Pauly and Paszkiewicz 2011). Our sequence data sets do not reveal a preponderance of toxin genes or toxin synthesis elements. However, a small group of orthologs for a series of important membrane damaging toxins were found in all the STPs studied (Table 1). Toxin orthologs seemed to have function similarities to pore-forming or extracellular cell surface disruption, suggesting potential epithelial cell membranes disturbance. Hemolysins for example are considered virulence factors commonly observed in pathogenicity mechanisms where pores are formed in host cell membranes causing disruption and cellular death. Likewise, exfoliative toxins are at the center of pathogenicity (i.e., blistering skin disease) where a serine protease disrupt epithelial tissue rigidity and structure (Alouf and Muller-Alouf 2003). Previous studies reported

Bacillus pumillus, *B. subtilis*, and other closely related bacilli as dominant taxa in smokeless tobacco. The same study suggested that increasing risk of exposure to toxins could be attributed to these species (Han et al. 2016). Although the metagenomic data reported here showed a series of dominant bacilli, corresponding taxonomic makeup or the potential associated toxins were not observed (Table 1). However, abundant species related to *Staphylococcus*, *Lactobacillus*, and *Tetragenococcus* were identified as the potential source of toxins in the STPs studied here (Tables S3a-e).

In soil, microorganisms are constantly engaged in what could be best described as biochemical warfare (Czárán et al. 2002). It is well known that synthesis of antimicrobials occurs in soil dueling communities as strategies to displace competitors (Hibbing et al. 2010). Naturally, the development of defense mechanisms to these assaults is established as well (Walsh and Duffy 2013). A number of studies suggest that antimicrobial resistance (AMR) genes in clinical environments are found to be closely associated with those in environments like soils (D'Costa et al. 2007). Such ubiquity suggests that genes encoding these resistance mechanism components could be part of microbial genomes associated to tobacco plant and persist throughout tobacco product manufacturing.

We observed AMR-associated gene sequences in both dry and moist STPs with predominant function categories linked to antibiotic insensitivity or deactivation as in the case of the genes found for penicillin, macrolides, and vancomycin resistance. Most of these genes play roles in peptidoglycan biogenesis and cell wall assembly which makes them important for the microorganism and thus hold potential for transference to the endogenous microbiome of the STP user. We also observed transport-mediated antibiotic resistance in the case of genes classified under the major facilitator super family (MFS) and multidrug resistance (MDR) transporters. These transporters are known to be part of the largest classes of secondary active transporters and are widely expressed in multiple domains of life. They can transport small solutes because of chemiosmotic ion gradients (Pao et al. 1998). They are also capable of mediating the export of structurally unrelated drugs independent of accessory proteins or cytoplasmic components (Putman et al. 2000). Such genetic traits could provide an advantage to those microorganisms.

Although no relevant clinical pathogens were identified, *Staphylococci* species were found to be dominant and contributed with several AMR-associated genes. Reports suggest that the antibiotic resistance in species like *Staphylococcus aureus* could be due to a transposable element responsible for resistance against major classes of antibiotics including beta-lactams, macrolides, and aminoglycosides. However, the genes responsible for the maintenance of the transposable element are widely distributed within the *Staphylococcaceae* family, indicating frequent exchanges of these genes between *Staphylococcus* species (John et al. 2019).

Several different species' genomes appeared to have resistance mechanisms or at least part of them in their genetic code. Considering the ubiquity and variability of these mechanisms, it is reasonable to assume that in addition to the species identified, these genes are carried by several other species in STPs. Such postulate requires further study to determine frequency and prevalence across numerous smokeless tobacco product microbiotas.

A large array of genetic transfer mechanism-related functions was observed in significant numbers across all STPs with elevated distributions in moist snuff (Fig. 5). Among these elements, a series of components for transposition was identified in both dry and moist products. Although the importance of mobilizable genetic elements like transposons is well recognized, the contributions by bacteriophages are also significant (Salyers et al. 2004; Wellington et al. 2013). Our data showed sequences associated to structural and functional genes that code for phage components as well as replication and regulatory functions. Though purely hypothetical, this suggests the possibility of gene transfer events occurring in the smokeless tobacco consumer's GI tract microbiotas via transduction mechanisms. Interestingly, preferential gene transfer has been observed to be strongly correlated with gene functions where operational genes (e.g., housekeeping, metabolic, or antimicrobial resistance functions) are more favorably transferred over informational ones (e.g., transcription and translation) (Jain et al. 1999). Moreover, there is growing evidence that lateral gene transfer has played an integral role in the widespread distribution of antibiotic resistance gene determinants (ARGD), pathogenicity, gene clusters encoding biodegradative pathways, and other important functions in microbial communities across various environments (de la Cruz and Davies 2000; Jain et al. 2002). Bearing in mind that a smokeless tobacco product can be a unique environment in itself, it is important to consider the potential negative aspects that genetic transference may present in an STP consumption situation. For instance, as long-standing STP consumers inoculate their mouth with STP microbial species, the prospect of genetic material transference could increase. Although support for this is outside the scope of our study, this assertion is supported in other environments where studies have reported *Bacteria* and *Eukaryotes* obtaining genes through horizontal transfer (Jain et al. 1999; Le Roux and Blokesch 2018).

Bacterial dominance is marked by species that appeared to be mainly fermenters, explaining the observed similarities to those used in the production of select foods. Whether these species are introduced as starter cultures to promote fermentation or are naturally occurring from tobacco plants and adjacent terroir is not entirely clear at this point, and more research is needed to answer this question. Remarkably, the microbial constituency of the STPs studied here carried genes with relevant functions and the potential to introduce additional harm to the consumer. Thus, we underscore the need to

further investigate short- and long-term effects on oral health, like endogenous microbiota displacement, biofilm formations, GI tract dysbiosis, and other microbiome disruptions, that in combination with epidemiological studies of STP users including microbial infection frequencies, antibiotic resistance assurgency, toxin-associated effects prevalence, and cardiovascular disease. These aspects could reveal a clearer picture of all the potential harm the microbial component of smokeless tobacco holds.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00253-020-10999-w.

Author contributions AJR conceived and design research. AJR and LMK conducted experiments. AJR and RET conducted bioinformatics quality and organization, and analyzed data. CHW contributed new reagents and analytical tools, and SBS provided tobacco technical expertise. AJR wrote the manuscript. All authors read and approved the manuscript.

Funding Funding for this study was provided by the Centers for Disease Control and Prevention, National Center for Environmental Health, Division of Laboratory Sciences. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Compliance with ethical standards

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

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

Disclaimer The findings and conclusions in this report are those of the author(s) and do not necessarily represent the views of the Centers for Disease Control and Prevention. Use of trade names is for identification only and does not imply endorsement by the Centers for Disease Control and Prevention, the Public Health Service, or the US Department of Health and Human Services.

References

Al-Hebshi NN, Alharbi FA, Mahri M, Chen T (2017) Differences in the bacteriome of smokeless tobacco products with different oral carcinogenicity: compositional and predicted functional analysis. Genes (Basel) 8(4). https://doi.org/10.3390/genes8040106

Alouf JE, Muller-Alouf H (2003) *Staphylococcal* and *streptococcal* superantigens: molecular, biological and clinical aspects. Int J Med Microbiol 292(7-8):429–440. https://doi.org/10.1078/1438-4221-00232

Amanuma H, Strominger JL (1984) Purification and properties of penicillin-binding proteins 5 and 6 from the dacA mutant strain of Escherichia coli (JE 11191). J Biol Chem 259(2):1294–1298

Antunes CA, dos Santos LS, Hacker E, Köhler S, Bösl K, Ott L, de Luna MG, Hirata R Jr, de Carvalho Azevedo VA, Mattos-Guaraldi A-LJM (2015) Characterization of DIP0733, a multi-functional virulence factor of *Corynebacterium diphtheriae*. Microbiology 161(3): 639–647

Ball J, Salvi F, Gadda G (2016) Functional annotation of a presumed nitronate monoxygenase reveals a new class of NADH:Quinone reductases. J Biol Chem 291(40):21160–21170. https://doi.org/10.1074/jbc.M116.739151

Behzad H, Gojobori T, Mineta K (2015) Challenges and opportunities of airborne metagenomics. Genome Biol Evol 7(5):1216–1226. https://doi.org/10.1093/gbe/evv064

Brennan NM, Ward AC, Beresford TP, Fox PF, Goodfellow M, Cogan TM (2002) Biodiversity of the bacterial flora on the surface of a smear cheese. Appl Environ Microbiol 68(2):820–830

Bulgarelli D, Schlaeppi K, Spaepen S, Ver Loren van Themaat E, Schulze-Lefert P (2013) Structure and functions of the bacterial microbiota of plants. Annu Rev Plant Biol 64:807–838. https://doi.org/10.1146/annurev-arplant-050312-120106

Caplice E, Fitzgerald GF (1999) Food fermentations: role of microorganisms in food production and preservation. Int J Food Microbiol 50(1-2):131–149

Carr FJ, Chill D, Maida N (2002) The lactic acid bacteria: a literature survey. Crit Rev Microbiol 28(4):281–370. https://doi.org/10.1080/1040-840291046759

Chang HX, Haudenshield JS, Bowen CR, Hartman GL (2017) Metagenome-wide association study and machine learning prediction of bulk soil microbiome and crop productivity. Front Microbiol 8:519. https://doi.org/10.3389/fmicb.2017.00519

Chen IA, Markowitz VM, Chu K, Palaniappan K, Szeto E, Pillay M, Ratner A, Huang J, Andersen E, Huntemann M, Varghese N, Hadjithomas M, Tennessen K, Nielsen T, Ivanova NN, Kyrpides NC (2017) IMG/M: integrated genome and metagenome comparative data analysis system. Nucleic Acids Res 45(D1):D507–D516. https://doi.org/10.1093/nar/gkw929

Chen IA, Chu K, Palaniappan K, Pillay M, Ratner A, Huang J, Huntemann M, Varghese N, White JR, Seshadri R, Smirnova T, Kirton E, Jungbluth SP, Woyke T, Eloe-Fadrosh EA, Ivanova NN, Kyrpides NC (2019) IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. Nucleic Acids Res 47(D1):D666–D677. https://doi.org/10.1093/nar/gky901

Cocolin L, Dolci P, Rantsiou K (2011) Biodiversity and dynamics of meat fermentations: the contribution of molecular methods for a better comprehension of a complex ecosystem. Meat Sci 89(3): 296–302. https://doi.org/10.1016/j.meatsci.2011.04.011

Coyle MB, Lipsky BA (1990) Coryneform bacteria in infectious diseases: clinical and laboratory aspects. Clin Microbiol Rev 3(3):227–246

Critchley JA, Unal BJT (2003) Health effects associated with smokeless tobacco: a systematic review. Thorax 58(5):435–443

Czárán TL, Hoekstra RF, Pagie L (2002) Chemical warfare between microbes promotes biodiversity. Proc Natl Acad Sci 99(2):786–790

D'Costa VM, Griffiths E, Wright GD (2007) Expanding the soil antibiotic resistome: exploring environmental diversity. Curr Opin Microbiol 10(5):481–489

de la Cruz F, Davies J (2000) Horizontal gene transfer and the origin of species: lessons from bacteria. Trends Microbiol 8(3):128–133. https://doi.org/10.1016/s0966-842x(00)01703-0

Deng W, Wang Y, Liu Z, Cheng H, Xue Y (2014) HemI: a toolkit for illustrating heatmaps. PLoS One 9(11):e111988. https://doi.org/10.1371/journal.pone.0111988

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. https://doi.org/10.1128/AEM.02378-06

Faruque SM, Mekalanos JJ (2012) Phage-bacterial interactions in the evolution of toxigenic *Vibrio cholerae*. Virulence 3(7):556–565. https://doi.org/10.4161/viru.22351

Francis K, Russell B, Gadda G (2005) Involvement of a flavosemiquinone in the enzymatic oxidation of nitroalkanes catalyzed by 2-nitropropane dioxygenase. J Biol Chem 280(7):5195–5204. https://doi.org/10.1074/jbc.M411249200

Gabriel MF, Postigo I, Tomaz CT, Martinez J (2016) *Alternaria alternata* allergens: markers of exposure, phylogeny and risk of fungi-induced respiratory allergy. Environ Int 89-90:71–80. https://doi.org/10.1016/j.envint.2016.01.003

Gadda G, Francis K (2010) Nitronate monooxygenase, a model for anionic flavin semiquinone intermediates in oxidative catalysis. Arch Biochem Biophys 493(1):53–61

Ge Y, Wang Y, Liu Y, Tan Y, Ren X, Zhang X, Hyde KD, Liu Y, Liu Z (2016) Comparative genomic and transcriptomic analyses of the Fuzhuan brick tea-fermentation fungus *Aspergillus cristatus*. BMC Genomics 17:428. https://doi.org/10.1186/s12864-016-2637-y

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. https://doi.org/10.1128/AEM.01612-16

Hibbing ME, Fuqua C, Parsek MR, Peterson SB (2010) Bacterial competition: surviving and thriving in the microbial jungle. Nat Rev Microbiol 8(1):15–25

Hoffmann D, Hoffmann IJS, Monograph TC (2001) The changing cigarette: chemical studies and bioassays Smoking and Tobacco Control. vol 13, U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES Public Health Service National Institutes of Health National Cancer Institute, pp 159-192

Hollenberg S (2017) Valvular heart disease in adults: etiologies, classification, and diagnosis. FP Essentials 457:11–16

Horwath W, Paul E, Biochemistry (2015) Carbon cycling: the dynamics and formation of organic matter. 4th, Elsevier Inc

Huntemann M, Ivanova NN, Mavromatis K, Tripp HJ, Paez-Espino D, Palaniappan K, Szeto E, Pillay M, Chen IM, Pati A, Nielsen T, Markowitz VM, Kyrpides NC (2015) The standard operating procedure of the DOE-JGI Microbial Genome Annotation Pipeline (MGAP v.4). Stand Genomic Sci 10:86. https://doi.org/10.1186/s40793-015-0077-y

Jain R, Rivera MC, Lake JA (1999) Horizontal gene transfer among genomes: the complexity hypothesis. Proc Natl Acad Sci 96(7):3801–3806

Jain R, Rivera MC, Moore JE, Lake JA (2002) Horizontal gene transfer in microbial genome evolution. Theor Popul Biol 61(4):489–495. https://doi.org/10.1006/tpbi.2002.1596

John J, George S, Nori SRC, Nelson-Sathi S (2019) Phylogenomic analysis reveals the evolutionary route of resistant genes in *Staphylococcus aureus*. Genome Biol Evol 11(10):2917–2926. https://doi.org/10.1093/gbe/evz213

Joshi NA, Fass JN (2011) Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software]

Juhas M (2015) Horizontal gene transfer in human pathogens. Crit Rev Microbiol 41(1):101–108. https://doi.org/10.3109/1040841X.2013.804031

Jung JY, Lee SH, Kim JM, Park MS, Bae JW, Hahn Y, Madsen EL, Jeon CO (2011) Metagenomic analysis of kimchi, a traditional Korean fermented food. Appl Environ Microbiol 77(7):2264–2274. https://doi.org/10.1128/AEM.02157-10

Kandel SL, Joubert PM, Doty SL (2017) Bacterial endophyte colonization and distribution within plants. Microorganisms 5(4):77. https://doi.org/10.3390/microorganisms5040077

Khan ZA, Hollenberg SM (2017) Valvular heart disease in adults: infective endocarditis. FP Essent 457:30–38

Knight S, Klaere S, Fedrizzi B, Goddard MR (2015) Regional microbial signatures positively correlate with differential wine phenotypes: evidence for a microbial aspect to terroir. Sci Rep 5:14233. https://doi.org/10.1038/srep14233

Lagier JC, Khelaifia S, Azhar EI, Croce O, Bibi F, Jiman-Fatani AA, Yasir M, Helaby HB, Robert C, Fournier PE, Raoult D (2015) Genome sequence of *Oceanobacillus picturae* strain S1, an halophilic bacterium first isolated in human gut. Stand Genomic Sci 10:91. https://doi.org/10.1186/s40793-015-0081-2

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. https://doi.org/10.1186/1617-9625-4-4

Law AD, Fisher C, Jack A, Moe LA (2016) Tobacco, microbes, and carcinogens: correlation between tobacco cure conditions, tobacco-specific nitrosamine content, and cured leaf microbial community. Microb Ecol 72(1):120–129. https://doi.org/10.1007/s00248-016-0754-4

Le Roux F, Blokesch M (2018) Eco-evolutionary dynamics linked to horizontal gene transfer in *Vibrios*. Annu Rev Microbiol 72:89–110. https://doi.org/10.1146/annurev-micro-090817-062148

Ma Y, Paulsen IT, Palenik B (2012) Analysis of two marine metagenomes reveals the diversity of plasmids in oceanic environments. Environ Microbiol 14(2):453–466. https://doi.org/10.1111/j.1462-2920.2011.02633.x

MacKown CT, Crafts-Brandner SJ, Sutton TG (1999) Relationships among soil nitrate, leaf nitrate, and leaf yield of burley tobacco. Agron J 91:613–621

Mandic-Mulec I, Stefanic P, van Elsas JD (2015) Ecology of *Bacillaceae*. Microbiol Spectr 3(2) TBS-0017-2013. https://doi.org/10.1128/microbiolspec.TBS-0017-2013

Marsh AJ, O'Sullivan O, Hill C, Ross RP, Cotter PD (2013) Sequencing-based analysis of the bacterial and fungal composition of kefir grains and milks from multiple sources. PLoS One 8(7):e69371. https://doi.org/10.1371/journal.pone.0069371

Meersman E, Steensels J, Mathawan M, Wittocx PJ, Saels V, Struyf N, Bernaert H, Vrancken G, Verstrepen KJ (2013) Detailed analysis of the microbial population in Malaysian spontaneous cocoa pulp fermentations reveals a core and variable microbiota. PLoS One 8(12):e81559. https://doi.org/10.1371/journal.pone.0081559

Meng F, Babujee L, Jacobs JM, Allen C (2015) Comparative transcriptome analysis reveals cool virulence factors of *Ralstonia solanacearum* race 3 biovar 2. PLoS One 10(10):e0139090. https://doi.org/10.1371/journal.pone.0139090

Monika S, Dineshkumar T, Priyadharini S, Niveditha T, Sk P, Rajkumar K (2020) Smokeless tobacco products (STPs) harbour bacterial populations with potential for oral carcinogenicity. Asian Pac J Cancer Prev 21(3):815–824. https://doi.org/10.31557/APJCP.2020.21.3.815

Nithya A, Babu S (2017) Prevalence of plant beneficial and human pathogenic bacteria isolated from salad vegetables in India. BMC Microbiol 17(1):64. https://doi.org/10.1186/s12866-017-0974-x

Novakova D, Pantucek R, Petras P, Koukalova D, Sedlacek I (2006a) Occurance of *Staphylococcus nepalensis* strains in different sources including human clinical material. FEMS Microbiol Lett 263(2):163–168. https://doi.org/10.1111/j.1574-6968.2006.00408.x

Novakova D, Sedlacek I, Pantucek R, Stetina V, Svec P, Petras P (2006b) *Staphylococcus equorum* and *Staphylococcus succinus* isolated from human clinical specimens. J Med Microbiol 55(Pt 5):523–528. https://doi.org/10.1099/jmm.0.46246-0

Nurk S, Meleshko D, Korobeynikov A, Pevzner PA (2017) metaSPAdes: a new versatile metagenomic assembler. Genome Res 27(5):824–834. https://doi.org/10.1101/gr.213959.116

Pao SS, Paulsen IT, Saier MH Jr (1998) Major facilitator superfamily. Microbiol Mol Biol Rev 62(1):1–34

Parada JL, Caron CR, Medeiros ABP, Soccol CR (2007) Bacteriocins from lactic acid bacteria: purification, properties and use as biopreservatives. Braz Arch Biol Technol 50:512–542

Pauly JL, Paszkiewicz G (2011) Cigarette smoke, bacteria, mold, microbial toxins, and chronic lung inflammation. J Oncol 2011:13

Planchon S, Gaillard-Martinie B, Dordet-Frisoni E, Bellon-Fontaine MN, Leroy S, Labadie J, Hebraud M, Talon R (2006) Formation of biofilm by *Staphylococcus xylosus*. Int J Food Microbiol 109(1-2):88–96. https://doi.org/10.1016/j.ijfoodmicro.2006.01.016

Plengvidhya V, Breidt F Jr, Lu Z, Fleming HP (2007) DNA fingerprinting of lactic acid bacteria in sauerkraut fermentations. Appl Environ Microbiol 73(23):7697–7702. https://doi.org/10.1128/AEM.01342-07

Putman M, van Veen HW, Konings WN (2000) Molecular properties of bacterial multidrug transporters. Microbiol Mol Biol Rev 64(4):672–693. https://doi.org/10.1128/mmbr.64.4.672-693.2000

Rezzonico F, Smits TH, Montesinos E, Frey JE, Duffy B (2009) Genotypic comparison of *Pantoea agglomerans* plant and clinical strains. BMC Microbiol 9:204. https://doi.org/10.1186/1471-2180-9-204

Ryan MP, Adley CC, diseases i (2014) *Ralstonia spp*.: emerging global opportunistic pathogens. Eur J Clin Microbiol Infect Dis 33(3):291–304

Saleem S, Naz SA, Shafique M, Jabeen N, Ahsan SW (2018) Fungal contamination in smokeless tobacco products traditionally consumed in Pakistan. JPMA J Pakistan Med Assoc 68(10):1471–1477

Salyers AA, Gupta A, Wang Y (2004) Human intestinal bacteria as reservoirs for antibiotic resistance genes. Trends Microbiol 12(9):412–416

Sami A, Stanton C, Ross P, Ryan T, Elimairi I (2020) Ultra- structure of Toombak; smokeless tobacco of Sudan and its effects on oral and systemic health. Access Microbiol 2(7A):836. https://doi.org/10.1099/acmi.ac2020.po0722

Schroder J, Maus I, Trost E, Tauch A (2011) Complete genome sequence of *Corynebacterium variabile* DSM 44702 isolated from the surface of smear-ripened cheeses and insights into cheese ripening and flavor generation. BMC Genomics 12:545. https://doi.org/10.1186/1471-2164-12-545

Shi H, Wang R, Bush LP, Zhou J, Yang H, Fannin N, Bai R (2013) Changes in TSNA contents during tobacco storage and the effect of temperature and nitrate level on TSNA formation. J Agric Food Chem 61(47):11588–11594. https://doi.org/10.1021/jf404813m

Smyth EM, Kulkarni P, Claye E, Stanfill S, Tyx R, Maddox C, Mongodin EF, Sapkota AR (2017) Smokeless tobacco products harbor diverse bacterial microbiota that differ across products and brands. Appl Microbiol Biotechnol 101(13):5391–5403. https://doi.org/10.1007/s00253-017-8282-9

Talon R, Walter D, Chartier S, Barriere C, Montel MC (1999) Effect of nitrate and incubation conditions on the production of catalase and nitrate reductase by *staphylococci*. Int J Food Microbiol 52(1-2):47–56

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. https://doi.org/10.1371/journal.pone.0146939

Uyaguari-Diaz MI, Chan M, Chaban BL, Croxen MA, Finke JF, Hill JE, Peabody MA, Van Rossum T, Suttle CA, Brinkman FS, Isaac-Renton J, Prystajecky NA, Tang P (2016) A comprehensive method for amplicon-based and metagenomic characterization of *viruses*, *bacteria*, and *eukaryotes* in freshwater samples. Microbiome 4(1):20. https://doi.org/10.1186/s40168-016-0166-1

Walsh F, Duffy B (2013) The culturable soil antibiotic resistome: a community of multi-drug resistant bacteria. PLoS One 8(6):e65567

Wellington EM, Boxall AB, Cross P, Feil EJ, Gaze WH, Hawkey PM, Johnson-Rollings AS, Jones DL, Lee NM, Otten WJ, Thomas CM, Williamns AP (2013) The role of the natural environment in the emergence of antibiotic resistance in Gram-negative bacteria. Infect Dis Ther 13(2):155–165

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. https://doi.org/10.1007/s00253-006-0625-x