



The Effects of Smoking on Human Pharynx Microbiota Composition and Stability

 Lydia Bach,^a Asha Ram,^a  Umer Z. Ijaz,^b  Thomas J. Evans,^c  Daniel T. Haydon,^a  Jan Lindström^a

^aSchool of Biodiversity, One Health and Veterinary Medicine, University of Glasgow, United Kingdom

^bSchool of Science and Engineering, University of Glasgow, United Kingdom

^cSchool of Infection and Immunity, Glasgow Biomedical Research Centre, University of Glasgow, United Kingdom

ABSTRACT The oral microbiota is essential to the health of the host, yet little is known about how it responds to disturbances. We examined the oropharyngeal microbiota of 30 individuals over 40 weeks. As the oropharynx is an important gateway to pathogens, and as smoking is associated with increased incidence and severity of respiratory infections, we compared the microbiota of smokers and nonsmokers to shed light on its potential for facilitating infections. We hypothesized that decreased species diversity, decreased community stability, or increased differences in community structure could facilitate invading pathogens. We found that smoking is associated with reduced alpha diversity, greater differences in community structure, and increased environmental filtering. The effects of short-term perturbations (antibiotic use and participants exhibiting cold symptoms) were also investigated. Antibiotic use had a negative effect on alpha diversity, irrespective of smoking status, and both antibiotic use and cold symptoms were associated with highly unique bacterial communities. A stability analysis of models built from the data indicated that there were no differences in local or global stability in the microbial communities of smokers, compared to nonsmokers, and that their microbiota are equally resistant to species invasions. Results from these models suggest that smoker microbiota are perturbed but characterized by alternative stable states that are as stable and invasion-resistant as are the microbiota of nonsmokers. Smoking is unlikely to increase the risk of infectious disease through the altered composition and ecological function of the microbiota; this is more likely due to the effects of smoking on the local and systemic immune system.

IMPORTANCE Smoking is associated with an increased risk of respiratory infections. Hypothetically, the altered community diversity of smokers' pharyngeal microbiota, together with changes in their ecological stability properties, could facilitate their invasion by pathogens. To address this question, we analyzed longitudinal microbiota data of baseline healthy individuals who were either smokers or nonsmokers. While the results indicate reduced biodiversity and increased species turnover in the smokers' pharyngeal microbiota, their ecological stability properties were not different from those of the microbiota of nonsmokers, implying, in ecological terms, that the smokers' microbial communities are not less resistant to invasions. Therefore, the study suggests that the increased propensity of respiratory infections that is seen in smokers is more likely associated with changes in the local and systemic immune system than with ecological changes in the microbial communities.

KEYWORDS microbiota, smoking, community stability, perturbation, invasion, cold, antibiotics, microbiome, oropharynx, stability

The oral microbiota comprises a large number of species that are found in high densities, occurring in several distinct habitats, including the teeth, tongue, buccal mucosa, and oropharynx (1). It is well-established that a stable microbiota contributes to a healthy status for the host (2). Correspondingly, changes in community composition and abundance within these microbial communities are associated with the pathogenesis of several

Editor Paul A. Jensen, University of Michigan-Ann Arbor

Copyright © 2023 Bach et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Lydia Bach, lydia.bach@glasgow.ac.uk, or Jan Lindström, jan.lindstrom@glasgow.ac.uk.

The authors declare no conflict of interest.

Received 5 November 2021

Accepted 16 January 2023

Published 14 February 2023

diseases (3). These changes can be further linked with environmental stimuli that affect colonization and growth, specifically, press perturbations that stay in place for a long time and pulse perturbations that tend to be of short duration (4).

Smoking is an example of a press perturbation that has been shown to change the oral microbial community and promote pathogenetic species (5–8). Indeed, a major health outcome of smoking is the increased propensity for infections (9). There are three possible pathways of how smoking-related changes in the oropharyngeal microbiota could be associated with infection risk: changes in microbiota merely reflect a changed immune response in smokers, microbiota play an active role in deferring or facilitating invasions of pathogens, or both of these processes are involved. It is well-established that the harmful components in smoke mediate the host immune response, affecting, in particular, humoral and cell-mediated responses, which consequently affects the microbial community composition (10–13). Moreover, smoking has also been shown to affect many cellular processes, including neutrophil chemotaxis, adherence, phagocytosis, and function (14–17). It has been hypothesized that the main cause of the increased infection risk is associated with a compromised anti-bacterial function of leukocytes, including neutrophils, monocytes, T cells, and B cells (18), as well as pathogen enrichment (8, 19). In addition, the toxic substances contained in cigarette smoke, as well as oxygen deprivation, have been hypothesized to contribute to the altered microbiota in smokers (6, 20). Observational and interventional studies have found that the composition of the intestinal microbiota is also altered in smokers, with some taxa occurring in higher abundances (e.g., *Parvimonas*, *Fusobacterium*, *Campylobacter*, *Bacteroides*, and *Treponema*), while others have been shown to decrease in abundance (e.g., *Veillonella*, *Neisseria*, and *Streptococcus*) (7, 8, 20). However, there are also indications that the increased infection risk in smokers could be more directly linked to the changes in the oropharyngeal microbiota, as the changed community might be more prone to invasions by pathogens (9, 21, 22). It remains to be determined how smoking affects the stability and invasibility of these communities in ecological terms.

In addition to the potential effects of long-term perturbations, such as smoking, short-term perturbations can affect the oral microbiota. Examples of pulse perturbations in microbiota include short-term antibiotic use or the common cold. Much of the work investigating the impact of antibiotics on microbial communities has been conducted in the gut, with findings relating to the oral microbiota remaining ambiguous (23, 24). Those focusing on the oral microbiota showed that antibiotics can cause shifts in the microbial community composition and reduce microbial diversity (25, 26). These effects are short-lived, which the authors suggested may be associated with a greater resilience of oral, compared to gut, communities (24, 27).

Viral infections, such as the common cold, are gaining interest in the microbiota literature (28–30), but few studies have investigated their impact on the human oropharyngeal microbiota (31). Evidence suggests that infection with Rhinovirus, one of the viruses that is responsible for the common cold, also results in increases in *Neisseria* and *Haemophilus* in the nasopharynx, which possibly leads to secondary bacterial infections (32).

While evidence supports the individual impacts of these perturbations on the oral microbiota, the purpose of the present investigation is to examine the combined effects of smoking, antibiotic use, and the acquisition of a common cold on the microbial community in terms of community composition, alpha diversity, beta diversity, and environmental filtering (i.e., abiotic factors preventing species from persisting in a location) (33). Going beyond those traditional measures, we also investigated stability (local and global stability) and invasibility in smokers and nonsmokers, providing us with a deeper understanding of the microbial community dynamics that are associated with pulse and press perturbations.

The relationship between diversity and ecological stability has been a long-standing question in ecology (34–36). Some studies suggest that more diverse and complex communities tend to be more stable and are thus generally more resistant to perturbations (37). Species removals or additions (e.g., through invasions), can result in major shifts in community structure and dynamics, which can thereby result in changes in community stability and function. In the context of the microbiota, this stability is

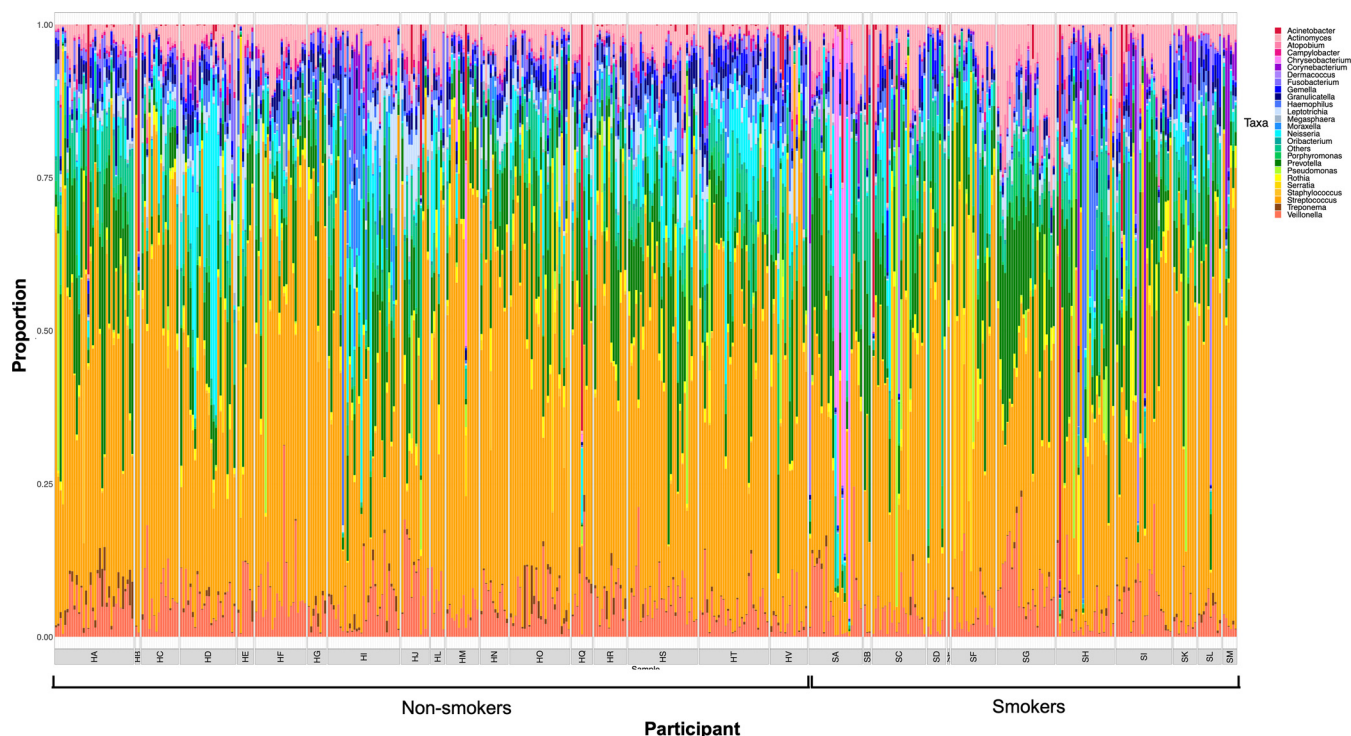


FIG 1 Relative abundance (%) of the 25 dominant genera for smokers and nonsmokers for each week over the sampling period.

critical for the well-being of the host, as it ensures that beneficial taxa persist over long periods, providing vital ecosystem functions (38–41).

Here, we test four specific perturbation-related hypotheses. We hypothesize that (i) smoking is equivalent to an ecological perturbation that is likely to decrease alpha diversity and cause greater differences in community structure and filtering; (ii) the community composition of the smoker microbiota will be different from that of nonsmokers; (iii) thereby, microbial communities in smokers will be less stable (locally and/or globally); (iv) meaning that the microbial communities in smokers will be more invasible, compared to those of nonsmokers.

RESULTS

The microbial communities were comprised of 9 phyla, 202 genera, and 1,438 assignments at the operational taxonomic unit (OTU) level. At the phylum level, 99.7% of reads were taxonomically classified, with the remaining 0.3% belonging to unknown or unclassified bacteria. Firmicutes, Bacteroidetes, and Proteobacteria dominated both the smoker and nonsmoker communities.

The relative abundances of the 25 most abundant genera were similar between time points for each participant, and they were broadly comparable across participants, irrespective of smoking status (Fig. 1). The three most dominant genera in healthy smokers and nonsmokers were *Streptococcus* (mean: 0.470, SD: 0.011), *Prevotella* (mean: 0.094, SD: 0.004), and *Veillonella* (mean: 0.058, SD: 0.002).

The most abundant OTUs in healthy smokers belonged to *Streptococcus* species, reflecting the general abundance of Firmicutes. These included *Streptococcus mitis*, *Streptococcus salivarius*, and *Streptococcus parasanguinis*, which were also the most dominant OTUs in the healthy samples from nonsmoking participants. Further details on the microbial composition of nonsmokers can be found in our previous publication (42). For plots showing the temporal changes in the community composition for all participants over the course of the study period at the phylum, family, and OTU levels, refer to Fig. S2, Fig. S3, and Fig. S4, respectively.

Alpha diversity was significantly higher in healthy nonsmokers, compared to healthy smokers (Fig. 2). Alpha diversity was significantly lower in nonsmokers exhibiting cold

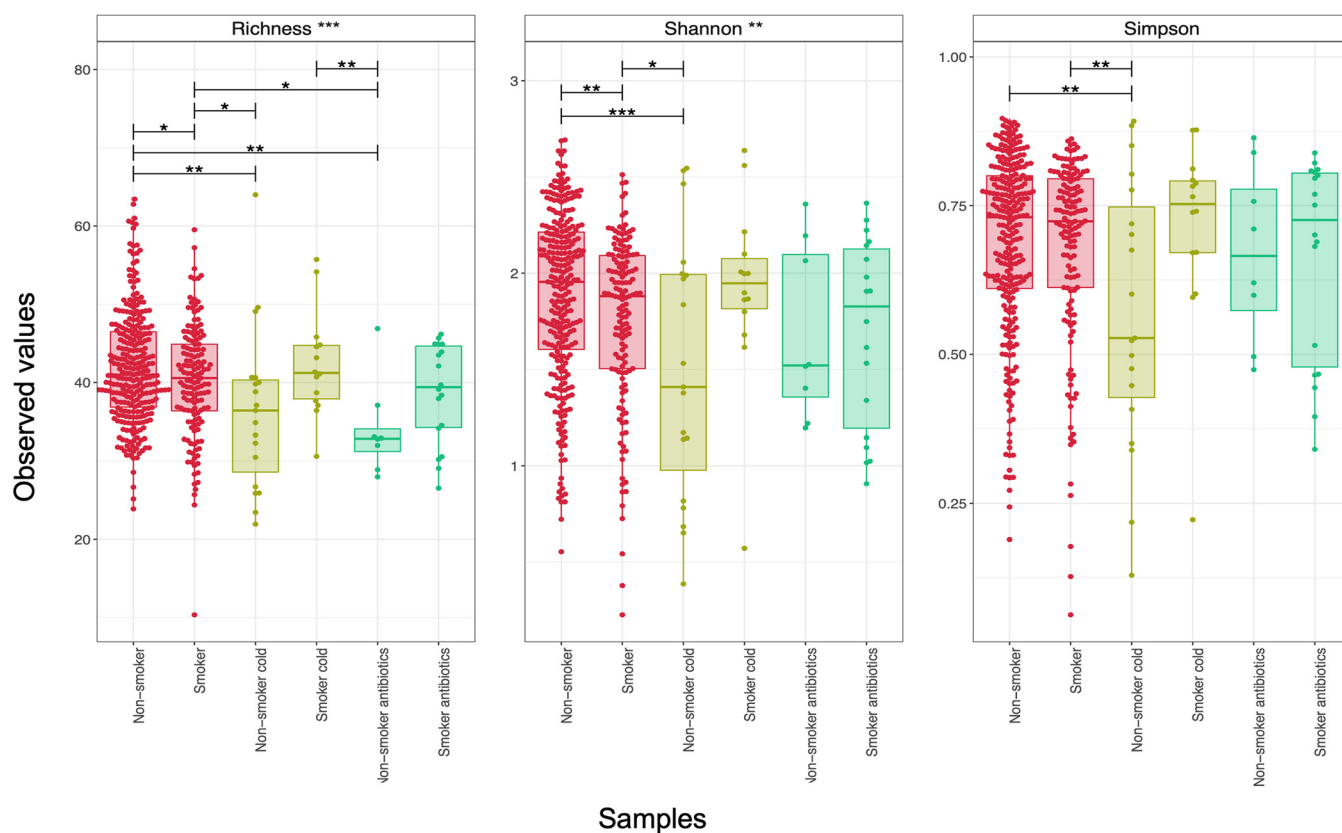


FIG 2 Three alpha diversity measures: species (OTU) richness, Shannon diversity, and Simpson diversity, calculated based on the oropharyngeal microbiota of smokers and nonsmokers when healthy, expressing cold symptoms, or taking antibiotics. The lines connect two categories if the values are significant per an ANOVA, with one within subject factor accounting for the participant IDs.

symptoms, compared to healthy nonsmokers. However, there were no differences in alpha diversity between healthy smokers and smokers experiencing cold symptoms, suggesting that cold had little or no effect on alpha diversity in smokers. Antibiotic use resulted in a decreased diversity in nonsmokers and smokers (with the exception of Simpson's index), but the use of antibiotics resulted in a greater similarity in alpha diversity measures between smokers and nonsmokers.

A beta dispersion analysis suggested that data homogeneity differed significantly with all factors (smoking status: smoker/nonsmoker, sex, age, and participant), with individual participants of the study being the single most important factor explaining the data variability for Bray Curtis (R^2 : 0.131%), unifracs (R^2 : 0.159), and weighted unifracs (R^2 : 0.132%), and this was followed by health status (R^2 between 0.054 and 0.065 for Bray-Curtis, unifracs, and weighted unifracs), whereas all of the other factors contributed significantly less (Table 1).

The local contribution to beta diversity (LCBD) is a measure providing insight into the ecological uniqueness (i.e., unusual species combinations) of a sample. Based on the abundance count (Bray Curtis) (Fig. 3), it was evident that nonsmokers with cold symptoms had significantly higher degrees of ecological uniqueness, compared to smokers and smokers with colds. When considering phylogenetic distances only (unifracs) (Fig. 3), nonsmokers on antibiotic treatments had particularly high LCBD values, which are indicative of relatively high degrees of ecological uniqueness. Using both abundances and phylogenetic distances, this pattern disappeared (weighted unifracs) (Fig. 3), but nonsmokers experiencing cold symptoms generally showed the greatest variability in LCBD.

We subsequently used a subset analysis (BVSTEP routine) to determine which OTUs were responsible for the changes in beta diversity, comparing samples between smokers and nonsmokers in three groups: healthy participants, participants exhibiting cold symptoms, and participants on antibiotics. To do so, we collapsed the abundance table to the minimum set

TABLE 1 Statistics for beta dispersion comparison, based on Bray-Curtis, unfrac, and weighted unfrac, based on a PERMANOVA analysis

Measure	Factor	Degrees of freedom	Sum of squares	Mean square	F	R ²	Pr(>F)
Bray-Curtis	Health status	5	7.98	1.60	6.09	0.054	0.001
	Sex	1	0.82	0.82	3.11	0.006	0.002
	Age	1	1.57	1.57	5.96	0.011	0.001
	Participant	28	19.32	0.69	2.63	0.131	0.001
	Residuals	447	117.28	0.26	0.80		
	Total	482	146.96	1.00			
Unfrac	Health status	5	2.99	0.60	7.75	0.065	0.001
	Sex	1	0.35	0.35	4.49	0.008	0.001
	Age	1	0.89	0.89	11.52	0.019	0.001
	Participant	28	7.35	0.26	3.40	0.159	0.001
	Residuals	447	34.53	0.08	0.75		
	Total	482	46.11	1.00			
Weighted unfrac	Health status	5	0.81	0.16	6.40	0.055	0.001
	Sex	1	0.09	0.09	3.57	0.006	0.006
	Age	1	0.47	0.47	18.76	0.033	0.001
	Participant	28	1.91	0.07	2.71	0.132	0.001
	Residuals	447	11.24	0.03	0.77		
	Total	482	14.52	1.00			

of OTUs explaining differences in community composition between groups. This resulted in a reduced-order abundance table that correlated highly with the full OTU table (Supporting information Table S2; Fig. S1). There were 24 OTUs associated with differences in community composition between healthy smokers and nonsmokers, 15 OTUs associated with nonsmokers and smokers experiencing cold symptoms, and 17 OTUs driving differences in nonsmokers and smokers on antibiotics. These three groups shared seven OTUs belonging to different genera. For further information on the relationships between features, the microbial communities (based on abundance tables), and the sources of variation (age, years smoking, and cigarettes smoked/week), refer to Fig. S6.

A heat tree analysis revealed OTUs in significantly different abundances across all samples, based on smokers, nonsmokers, smokers and nonsmokers expressing common cold

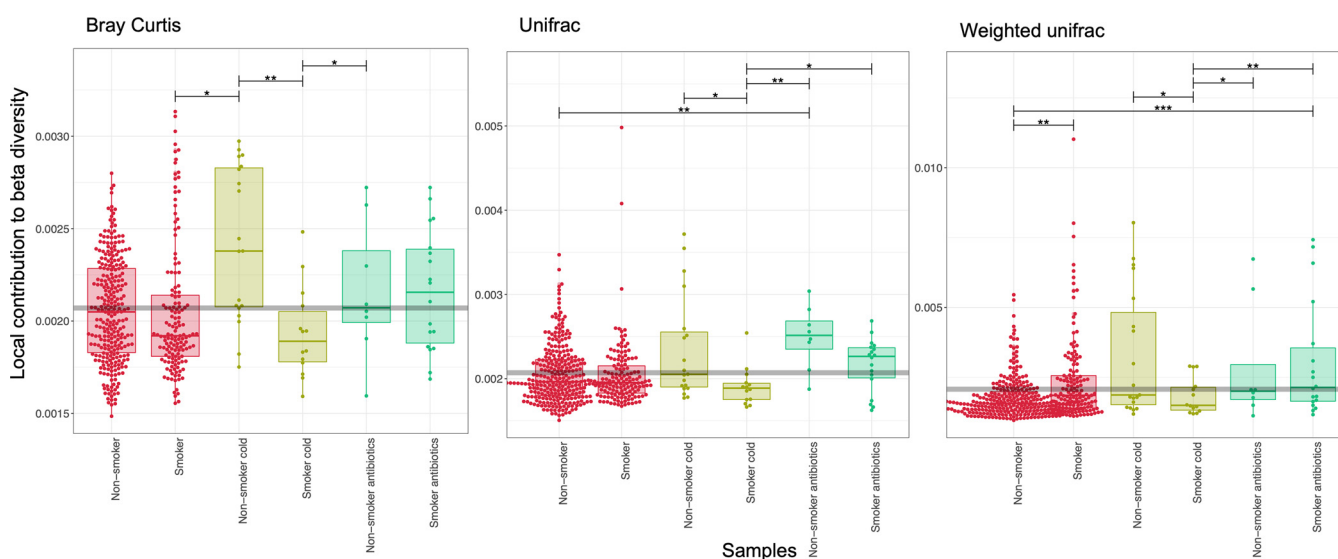


FIG 3 Beta diversity measures using Bray-Curtis (counts), unweighted unfrac (phylogenetic distance), and weighted unfrac (phylogenetic distance weighted by abundance counts) of the oropharyngeal microbiota of nonsmokers and smokers when healthy, expressing common cold symptoms, or taking antibiotics. Note the difference in scale. Means are given as horizontal gray lines. The lines connect two categories if the values are significant per an ANOVA, with one within subject factor accounting for the participant IDs.

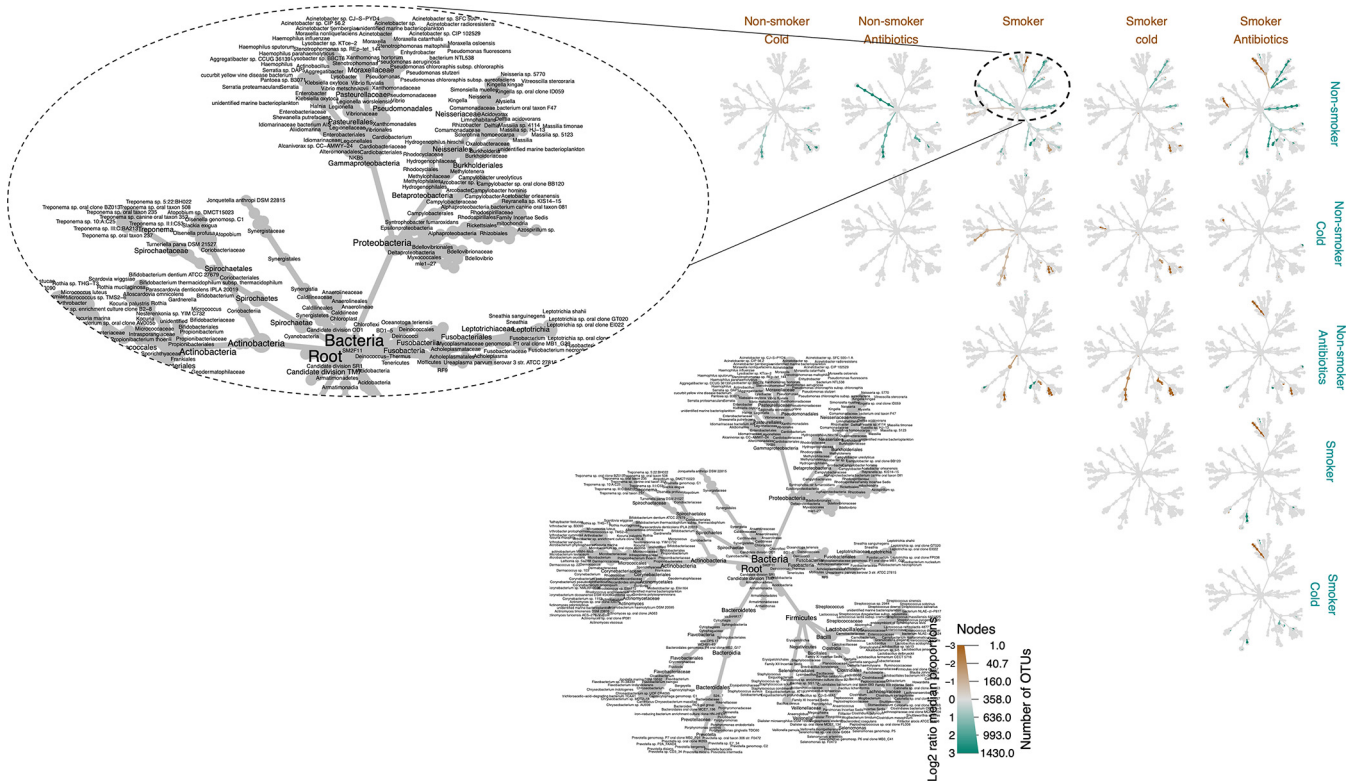


FIG 4 Heat tree matrix comparing the OTUs of the bacterial core of the oropharyngeal microbiota of smokers and nonsmokers when healthy, expressing common cold symptoms, or taking antibiotics. The lower left-hand side diagram shows the phylogeny of the pooled data set, and the sizes of the circles associated with different taxa indicate their relative abundances. The brown and cyan colors indicate significant differences across pairwise abundances, with the color indicating in which group the abundances were greater (with deeper colors indicating higher abundances). The gray color represents no significant difference in relative abundance. Refer to a high resolution rendering in Fig. S5 in the supplementary materials.

symptoms, and smokers and nonsmokers on antibiotics (Fig. 4) (refer to a high resolution rendering in Fig. S5). A heat tree analysis of healthy smokers and healthy nonsmokers showed greater a relative abundance of Fusobacteriaceae, Neisseriales, and Spirochaetida in nonsmokers and a greater abundance of Actinobacteria in smokers.

Healthy nonsmokers had a greater abundance of Bacteroidales and Fusobacteria, compared to nonsmokers with colds. Moreover, healthy nonsmokers showed a greater abundance of Spirochaetidae and Porphyromonadacea, compared to nonsmokers on antibiotics. Interestingly, there were no differences in abundance between nonsmokers with colds and nonsmokers on antibiotics. Healthy smokers and smokers with cold symptoms showed no differences in microbial community abundance, whereas smokers on antibiotics had greater abundance of *Klebsiella* and *Enterobacter* but lower *Porphyromonas*, compared to healthy smokers.

Next, we explored whether the microbial assemblage composition was driven by the host environment or by competition via the use of the nearest-relative-index (NRI) and the nearest-taxon-index (NTI). Values greater than 2 indicate strong phylogenetic clustering (driven by environmental filtering), whereas those less than 2 indicate phylogenetic overdispersion (indicating the role of ecological interactions, rather than the environment, in shaping the communities). Our analysis suggested that environmental filtering (rather than competition) was the process driving community composition in the oropharynx (Fig. 5). Moreover, for NTI, environmental filtering was greater in smokers than in nonsmokers. It is worth considering that NRI reflects phylogenetic clustering in a broad sense (considering the whole phylogenetic tree), with the negative values representing an evenly spread community. Nonsmokers on antibiotics had the greatest NRI values, suggestive of the highest degree of environmental filtering, when considering the phylogenetic tree as a whole. NTI focuses on the tips of the tree, with positive values indicating that species co-occur with

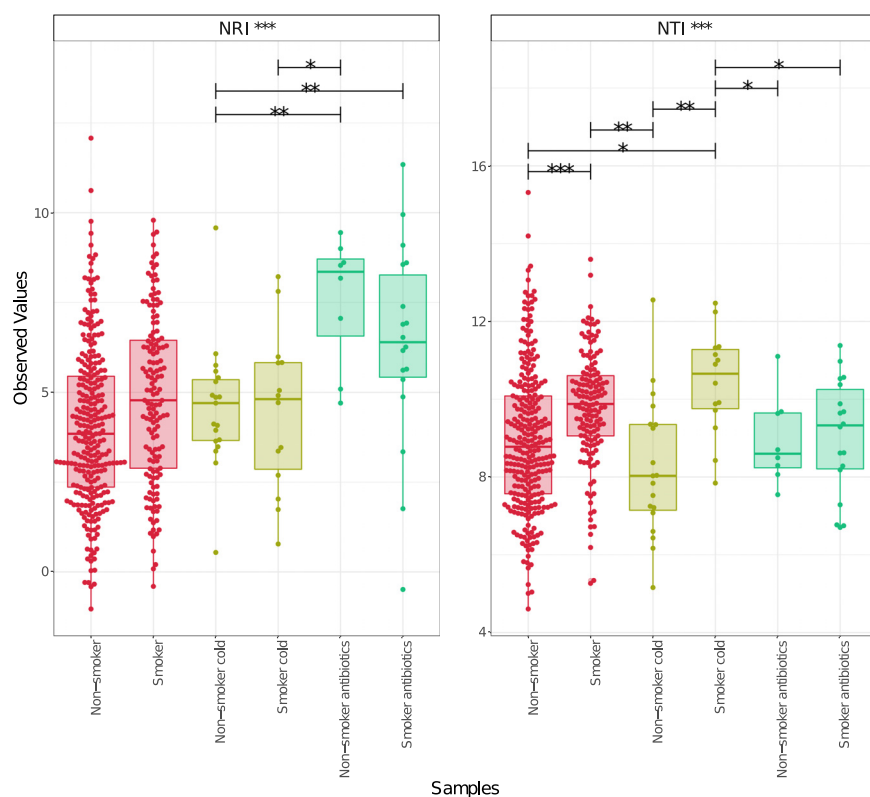


FIG 5 Nearest-taxon-index (NTI) and nearest-relative-index (NRI), considering the presence/absence of OTUs in the microbiota samples of smokers and nonsmokers when healthy, expressing common cold symptoms, or taking antibiotics. Note the difference in scale. The lines connect two categories if the values are significant per an ANOVA, with one within subject factor accounting for the participant IDs.

more closely related species than expected and negative values indicating that closely related species do not co-occur. Evidently, when showing cold symptoms, environmental filtering was significantly higher in smokers than in nonsmokers, suggesting that more species co-occurred with closely related species. In general, environmental filtering was greater in relation to closely-related species (NRI), compared to when accounting for the whole phylogenetic tree (NTI).

For each participant, the core microbiota was identified to infer interaction matrices. These were used to determine the dynamical properties of communities by inferring the Jacobian matrix and its dominant eigenvalue, indicating that all communities were locally stable. We subsequently determined whether the local stability of the smokers and nonsmokers' oropharyngeal microbiota differed by comparing their dominant eigenvalues (nonsmokers mean = -5.908 , SD = $+3.537$; smokers mean = -6.953 , SD = $+2.989$). An analysis of variance (ANOVA) indicated no difference between them ($F_{1,21} = 0.495$, $P = 0.490$).

To determine whether they were globally stable (i.e., permanent), interaction matrices were then tested for the existence of an average Lyapunov function. Out of the 15 inferred interaction networks from nonsmokers, 6 (40%) were globally stable. Out of the 8 smoker networks, 4 (50%) were globally stable. Thus, our results provide no evidence for differences in local or global stability between smokers and nonsmokers.

With the exception of one nonsmoker interaction network, all smoker ($n = 8$) and nonsmoker ($n = 14$) communities were locally stable after a new species was added to the Jacobian matrix, indicative of resistance to species invasion, irrespective of smoking status.

DISCUSSION

The oral microbiota plays a vital role in the maintenance of oral health, and dysbiosis is associated with a number of diseases, including periodontal disease (43, 44). Findings of

studies investigating the effects of smoking on the oral microbiota have been inconsistent, with some reporting changes in community composition that are associated with varying taxa (5, 20, 45, 46), while others reported no association between the oral microbiota and smoking (47–49). Variability in the definition of smoking, sample site selection, methodology, study population, and inclusion/exclusion of samples based on preconditions are likely to account for at least some of these inconsistencies. Here, we collected longitudinal samples of the oropharyngeal microbiota of 30 healthy participants (18 nonsmokers and 12 smokers), using a standardized sampling approach by sampling the same site and using the same methodology (as described above) in the same participants and study population, avoiding the issues noted above.

Smoking constitutes a press perturbation that could result in changes in the richness, composition, and abundance of the microbiota. As hypothesized, in our results, smoking decreased alpha diversity in line with that reported by other studies (50, 51), similar to antibiotic use, which we consider to be a short-term pulse perturbation (52, 53). Our results suggest that this process results in lower diversity in communities undergoing a simultaneous, significant press perturbation (i.e., smoking). Fewer niches being available due to a relatively more hostile microclimate may result in a lower species richness and a loss of protection by mutualistic species and, thus, disease (21, 54). However, others (6) have found that smokers had increased alpha diversity (Shannon diversity), compared to nonsmokers, arguing that smoking decreases the niche saturation abilities of a community that selects for a specific group of organisms.

As hypothesized, our study provided evidence that press perturbations, such as smoking, impacted beta diversity, even within the context of short-duration pulse perturbations (cold or antibiotic use). Our results showed that the seven OTUs that were responsible for the changes in the beta diversity patterns were shared across smokers and nonsmokers, regardless of their health status. The effect of the short-term pulse perturbations in the form of cold and antibiotic use was evident when measuring the local contribution to beta diversity (LCBD), which indicated a high uniqueness of the bacterial community composition. Our analysis showed that nonsmokers exhibiting cold symptoms and participants on antibiotics (irrespective of smoking status) displayed the greatest LCBD values. Colds only affected undisturbed communities (i.e., nonsmokers), whereas sufficiently large pulse perturbations (antibiotic treatment) impacted smokers and nonsmokers (i.e., disturbed and undisturbed communities) equally, as evidenced by the alpha diversity values becoming more similar between smokers and nonsmokers. This suggests that disturbed communities become more similar to one another, irrespective of smoking status.

High values of the nearest-relative-index (NRI) and nearest-taxon-index (NTI) provide evidence that these oropharyngeal communities are dominated by environmental filtering. This may be related to nutritional, spatial, and metabolic factors selecting for species that possess suitable traits to exist within specific conditions (55). As hypothesized, smokers show greater environmental filtering, compared to nonsmokers, and participants on antibiotic treatment show the greatest degree of environmental filtering, irrespective of smoking status, suggesting that these forms of environmental perturbations are strong enough to override the importance of ecological interactions in shaping the microbial communities. Smoking is likely to affect the microenvironment, specifically, the oxygen, pH, and acid production, thereby resulting in the selection of specific microbial community members of microaerophilic and anaerobic bacteria that can dominate due to the lower oxygenation (6, 56).

Studies have shown that more diverse and complex communities tend to be generally more resistant to perturbations (37). Contrary to our expectations, stability models based on our data indicated that local and global stability did not differ between smokers and nonsmokers, suggesting that microbial communities of smokers are resistant and resilient to small and large perturbations. Moreover, our results indicated no differences in invasibility between smokers and nonsmokers when invasibility was measured as instability.

The human microbiota is often found to be stable and generally able to maintain

homeostasis, even in the face of perturbations (57). Other authors have argued that the oral cavity is characterized by an even greater degree of stability, compared to that of other body sites (58). The heterogeneity of the oral cavity increases niche space, in which many microbial community members live on surfaces in biofilms, which likely allows them to survive adverse conditions, such as smoking and antibiotic exposure (24, 51, 59). Moreover, others have found that even upon smoking cessation, the differences in microbial community composition generally persisted over the time period of this study, suggesting that smoking can introduce an alternative but stable state (22).

The microbiota is not just a passive passenger; its modification can play a key role in contributing to or causing many pathophysiological processes (21, 42, 60). Others have shown that smoking creates pathogen-rich and commensal-poor communities (5, 6). Similar to the findings of others, there were decreases in the relative abundance of *Prevotella* and *Neisseria* spp. (61, 62) and an increased relative abundance of *Veillonella* spp. in smokers, compared to nonsmokers (50). These differences may be associated with differences in oxygen tension, with smoking promoting the presence of anaerobes (63). Overall, we found no evidence suggesting that smoking communities in this particular group of participants were characterized by pathogen-rich and commensal-poor species. It may be worth noting, however, that the excess of Gram-negative *Klebsiella* and *Enterobacter* in smokers on antibiotics are increasingly seen in aspiration pneumonia (64), which could suggest that smoking/antibiotics are possible risk factors for aspiration pneumonia caused by these organisms in older age groups.

In conclusion, our study provides evidence that microbiota changes with smoking are likely a combination of parts of the microbiota not surviving the direct toxicity of cigarette smoke and its indirect consequences via the associated changes in the immune system and cellular processes, in general. This manifested in the analyses as the importance of environmental filtering in the structuring of the microbial communities. This suggests that environmental filtering dictates which OTUs may be present in the community, resulting in the perturbed communities' final, alternative states. Even though the oropharyngeal microbiota of smokers is reduced in terms of diversity, compared to that of nonsmokers, it is ecologically no less stable or more invasible. Therefore, the increased risk of infectious diseases in smokers is not necessarily facilitated by the effect of smoking on ecological changes in the composition or function of the oropharyngeal microbiota but is more likely due to its effects on the local and systemic immune system (65), which may alter the susceptibility to contract an infection and the course of disease (66, 67). A next step to support the findings from these analyses would be to conduct experiments or generate empirical data that investigate the stability and invasibility of microbiota in response to cigarette smoke, cold-symptoms, or antibiotic disturbances.

MATERIALS AND METHODS

Participants and sample collection. 30 participants (18 nonsmokers and 12 smokers) were recruited for this study, representing both genders (21 females, 9 males) and ranging in age from 18 to 40 (mean = 25.18 years, SD = 6.71). 10 of the females were smokers (smoking on average 32.25 [SD = 25.76] cigarettes per week, having smoked for 5.1 [SD = 3.70] years on average), and 2 of the male participants were smokers (smoking on average 43.14 [SD = 24.12] cigarettes per week, having smoked for 9.5 [SD = 9.19] years on average). Further information on sample sizes is provided in Table S1.

All participants were healthy (without illnesses or undergoing medical treatment), not on any long-term medication, and had not stopped smoking recently. The study was approved by the University of Glasgow Ethics Committee (Ethics Applications 2012107 and 200140021).

Participants were provided with Sigma Transwabs swabs in liquid amies (Medical Wire Ltd., United Kingdom) for bacterial detection, and samples were taken from the tonsils and the posterior wall of the tonsil, as suggested in the literature (68). Participants were asked to take a swab once weekly, early in the morning, prior to brushing their teeth and having breakfast, as well as to keep a diary providing information about their health status, self-reporting any changes there. Samples were collected weekly, and bacterial swabs were processed, after being kept on ice, within 2 h after collection. Sampling occurred in Glasgow (United Kingdom) for the nonsmokers from January until May of 2013, and sampling recommenced from September until December of 2013. The sampling period for the smokers took place from November of 2014 to June of 2015, giving a sampling period of 30 weeks. Staffing constraints did not allow for the sampling of smokers and nonsmokers within exactly the same time period. While we recognize this issue, we argue it unlikely that the differences observed between the smokers and nonsmokers are driven solely by annual fluctuations in microbiota.

DNA extractions. DNA was extracted using a QIAamp DNA Minikit (Qiagen Ltd., United Kingdom), following the manufacturer's instructions. The extracted DNA was quantified using a Qubit 2.0 (Thermo Fisher Scientific, Q32866) and a Picogreen HS DNA Assay (Invitrogen Ltd., United Kingdom). 5 μ L of DNA were mixed with 2 μ L of loading dye on a 1% agarose gel (1 g agarose to 100 mL TBE), along with a 1 kb Invitrogen DNA ladder, and run at 100 V for 60 min to check purity, and they were subsequently stored at -20°C until required (69). The details on processing the samples are given in Chapter 2 of <http://theses.gla.ac.uk/8163/>, and they have also been described in Rani Ram (2017) (70).

A clone library was prepared using an Invitrogen Topo-Seq Kit (Invitrogen Ltd., United Kingdom) as a quality control step to produce a mock community for future Illumina MiSeq runs. The purpose of the mock community was to act as a positive control for each MiSeq run so as to ensure that the correct sequences were being sequenced. To reduce PCR biases, all of the template DNA was diluted to the same concentration, and the minimum number of cycles was used in each PCR run to reduce nonspecific binding, with appropriate controls in place. A positive control (DNA from the mock community) and a negative control (for each different reverse barcode, using nuclease free water) was set up for each PCR run. To ensure that the source of bacterial sequences was not the swab itself or the DNA isolation reagents, PCR was performed on DNA isolated from an unused swab. To confirm that the PCR reagents were not sources of bacterial sequences, a PCR of the no-template extraction control was also performed.

Bioinformatics. Abundance tables were obtained by constructing OTUs (a proxy for species) as previously described (42) and as follows. The trimming and filtering of the paired-end sequencing reads was performed using Sickle (version 1.2) by applying a sliding window approach and trimming the regions where the average base quality dropped below 20 (71). This applied a 10 bp length threshold to discard reads that fell below this length. BayesHammer (72) was used from the SPAdes assembler (version 2.5) to error correct the paired-end reads, and this was followed by PANDAsq (version 2.4) with a minimum overlap of 50 bp to assemble the forward and reverse reads into a single sequence that spanned the entire V1-V2 region (73). The above choice of software showed a reduction in substitution errors by 77 to 98%, with an average of 93.2% for the MiSeq data sets (74). UPARSE (version 7.0.1001) was used for OTU construction (75). The approach pools together the reads from different samples and adds barcodes to keep an account of the samples from which these reads originate. The reads are then dereplicated and sorted by decreasing abundance, discarding singletons. The reads were clustered based on 97% similarity, discarding any reads that were shorter than 32 bp. The original barcoded reads were then matched against OTUs with 97% similarity (a proxy for species-level separation) to generate OTU tables for different samples. The representative OTUs were then taxonomically classified against the RDP database, using the standalone RDP classifier (version 2.6) (76). The phylogenetic distances between OTUs were produced by first using MAFFT (version 7.040) (77) to align the OTUs against each other and then by using FastTree (version 2.1.7) on these alignments to generate an approximately maximum likelihood phylogenetic tree (78). Potentially contaminating amplicon sequence variants were identified and removed using the decontam package in R (79). The OTU table, phylogenetic tree, taxonomic information, and metadata were then used in a statistical analysis. Chimeras were removed using two approaches: an external Chimera Slayer Gold database and a *de novo* approach in which the abundant reads served as a reference database. Traditional pipelines were modified to get the optimum accuracy for amplicons, based on benchmarking studies (73, 80, 81). Subsequently, the OTU table, phylogenetic tree, taxonomic information, and metadata were used in a multivariate statistical analysis. While the *de novo* chimera removal step removes reads that have chimeric models that are built from more abundant reads, a few chimeras may be missed, especially if they have parents that are absent from the reads or are present in low abundance. To deal with that, we used a reference-based chimera filtering step from a Gold database (http://drive5.com/uchime/uchime_download.html) that is derived from the ChimeraSlayer reference database in the Broad Microbiome Utilities (<http://microbiomeutil.sourceforge.net/>).

Statistical analysis. Statistical analyses were performed in R, using the tables and data that were generated as well as the metadata associated with the study. All samples that contained less than 5,000 reads were discarded in the analysis to allow for the comparison of all samples with sufficient statistical power.

For the community analysis (including the alpha and beta diversity analyses), we used the vegan package (82). For the alpha diversity measures, which provided us with an insight into the changes in richness and evenness with smoking, antibiotic, and cold status, we calculated: OTU richness (the estimated number of OTUs per sample), Shannon entropy (a commonly used index by which to measure the richness and evenness of a community within a sample), and Simpson diversity, which takes both sample richness and evenness into account but is less sensitive to species richness, using Renyi-Hill numbers (83). The alpha diversity measures were calculated after rarefying the abundance table to the minimum library size by using rarefaction curves that were obtained for each participant to approximate the OTUs that were detected as a function of the sequencing depth. As a result, the samples were rarefied to the minimum number of reads (5,118) to test for alpha diversity, and the relative abundance of taxa for each sample was calculated by dividing the read counts of a taxon by the sample size. This value ranges from 0 to 1. Rarefaction curves can be found in Rani Ram (2017) (70). The data were analyzed using an ANOVA with Tukey's *post hoc* test, after checking the assumptions (normality, heterogeneity of variance, and independence). To adjust for patient IDs, all of the pairwise statistics were done using one within subject factor in anova, as `aov(value ~ Groups + Error(ParticipantID/Groups))`, by incorporating Error().

In order to investigate the multivariate homogeneity of the group dispersion (i.e., how variable the groups are) between multiple conditions (i.e., smoking, antibiotic, and cold status) we used vegan's `betadisper()` function to calculate three different distance measures: Bray-Curtis (which considers the species abundance count), unweighted unifracs (which considers the phylogenetic distance between the

branch lengths of OTUs that are observed in different samples without taking into account the abundances) and weighted unifracs (unweighted unifracs distance, weighted by the abundances of the OTUs), using the `metamds()` function in `vegan`. The distances between objects and group centroids are handled by reducing the original distances to principal coordinates and subsequently performing an ANOVA on them, according to a four-factor design, including health status (six levels: nonsmoker, smoker, nonsmoker with cold, smoker with cold, nonsmoker on antibiotics, smoker on antibiotics), sex (two levels: male, female), age (continuous), and participant (30 levels). We used `vegan`'s `adonis()` function for an analysis of variance using distance matrices to partition sources of variation. This function will be referred to as PERMANOVA, and it fits linear models to distance matrices and uses a permutation test with pseudo-F ratios (84).

To measure beta diversity, we performed a local contribution to beta diversity (LCBD) analysis (85). In the context of this longitudinal study, LCBD provides a means by which to show the difference of the microbial community structure from a single sample, based on the average of all samples (with higher LCBD values representing communities with greater beta diversity values or outliers). We used the `LCBD.comp()` function from the `adespatial` package (86) to calculate three different measures: Hellinger distance (abundances), unweighted unifracs (phylogenetic distance), and weighted unifracs (phylogenetic distance weighted by abundance) dissimilarities. LCBD gives the sample-wise local contributions to beta diversity, which are derived from the proportions of the total beta diversity.

We subsequently used the BVSTEP routine to complement the beta diversity analysis, as it is an algorithm that allows for the identification of the OTUs that are causing the major shifts in beta diversity. The BVSTEP algorithm (87) searches for the highest correlation between the dissimilarities of fixed and variable multivariate data sets using the `bvStep()` function from the `sinker` package, using a permutation of the $2^n - 1$ features in the data set (i.e., OTUs) (88). Testing all feature combinations is unrealistic and computationally intractable when the species richness in the data set is high. To deal with this, we used the abundance table with the 1,000 most abundant OTUs to best correlate with the overall similarities in the features (i.e., grouping factors, smoking, cold, and antibiotic status), given all of the OTUs, assuming that the most abundant species may play a significant role.

To determine whether the phylogenetic community structure in each sample was driven by competition among taxa or by strong environmental pressure (i.e., the host environment) we calculated the mean-nearest-taxon-distance (MNTD, using the `mntd()` function), nearest-taxon-index (NTI, using the `ses.mntd()` function), mean-phylogenetic-diversity (MPD, using the `mpd()` function) and nearest-relative-index (NRI, using the `ses.mpd()` function) in the `picante` package in R (89). NTI and NRI are the negative outputs from the `ses.mntd()` and `ses.mpd()` functions, respectively.

NTI and NRI quantify the number of standard deviations that the observed MNTD and MPD values, respectively, is from the mean of the null distribution (999 randomizations using `null.model = "richness"`) in the `ses.mntd()` and `ses.mpd()` functions, only considering the taxa as present/absent, without taking their abundances into account. Based on recommendations, we used the top 1,000 most abundant OTUs for the calculation of these measures (90).

Subset regression is a model selection approach that provides evidence for the important factors that are associated with changes in the microbial community composition, including the effects of different disturbances (e.g., smoking, cold, and antibiotic status), thereby providing further support for the above analysis. Subset regression was performed against different microbiota metrics, testing all possible combinations of explanatory variables and subsequently selecting the best model, according to statistical criteria, based on recommendations (91). The code is available at (<http://www.sthda.com/english/articles/37-model-selection-essentials-in-r/155-best-subsets-regression-essentials-in-r/>) (91). We used the `regsubsets()` function from the `leaps` package (92) to identify the best models of different sizes by specifying the option `nvmax` set to the maximum number of predictors to be incorporated in the model. After having obtained the best subsets, the k-fold cross-validation consisted of first dividing the data into k subsets, of which each subset (10%) served successively as test data and the remaining subset (90%) as training data, using a custom function that utilized the `train()` function from the `caret` package (93). The average cross-validation error was calculated as the model prediction error. We used the `tab_model()` function from the `sjPlot` package to obtain the statistics for each model (94).

There has been no single standardized approach by which to identify the core microbiota, with most studies reporting the core based on the presence/absence of taxa at specific threshold values of abundance and prevalence (95). Here, we defined the core microbiota by setting a detection threshold of 0.001 (i.e., 0.1% compositional abundance threshold in a sample) and a prevalence of the OTUs to be found in samples between 85 and 92% (depending on the total sample size available for each longitudinal data set) for each participant of this study, following advice given in Shetty (96) and using the `microbiota` package (97).

We subsequently used the `heat_tree_matrix()` function in the `metacoder` package (98) to detect differences in the relative abundances in the core OTUs (based on the \log_2 median ratio) between participants, using a Wilcoxon test.

Local and global stability and invisibility. Here, we refer to stability in the context of ecological community dynamics, defined broadly as the ability of an ecosystem to persist through perturbations, which is influenced by changes in the interactions between different species. Understanding under which conditions the species in a community coexist in the long term is a central question in ecology that can be addressed by investigating local and global stability. Local stability is a measure of the persistence of communities subjected to small perturbations, whereas global stability (also referred to as permanence) refers to a global property of a community that only requires the densities of species to increase when they are rare (99, 100). To address the question of whether the potentially reduced stability of smokers' microbiotas could be involved in pathogen invasions, we characterized the ecological

stability of the microbiotas of both smokers and nonsmokers by using two methods, namely, local and global stability analysis, working under the assumption that microbiotas, including the pharyngeal microbiotas, are stable over time (45, 101, 102).

The calculation of these stability measures requires the inference of the microbial core for each participant as described above; we used the `limits()` function in the `seqtime` package (103, 104), which utilizes time-series data, to infer the microbial interaction matrices that corresponded to the classical generalized Lotka-Volterra dynamics through deterministic model fitting. We only selected participants for which we had at least nine samples (after discussion on sample size and network inference with the package authors). This resulted in estimated interaction matrices for 15 nonsmokers and 8 smokers.

Interaction matrices can be used to determine the dynamical properties of communities by inferring the Jacobian matrix and its dominant eigenvalue, where the equilibrium is locally stable if all of the eigenvalues of the interaction matrix have negative real parts (100, 105, 106). For each inferred interaction matrix, we computed the dominant eigenvalue (the most positive eigenvalue) to evaluate local stability.

Local stability applies only to small perturbations, and its ecological relevance has been debated (100, 107, 108). A dynamical system is globally stable if it does not experience extinction following arbitrarily large stochastic perturbations. Communities are defined as globally stable if an average Lyapunov function exists near the boundary of the state space and if the system is dissipative (109), meaning that even if a system experiences a large perturbation, it will return to some form of equilibrium without resulting in species extinction. To determine whether the interaction matrices were globally stable, we tested our interaction matrices for the existence of an average Lyapunov function using (110) “sufficient condition”, as laid out prior (111), solving the problem as a linear programming problem in MATLAB. Here, we calculated both measures of stability for the interaction matrix estimated for the core communities of each participant to determine if those in smokers were less stable, globally and locally, compared to those of nonsmokers (110, 111).

To test whether the microbial communities in smokers were more invasible, for each participant, we recomputed the equilibrium point of the Jacobian matrix by adding a new species to the community (where all of the rows and columns of the “invader” OTU were zero) and calculating the dominant eigenvalue of this augmented interaction matrix. If stable, then the species were at equilibrium and were not increasing in abundance (i.e., the community was not invasible). If unstable, then the community is invasible, with the invader increasing in abundance.

Data availability. The following data: `otu_table.biom`, `otus.fasta`, `otus.tre`, and `meta_data.csv` file are available on GitHub (<https://github.com/umerijaz/pharynxmicrobiome>). The scripts are available as part of `microbiomeSeq` (https://userweb.eng.gla.ac.uk/umer.ijaz/projects/microbiomeSeq_Tutorial.html) and at <https://userweb.eng.gla.ac.uk/umer.ijaz/bioinformatics/ecological.html>.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, XLSX file, 0.01 MB.

SUPPLEMENTAL FILE 2, PDF file, 4.1 MB.

ACKNOWLEDGMENTS

We are grateful for the support by the Glasgow University Kelvin-Smith studentship (A.R.), the IISF Catalyst 419 Grant (097821/Z/11/Z) by the Wellcome Trust (J.L. and T.E.), and the Leverhulme Trust grant (RPG-2017-279; J.L., D.H., and L.B.).

No conflict of interest is declared.

REFERENCES

- Huttenhower C, Gevers D, Knight R, Abubucker S, Badger JH, Chinwalla AT, Creasy HH, Earl AM, FitzGerald MG, Fulton RS. 2012. Structure, function and diversity of the healthy human microbiome. *Nature Lett* 486:207.
- Avila M, Ojcius DM, Yilmaz Ö. 2009. The oral microbiota: living with a permanent guest. *DNA Cell Biol* 28:405–411. <https://doi.org/10.1089/dna.2009.0874>.
- Cho I, Blaser MJ. 2012. The human microbiome: at the interface of health and disease. *Nat Rev Genet* 13:260–270. <https://doi.org/10.1038/nrg3182>.
- Weathers KC, Strayer DL, Likens GE. 2012. Fundamentals of ecosystem science. Academic Press.
- Kumar PS, Matthews CR, Joshi V, Jager M d, Aspiras M. 2011. Tobacco smoking affects bacterial acquisition and colonization in oral biofilms. *Infect Immun* 79:4730–4738. <https://doi.org/10.1128/IAI.05371-11>.
- Mason MR, Preshaw PM, Nagaraja HN, Dabdoub SM, Rahman A, Kumar PS. 2015. The subgingival microbiome of clinically healthy current and never smokers. *ISME J* 9:268–272. <https://doi.org/10.1038/ismej.2014.114>.
- Morris A, Beck JM, Schloss PD, Campbell TB, Crothers K, Curtis JL, Flores SC, Fontenot AP, Ghedin E, Huang L, Jablonski K, Kleeerup E, Lynch SV, Sodergren E, Twigg H, Young VB, Bassis CM, Venkataraman A, Schmidt TM, Weinstock GM, Lung HIV Microbiome Project. 2013. Comparison of the respiratory microbiome in healthy nonsmokers and smokers. *American J Respiratory and Critical Care Medicine* 187:1067–1075. <https://doi.org/10.1164/rccm.201210-1913OC>.
- Shchipkova AY, Nagaraja HN, Kumar PS. 2010. Subgingival microbial profiles of smokers with periodontitis. *J Dent Res* 89:1247–1253. <https://doi.org/10.1177/0022034510377203>.
- Jaspers I. 2014. Cigarette smoke effects on innate immune mechanisms in the nasal mucosa. Potential effects on the microbiome. *Annals of the American Thoracic Society* 11:S38–S42. <https://doi.org/10.1513/AnnalsATS.201306-154MG>.
- Gerrard JW, Heiner DC, Ko CG, Mink J, Meyers A, Dosman JA. 1980. Immunoglobulin levels in smokers and non-smokers. *Ann Allergy* 44: 261–262.
- Hughes DA, Haslam PL, Townsend PJ, Turner-Warwick M. 1985. Numerical and functional alterations in circulatory lymphocytes in cigarette smokers. *Clin Exp Immunol* 61:459–466.

12. Page RC, Kornman KS. 1997. The pathogenesis of human periodontitis: an introduction. *Periodontol* 2000 14:9–11. <https://doi.org/10.1111/j.1600-0757.1997.tb00189.x>.
13. Preber H, Bergström J, Linder LE. 1992. Occurrence of periopathogens in smoker and non-smoker patients. *J Clin Periodontol* 19:667–671. <https://doi.org/10.1111/j.1600-051x.1992.tb01716.x>.
14. Eichel B, Shahrik HA. 1969. Tobacco smoke toxicity: loss of human oral leukocyte function and fluid-cell metabolism. *Science* 166:1424–1428. <https://doi.org/10.1126/science.166.3911.1424>.
15. Kenney EB, Kraal JH, Saxe SR, Jones J. 1977. The effect of cigarette smoke on human oral polymorphonuclear leukocytes. *J Periodontol Res* 12: 227–234. <https://doi.org/10.1111/j.1600-0765.1977.tb00126.x>.
16. Kraal JH, Kenney EB. 1979. The response of polymorphonuclear leukocytes to chemotactic stimulation for smokers and non-smokers. *J Periodontol Res* 14:383–389. <https://doi.org/10.1111/j.1600-0765.1979.tb00235.x>.
17. Noble RC, Penny BB. 1975. Comparison of leukocyte count and function in smoking and nonsmoking young men. *Infect Immun* 12:550–555. <https://doi.org/10.1128/iai.12.3.550-555.1975>.
18. Bagaitkar J, Demuth DR, Scott DA. 2008. Tobacco use increases susceptibility to bacterial infection. *Tob Induc Dis* 4:12. <https://doi.org/10.1186/1617-9625-4-12>.
19. Al-Zyoud W, Hajjo R, Abu-Siniyeh A, Hajjaj S. 2020. Salivary microbiome and cigarette smoking: a first of its kind investigation in Jordan. *Int J Environmental Res and Public Health* 17:256. <https://doi.org/10.3390/ijerph17010256>.
20. Wu J, Peters BA, Dominianni C, Zhang Y, Pei Z, Yang L, Ma Y, Purdue MP, Jacobs EJ, Gapstur SM, Li H, Alekseyenko AV, Hayes RB, Ahn J. 2016. Cigarette smoking and the oral microbiome in a large study of American adults. *ISME J* 10:2435–2446. <https://doi.org/10.1038/ismej.2016.37>.
21. Huang C, Shi G. 2019. Smoking and microbiome in oral, airway, gut and some systemic diseases. *J Transl Med* 17:225. <https://doi.org/10.1186/s12967-019-1971-7>.
22. Ramakrishnan VR, Frank DN. 2015. Impact of cigarette smoking on the middle meatus microbiome in health and chronic rhinosinusitis. *Int Forum Allergy Rhinol* 5:981–989. <https://doi.org/10.1002/alf.21626>.
23. Kilian M. 2018. The oral microbiome—friend or foe? *European J Oral Sciences* 126:5–12. <https://doi.org/10.1111/eos.12527>.
24. Zaura E, Brandt BW, Teixeira de Mattos MJ, Buijs MJ, Caspers MPM, Rashid M-U, Weintraub A, Nord CE, Savell A, Hu Y, Coates AR, Hubank M, Spratt DA, Wilson M, Keijsers B, Crielaard W. 2015. Same exposure but two radically different responses to antibiotics: resilience of the salivary microbiome versus long-term microbial shifts in feces. *mBio* 6:e01693-15–e01615. <https://doi.org/10.1128/mBio.01693-15>.
25. Flanagan JL, Brodie EL, Weng L, Lynch SV, Garcia O, Brown R, Hugenholz P, DeSantis TZ, Andersen GL, Wiener-Kronish JP, Bristow J. 2007. Loss of bacterial diversity during antibiotic treatment of intubated patients colonized with *Pseudomonas aeruginosa*. *J Clin Microbiol* 45:1954–1962. <https://doi.org/10.1128/JCM.02187-06>.
26. Wilson MT, Hamilos DL. 2014. The nasal and sinus microbiome in health and disease. *Curr Allergy Asthma Rep* 14:485. <https://doi.org/10.1007/s11882-014-0485-x>.
27. Jakobsson HE, Jernberg C, Andersson AF, Sjölund-Karlsson M, Jansson JK, Engstrand L. 2010. Short-term antibiotic treatment has differing long-term impacts on the human throat and gut microbiome. *PLoS One* 5. <https://doi.org/10.1371/journal.pone.0009836>.
28. Hanada S, Pirzadeh M, Carver KY, Deng JC. 2018. Respiratory viral infection-induced microbiome alterations and secondary bacterial pneumonia. *Front Immunol* 9:2640. <https://doi.org/10.3389/fimmu.2018.02640>.
29. Kloepfer KM, Sarsani VK, Poroyko V, Lee WM, Pappas TE, Kang T, Grindle KA, Bochkov YA, Janga SC, Lemanske RF. 2017. Community-acquired rhinovirus infection is associated with changes in the airway microbiome. *Journal of Allergy and Clinical Immunology* 140:312–315.e8.
30. Rosas-Salazar C, Shilts MH, Tovchigrechko A, Schobel S, Chappell JD, Larkin EK, Shankar J, Yooseph S, Nelson KE, Halpin RA, Moore ML, Anderson LJ, Peebles RS, Das SR, Hartert TV. 2016. Differences in the nasopharyngeal microbiome during acute respiratory tract infection with human rhinovirus and respiratory syncytial virus in infancy. *J Infect Dis* 214:1924–1928. <https://doi.org/10.1093/infdis/jiw456>.
31. Yi H, Yong D, Lee K, Cho Y-J, Chun J. 2014. Profiling bacterial community in upper respiratory tracts. *BMC Infect Dis* 14:583. <https://doi.org/10.1186/s12879-014-0583-3>.
32. Hofstra JJ, Matamoros S, van de Pol MA, de Wever B, Tanck MW, Wendt-Knol H, Deijs M, van der Hoek L, Wolthers KC, Molenkamp R, Visser CE, Sterk PJ, Lutter R, de Jong MD. 2015. Changes in microbiota during experimental human Rhinovirus infection. *BMC Infect Dis* 15:336. <https://doi.org/10.1186/s12879-015-1081-y>.
33. Kraft NJ, Adler PB, Godoy O, James EC, Fuller S, Levine JM. 2015. Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology* 29:592–599. <https://doi.org/10.1111/1365-2435.12345>.
34. Lewontin RC. 1969. The meaning of stability. *Brookhaven Symp Biol* 22: 13–24.
35. MacArthur R. 1955. Fluctuations of animal populations and a measure of community stability. *Ecology* 36:533–536. <https://doi.org/10.2307/1929601>.
36. Margalef R. 1969. Diversity and stability: a practical proposal and a model of interdependence.
37. McCann KS. 2000. The diversity-stability debate. *Nature* 405:228–233. <https://doi.org/10.1038/35012234>.
38. Cruz PD, Pridaux L, Wagner J, Ng SC, McSweeney C, Kirkwood C, Morrison M, Kamm MA. 2012. Characterization of the gastrointestinal microbiota in health and inflammatory bowel disease. *Inflamm Bowel Dis* 18:372–390. <https://doi.org/10.1002/ibd.21751>.
39. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. 2012. Diversity, stability and resilience of the human gut microbiota. *Nature Lett* 489:220. <https://doi.org/10.1038/nature11550>.
40. Relman DA. 2012. The human microbiome: ecosystem resilience and health. *Nutrition Rev* 70:52–59. <https://doi.org/10.1111/j.1753-4887.2012.00489.x>.
41. Hu J, Amor DR, Barbier M, Bunin G, Gore J. 2022. Emergent phases of ecological diversity and dynamics mapped in microcosms. *Science* 378: 85–89. <https://doi.org/10.1126/science.abm7841>.
42. Bach LL, Ram A, Ijaz UZ, Evans TJ, Lindström J. 2020. A longitudinal study of the human oropharynx microbiota over time reveals a common core and significant variations with self-reported disease. *Front Microbiol* 11: 573969. <https://doi.org/10.3389/fmicb.2020.573969>.
43. Dewhirst FE, Chen T, Izard J, Paster BJ, Tanner ACR, Yu WH, Lakshmanan A, Wade WG. 2010. The human oral microbiome. *J Bacteriol* 192:5002–5017. <https://doi.org/10.1128/JB.00542-10>.
44. Wade WG. 2013. The oral microbiome in health and disease. *Pharmacol Res* 69:137–143. <https://doi.org/10.1016/j.phrs.2012.11.006>.
45. Kumar PS, Mason MR. 2015. Mouthguards: does the indigenous microbiome play a role in maintaining oral health? *Front Cell Infect Microbiol* 5:35. <https://doi.org/10.3389/fcimb.2015.00035>.
46. Charlson ES, Chen J, Custers-Allen R, Bittinger K, Li H, Sinha R, Hwang J, Bushman FD, Collman RG. 2010. Disordered microbial communities in the upper respiratory tract of cigarette smokers. *PLoS One* 5:e15216. <https://doi.org/10.1371/journal.pone.0015216>.
47. Boström L, Bergström J, Dahlén G, Linder LE. 2001. Smoking and subgingival microflora in periodontal disease. *J Clin Periodontol* 28:212–219. <https://doi.org/10.1034/j.1600-051x.2001.028003212.x>.
48. Gomes SC, Nonnenmacher C, Susin C, Oppermann RV, Mutters R, Marcantonio RA. 2008. The effect of a supragingival plaque-control regimen on the subgingival microbiota in smokers and never-smokers: evaluation by real-time polymerase chain reaction. *J Periodontol* 79:2297–2304. <https://doi.org/10.1902/jop.2008.070558>.
49. Haffajee AD, Socransky SS. 2001. Relationship of cigarette smoking to the subgingival microbiota. *J Clin Periodontol* 28:377–388. <https://doi.org/10.1034/j.1600-051x.2001.028005377.x>.
50. Shanahan ER, Shah A, Koloski N, Walker MM, Talley NJ, Morrison M, Holtmann GJ. 2018. Influence of cigarette smoking on the human duodenal mucosa-associated microbiota. *Microbiome* 6:1–12.
51. Yu G, Phillips S, Gail MH, Goedert JJ, Humphrys MS, Ravel J, Ren Y, Caporaso NE. 2017. The effect of cigarette smoking on the oral and nasal microbiota. *Microbiome* 5:3. <https://doi.org/10.1186/s40168-016-0226-6>.
52. Ferrer M, Méndez-García C, Rojo D, Barbas C, Moya A. 2017. Antibiotic use and microbiome function. *Biochem Pharmacol* 134:114–126. <https://doi.org/10.1016/j.bcp.2016.09.007>.
53. Shaw LP, Bassam H, Barnes CP, Walker AS, Klein N, Balloux F. 2019. Modelling microbiome recovery after antibiotics using a stability landscape framework. *ISME J* 13:1845–1856. <https://doi.org/10.1038/s41396-019-0392-1>.
54. Einarsson GG, Comer DM, McIlreavey L, Parkhill J, Ennis M, Tunney MM, Elborn JS. 2016. Community dynamics and the lower airway microbiota in stable chronic obstructive pulmonary disease, smokers and healthy non-smokers. *Thorax* 71:795–803. <https://doi.org/10.1136/thoraxjnl-2015-207235>.

55. Keddy PA. 1992. Assembly and response rules: two goals for predictive community ecology. *J Vegetation Science* 3:157–164. <https://doi.org/10.2307/3235676>.
56. Ganesan SM, Joshi V, Fellows M, Dabdoub SM, Nagaraja HN, O'Donnell B, Deshpande NR, Kumar PS. 2017. A tale of two risks: smoking, diabetes and the subgingival microbiome. *ISME J* 11:2075–2089. <https://doi.org/10.1038/ismej.2017.73>.
57. Faith JJ, Guruge JL, Charbonneau M, Subramanian S, Seedorf H, Goodman AL, Clemente JC, Knight R, Heath AC, Leibel RL, Rosenbaum M, Gordon JL. 2013. The long-term stability of the human gut microbiota. *Science* 341:1237439. <https://doi.org/10.1126/science.1237439>.
58. Zhou Y, Gao H, Mihindukulasuriya KA, La Rosa PS, Wylie KM, Vishnivetskaya T, Podar M, Warner B, Tarr PI, Nelson DE, Fortenberry JD, Holland MJ, Burr SE, Shannon WD, Sodergren E, Weinstock GM. 2013. Biogeography of the ecosystems of the healthy human body. *Genome Biol* 14:R1. <https://doi.org/10.1186/gb-2013-14-1-r1>.
59. Høiby N, Bjarnsholt T, Givskov M, Molin S, Ciofu O. 2010. Antibiotic resistance of bacterial biofilms. *Int J Antimicrob Agents* 35:322–332. <https://doi.org/10.1016/j.ijantimicag.2009.12.011>.
60. Sampaio-Maia B, Caldas I, Pereira M, Pérez-Mongioli D, Araujo R. 2016. The oral microbiome in health and its implication in oral and systemic diseases. *Adv Appl Microbiol* 97:171–210. <https://doi.org/10.1016/bs.aamb.2016.08.002>.
61. Colman G, Beighton D, Chalk AJ, Wake S. 1976. Cigarette smoking and the microbial flora of the mouth. *Aust Dent J* 21:111–118. <https://doi.org/10.1111/j.1834-7819.1976.tb02833.x>.
62. Ertel A, Eng R, Smith SM. 1991. The differential effect of cigarette smoke on the growth of bacteria found in humans. *Chest* 100:628–630. <https://doi.org/10.1378/chest.100.3.628>.
63. Paropkari AD, Leblebicioglu B, Christian LM, Kumar PS. 2016. Smoking, pregnancy and the subgingival microbiome. *Sci Rep* 6:30388–30389. <https://doi.org/10.1038/srep30388>.
64. Mandell LA, Niederman MS. 2019. Aspiration pneumonia. *N Engl J Med* 380:651–663. <https://doi.org/10.1056/NEJMra1714562>.
65. Huttunen R, Heikkinen T, Syrjänen J. 2011. Smoking and the outcome of infection. *J Intern Med* 269:258–269. <https://doi.org/10.1111/j.1365-2796.2010.02332.x>.
66. Stämpfli MR, Anderson GP. 2009. How cigarette smoke skews immune responses to promote infection, lung disease and cancer. *Nat Rev Immunol* 9:377–384. <https://doi.org/10.1038/nri2530>.
67. Arcavi L, Benowitz NL. 2004. Cigarette smoking and infection. *Arch Intern Med* 164:2206–2216. <https://doi.org/10.1001/archinte.164.20.2206>.
68. Van der Veen E, Sanders E, Videler W, Van Staaij B, Van Benthem P, Schilder A. 2006. Optimal site for throat culture: tonsillar surface versus posterior pharyngeal wall. *Eur Arch Otorhinolaryngol* 263:750–753. <https://doi.org/10.1007/s00405-006-0046-6>.
69. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, Walker AW. 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 12:87. <https://doi.org/10.1186/s12915-014-0087-z>.
70. Rani Ram A. 2017. Invasions of the oropharynx: microbiome of healthy and infected respiratory tissue. PhD thesis, University of Glasgow.
71. Joshi NA, Fass JN. 2011. Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33)[Software].
72. Nikolenko SI, Korobeynikov AI, Alekseyev MA. 2013. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. 14:57.
73. Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD. 2012. PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics* 13:31. <https://doi.org/10.1186/1471-2105-13-31>.
74. Schirmer M, Ijaz UZ, D'Amore R, Hall N, Sloan WT, Quince C. 2015. Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Res* 43:e37. <https://doi.org/10.1093/nar/gku1341>.
75. Edgar RC. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* 10:996–998. <https://doi.org/10.1038/nmeth.2604>.
76. 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* 73:5261–5267. <https://doi.org/10.1128/AEM.00062-07>.
77. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>.
78. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>.
79. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. 2018. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 6:1–14.
80. D'Amore R, Ijaz UZ, Schirmer M, Kenny JG, Gregory R, Darby AC, Shakya M, Podar M, Quince C, Hall N. 2016. A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling. *BMC Genom* 17:55. <https://doi.org/10.1186/s12864-015-2194-9>.
81. Gerasimidis K, Bertz M, Quince C, Brunner K, Bruce A, Combet E, Calus S, Loman N, Ijaz UZ. 2016. The effect of DNA extraction methodology on gut microbiota research applications. *BMC Res Notes* 9:365. <https://doi.org/10.1186/s13104-016-2171-7>.
82. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H. 2013. Package 'vegan.' Community ecology package, version 2.
83. Hill MO. 1973. Diversity and evenness: a unifying notation and its consequences. *Ecology* 54:427–432. <https://doi.org/10.2307/1934352>.
84. Anderson MJ. 2001. Permutation tests for univariate or multivariate analysis of variance and regression. *Canadian J Fisheries and Aquatic Sciences* 58:626–639. <https://doi.org/10.1139/f01-004>.
85. Legendre P, Cáceres MD. 2013. Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecol Lett* 16:951–963. <https://doi.org/10.1111/ele.12141>.
86. Dray S, Blanchet G, Borcard D, Guenard G, Jombart T, Legendre P, Wagner H. 2018. Package 'adespatial': multivariate multiscale spatial analysis. See <https://cran.r-project.org/web/packages/adespatial/index.html>.
87. Clarke KR, Ainsworth M. 1993. A method of linking multivariate community structure to environmental variables. *Marine Ecology Progress Series* 92:205–205. <https://doi.org/10.3354/meps092205>.
88. Taylor M. 2015. sinkr: a collection of functions featured on the blog 'Me nugget'. R package version 1.0.
89. Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26:1463–1464. <https://doi.org/10.1093/bioinformatics/btq166>.
90. Stegen JC, Lin X, Konopka AE, Fredrickson JK. 2012. Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J* 6:1653–1664. <https://doi.org/10.1038/ismej.2012.22>.
91. Kassambara A. 2018. Machine learning essentials: practical guide in R. sthda.
92. Lumley T, Miller A. 2009. Leaps: regression subset selection. R Package Version 2:2366.
93. Kuhn M. 2008. Building predictive models in R using the caret package. *J Statistical Software* 28:1–26.
94. Lüdtke D. 2018. sjPlot: data visualization for statistics in social science. R package version 2.
95. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. 2007. The human microbiome project. *Nature Lett* 449:804. <https://doi.org/10.1038/nature06244>.
96. Shetty SA, Hugenholtz F, Lahti L, Smidt H, de Vos WM. 2017. Intestinal microbiome landscaping: insight in community assemblage and implications for microbial modulation strategies. *FEMS Microbiol Rev* 41:182–199. <https://doi.org/10.1093/femsre/fuw045>.
97. Lahti L, Shetty S, Blake T, Salojärvi J. 2017. microbiome R package. *Tools Microbiome Anal R*.
98. Foster ZS, Shapton TJ, Grünwald NJ. 2017. Metacoder: an R package for visualization and manipulation of community taxonomic diversity data. *PLoS Comput Biol* 13:e1005404. <https://doi.org/10.1371/journal.pcbi.1005404>.
99. Chen X, Cohen JE. 2001. Global stability, local stability and permanence in model food webs. *J Theor Biol* 212:223–235. <https://doi.org/10.1006/jtbi.2001.2370>.
100. Townsend SE, Haydon DT, Matthews L. 2010. On the generality of stability–complexity relationships in Lotka–Volterra ecosystems. *J Theor Biol* 267:243–251. <https://doi.org/10.1016/j.jtbi.2010.08.018>.
101. Coyte KZ, Schluter J, Foster KR. 2015. The ecology of the microbiome: networks, competition, and stability. *Science* 350:663–666. <https://doi.org/10.1126/science.aad2602>.
102. Rasiah IA, Wong L, Anderson SA, Sissons CH. 2005. Variation in bacterial DGGE patterns from human saliva: over time, between individuals and in corresponding dental plaque microcosms. *Arch Oral Biol* 50:779–787. <https://doi.org/10.1016/j.archoralbio.2005.02.001>.

103. Faust K, Bauchinger F, Laroche B, Buyl SD, Lahti L, Washburne AD, Gonze D, Widder S. 2018. Signatures of ecological processes in microbial community time series. *Microbiome* 6:1–13.
104. Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, Raes J, Huttenhower C. 2012. Microbial co-occurrence relationships in the human microbiome. *PLoS Comput Biol* 8:e1002606. <https://doi.org/10.1371/journal.pcbi.1002606>.
105. Jansen VA, Kokkoris GD. 2003. Complexity and stability revisited. *Ecology Lett* 6:498–502. <https://doi.org/10.1046/j.1461-0248.2003.00464.x>.
106. Wang YH, Gutierrez AP. 1980. An assessment of the use of stability analyses in population ecology. *The J Animal Ecology*:435–452. <https://doi.org/10.2307/4256>.
107. Berlow EL, Neutel AM, Cohen JE, Rüter PCD, Ebenman BO, Emmerson M, Fox JW, Jansen VA, Jones JJ, Kokkoris GD. 2004. Interaction strengths in food webs: issues and opportunities. *The J Animal Ecology* 73:585–598. <https://doi.org/10.1111/j.0021-8790.2004.00833.x>.
108. Haydon D. 1994. Pivotal assumptions determining the relationship between stability and complexity: an analytical synthesis of the stability-complexity debate. *The Am Nat* 144:14–29. <https://doi.org/10.1086/285658>.
109. Anderson HM, Hutson V, Law R. 1992. On the conditions for permanence of species in ecological communities. *The Am Nat* 139:663–668. <https://doi.org/10.1086/285350>.
110. Jansen VA, Sigmund K. 1998. Shaken, not stirred: on permanence in ecological communities.
111. Law R, Morton RD. 1996. Permanence and the assembly of ecological communities. *Ecology* 77:762–775.