The ISME Journal (2010) 4, 839–851

www.nature.com/ismej

ORIGINAL ARTICLE

A poke into the diversity and associations within human anterior nare microbial communities

Melissa L Wos-Oxley[1], Iris Plumeier[1], Christof von Eiff[2,4], Stefan Taudien[3], Matthias Platzer[3], Ramiro Vilchez-Vargas[1], Karsten Becker[2] and Dietmar H Pieper[1]
[1]*Microbial Interactions and Processes Research Group, Department of Medical Microbiology, Helmholtz Centre for Infection Research, Braunschweig, Germany;* [2]*Institute of Medical Microbiology, University Hospital, Münster, Germany and* [3]*Genome Analysis, Leibniz Institute for Age Research - Fritz Lipmann Institute, Jena, Germany*

The anterior nares are the major reservoir for *Staphylococcus aureus* in humans, where nasal carriage has a crucial function as a source for invasive infections. Despite various investigations on aerobic community members based on traditional cultivation methods, little is known on the overall microbial composition and complex *in situ* interactions, but such knowledge is highly warranted for effective *S. aureus* control strategies. As assessed using advanced culture-independent approaches, this study provides a comprehensive survey of the anterior nare bacterial community of 40 individuals. Previously undiscovered co-colonization patterns and natural variations in species composition were revealed and provide insights for future intervention strategies for the control of health-care- and community-associated *S. aureus* infections.
The ISME Journal (2010) **4,** 839–851; doi:10.1038/ismej.2010.15; published online 25 February 2010
Subject Category: microbial population and community ecology
Keywords: anterior nares; bacterial associations; microbial diversity; 454-pyrosequening; SSCP; *Staphylococcus aureus*

Introduction

The human anterior nares are the principle habitat for *Staphylococcus aureus*, which constitute to the normal and asymptomatic microbial community. Approximately 20% of humans are persistently colonized by this pathogen (Williams, 1963; van Belkum *et al.*, 2009). Although *S. aureus* colonization of the nares is asymptomatic, nasal carriage has a crucial function as a source of invasive infections in both community and hospital settings (von Eiff *et al.*, 2001) Therefore, elimination of *S. aureus* nasal carriage seems to be one of the most straightforward strategies to prevent *S. aureus* infections (van Rijen *et al.*, 2008).

Differences in colonization by *S. aureus* have been attributed to host factors such as host immunity, age, gender and/or environmental factors (García-Rodríguez and Fresnadillo Martínez, 2002). However, interactions between microorganisms themselves may also influence which of the species are able to persist (Uehara *et al.*, 2000). As an example, converse associations have been suggested between *S. aureus* and *S. pneumoniae* (Regev-Yochay *et al.*, 2004). In addition, a decreased incidence of *S. aureus* colonization was observed in individuals heavily colonized by *Corynebacterium* spp. (Uehara *et al.*, 2000). Such studies have been traditionally performed using culture-dependent methods, and thus have concentrated on the interaction between a few bacterial species assumed to be important and abundant within the nares.

In recent years, special attention has been directed towards using advanced culture-independent methods to study the bacterial diversity in humans. Various reports are now available on the microbiota inhabiting the human skin (Dekio *et al.*, 2005; Gao *et al.*, 2007; Fierer *et al.*, 2008). Two recent reports have given a broad overview on the diversity at distinct skin sites of a limited number (<10) of healthy humans (Costello *et al.*, 2009; Grice *et al.*, 2009), including an initial assessment of the bacterial diversity of the nares as a comparison to these other sites.

In this report, we characterized the bacterial diversity of the anterior nares of 40 individuals using 16S rRNA gene fingerprinting tools, then chose a further 6 representative individuals for a more rigorous evaluation of diversity using 454-pyrosequencing. With such approaches, we were able to make an in-depth assessment of the microbiota of the nares and their distributions in

Correspondence: DH Pieper, Microbial Interactions and Processes Research Group, Department of Medical Microbiology, Helmholtz Centre for Infection Research, Inhoffenstr. 7, Braunschweig 38124, Germany.
E-mail: dpi@helmholtz-hzi.de
[4]Current address: Wyeth Pharma GmbH, Wienburgstr. 207, Münster D-48159, Germany.
Received 2 November 2009; revised 18 January 2010; accepted 24 January 2010; published online 25 February 2010

respect to both converse and congregating relationships between those community members that reside within this niche.

Materials and methods

Subjects

The nasal microbial communities of 40 healthy volunteers (subjects with no history of *S. aureus* disease and no current antibiotic use) from two working locations in Germany (Braunschweig, $n=19$ and Münster, $n=21$) were sampled. In a preliminary test, all volunteers from Münster were screened for nasal colonization by *S. aureus*, as previously described (Becker *et al.*, 2006), to include in the study a significant amount of *S. aureus* carriers. The mean age of the volunteers (16 male, 24 females) was 37 years (19–57 years, with the majority of 23 volunteers aged between 26 and 35 years). The standard cotton swabbing technique was used to sample the anterior nares, and samples were immediately stored at −20 °C until further DNA extraction or directly processed for RNA extraction.

Nucleotide extraction

DNA was extracted from the whole swab using the FastDNA Spin Kit for Soil (MP Biomedicals, Solon, OH, USA). In addition, RNA from 3 of the 40 volunteers was simultaneously extracted along with the DNA using the AllPrep DNA/RNA Mini Kit (Qiagen, Hilden, Germany). Directly following swabbing, the RNA was stabilized with 200 µl of RNA protect Bacteria Reagent (Qiagen). In brief, the swabs were placed into Lysing Matrix B tubes (MP Biomedicals) containing 500 µl of buffer RLT (reconstituted with β-mercaptoethanol according to the manufacturer and cells disrupted in a Fast Prep-24 instrument (40 s, intensity 5.5). RNA and DNA components were simultaneously separated following the manufacturer's instructions. Potential DNA contamination of the RNA was digested with RNase-Free DNase (Qiagen). SuperScript III Reverse transcriptase (Invitrogen, Karlsruhe, Germany) was used for cDNA synthesis according to the instructions of the provider. As a negative control, a sterile cotton swab was processed in the same manner. Nucleotides were quantified using the NanoDrop spectrophotometer (Thermo Scientific, Waltham, MA, USA).

SSCP analysis

The bacterial diversity was analyzed using single-strand conformation polymorphism (Schwieger and Tebbe, 1998). As the Com-1 (forward) and Com2-Ph (reverse) primers that produce an approximately 410 bp fragment of the eubacterial 16S rRNA gene enclosing the variable regions IV and V do not differentiate between different *Staphylococcus* strains, the universal eubacterial primers 27F, 338R and 521R were evaluated (Supplementary Table 1). In addition, primer 63F previously described as superior for analyzing the bacterial diversity in environmental samples (Marchesi *et al.*, 1998) was modified (now referred to as 64F) to target without mismatch >85% of bacterial sequences deposited in the RDP database (http://www.rdp.cme.msu.edu/). Both forward and reverse primers were applied in all four possible combinations, and both a phosphorylated forward and a phosphorylated reverse primer were tested by amplifying 10 ng template DNA. Double-stranded PCR products were purified with QIAquick (Qiagen). Single-stranded DNA was produced by lambda exonuclease digestion and purified and quantified as previously described (Witzig *et al.*, 2006). For fingerprinting analysis, 100 ng of single-stranded DNA was mixed with gel loading buffer (95% formamide, 10 mM NaOH, 0.25% bromphenol blue, 0.25% xylene cyanol) in a final volume of 5 µl. After incubation for 3 min at 95 °C, the single-stranded DNA samples were stored on ice, loaded onto an MDE gel (0.6 × MDE gel solution; Cambrex Bio Science, Rockland, ME, USA), electrophoretically separated at 20 °C at 400 V for 18 h on a Macrophor sequencing apparatus (GE Healthcare, München, Germany) and silver stained (Schwieger and Tebbe, 1998).

As the combination of 64F and 521RP (phosphorylated) primers gave the most optimal resolution of an artificial mixture of PCR products from *S. aureus* Newman DSM1104, *S. warneri* DSM20316, *S. cohnii* DSM6718 and *S. simulans* DSM20322, SSCP was performed with amplicon mixtures generated by these primers. The SSCP images were read and the intensities of bands in each line were quantified by ImageJ software (http://www.rsb.info.nih.gov/ij/; Bethesda, MD, USA). Bands were excised and DNA recovered as previously described (Witzig *et al.*, 2006). DNA was reamplified using 64F/521RP primers with an annealing temperature of 55 °C. The purified PCR products were used as DNA templates in independent sequencing reactions using the 64F and the 521RP primers applying the BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) on an ABI PRISM 3100 Genetic Analyser (Applied Biosystems). Sequencing trace files were checked for quality using Sequencher (v.4.0.5, Gene Codes, Ann Arbor, MI, USA), giving an overall of 182 high-quality sequences.

Corynebacterial diversity

SSCP fragments of *Corynebacterium tuberculostearicum, C. pseudodiphtheriticum, C. accolens* and *C. macginleyi* were prepared from clone library sequences (see below) by a nested PCR approach using M13 primers followed by 64F/521RP primers. To evaluate corynebacterial diversity in nasal samples, PCR fragments were generated by nested PCR using primers 27F and Co667 with the later designed to target a phylogenetic group inside the

corynebacterial genus comprising *C. accolens, C. confusum, C. macginleyi, C. propinquum, C. pseudodiphtheriticum* and *C. tuberculostearicum* (annealing temperature 58 °C). One μl of the purified PCR product was subjected to a second PCR using the primers 64F/521RP followed by SSCP analysis as described above.

Clone libraries

DNA extracts of the nasal communities of two volunteers were used as template in PCR amplifications with primers 27F and 1389R (Supplementary Table 1) at an annealing temperature of 55 °C. Purified PCR products were cloned into the pGEM-T easy vector system (Promega, Mannheim, Germany) and amplified by PCR with vector-specific M13 primers on transformant colonies dissolved in water and previously incubated at 95 °C for 10 min. The purified PCR products were used as DNA templates in independent sequencing reactions using the 27F primer. Corynebacterial sequences representing *C. accolens, C. macginleyi, C. pseudodiphtheriticum* or *C. tuberculostearum* were retrieved from DNA extracts of volunteers by PCR using primers 27F and Co1475R with the latter designed to target members of the corynebacterial genus, followed by cloning of purified PCR products as described above. Appropriate clones were identified by sequencing and sequences deposited in the GenBank database under accession number GU74963–GU74966.

454-Pyrosequencing

In order to multiplex the 16S rRNA gene sequence analysis from six samples, 16S rRNA gene fragments were amplified using the universal 64F and 338R primers. The forward primer was modified to include the 454 Life Sciences Amplicon primer A (454 Life Sciences, Branford, CT, USA) at its 3′ end, and the reverse primer linked to the Amplicon primer B by a bridge consisting of 1 of 6 different 4-base barcodes and a CA linker (Hamady *et al.*, 2008). Each sample was assigned a different barcoded reverse primer before amplification (Supplementary Table 1).

PCR products were quantified by staining with PicoGreen fluorescent dye (Invitrogen), pooled in equimolar ratios, diluted and subjected to emPCR following the GS-FLX emPCR protocol for amplicons (Roche Diagnostics, Branford, CT, USA, December 2007) using the GS emPCR kit III (primer B) and sequenced by primer B on 1 lane of a 70×75 picotiterplate (GS-FLX, 16/16 gasket) resulting in 13 456 reads with an average length of 235 bp.

Initial sequence processing, alignment, clustering and phylotype classification

All pyrosequencing reads were initially processed through RDPs pyrosequencing pipeline (http://www.pyro.cme.msu.edu/). Raw reads were thereby sorted into those from each of the original samples, the tag and primers trimmed off and sequences of low quality removed. All reads shorter than 200 bases were deleted from the analysis. Sequences were aligned using RDP's pyrosequencing INFERNAL aligner (Nawrocki and Eddy, 2007) and grouped into phylotypes using the RDP complete linkage clustering tool at a distance cutoff of 0.02. The most abundant sequence type of each cluster was chosen manually as the representative sequence, and all representative sequences were aligned using MUSCLE, Cambridge, UK (Edgar, 2004). Further clustering was performed after estimating the evolutionary divergence through MEGA4 (Tamura *et al.*, 2007) by merging all sequence clusters that had a nucleotide difference of $\leqslant 2$ bases into a single cluster represented by the most dominant representative sequence. Finally, all single reads showing $\leqslant 4$ base differences were also grouped into these clusters.

Sequences obtained from clone libraries and excised SSCP bands showing $>99\%$ sequence identity were also grouped as 'phylotypes.' Finally, representative sequences obtained from pyrosequencing, clone libraries and SSCP were compared and defined as the same phylotype in case they differed by $<1\%$ over the common sequence stretch (Supplementary Table 2).

The most closely related 16S rRNA gene sequences were in all cases identified in the RDP-II database using the Seqmatch program against sequences obtained from validly described type strains and isolates. Alignments were performed using MUSCLE and edited using SEAVIEW, Villeurbanne Cedex, France (Galtier *et al.*, 1996). Phylogenetic analyses were performed with MEGA4 (Tamura *et al.*, 2007) using the neighbor-joining method with Jukes-Cantor correction and pairwise deletion of gaps/missing data. A total of 1000 bootstrap replications were performed to test for branch robustness.

Statistical analysis

Non-parametric multivariate statistical analysis was performed using PRIMER (v.6.1.6, PRIMER-E, Plymouth Marine Laboratory, Plymouth, UK) (Clarke, 1993; Clarke and Warwick, 2001). All multivariate routines were computed on standardized (the relative abundance of each species out of the total abundance of all species present within a sample equating to 100%), square root transformed abundance data. A sample-similarity matrix was generated using the Bray-Curtis coefficient by comparing the transformed abundances of each of the 30 phylotypes detected by SSCP in regard to every pairwise combination of all samples. Differences between anterior nare bacterial communities of the 40 volunteers as deduced by 16S rRNA gene profiling, where 6 volunteers were sampled twice and 1 volunteer was sampled thrice, the simultaneous 16SrRNA profile of 3 volunteers and the

community structure as deduced by 454-pyrosequencing (6 volunteers) or through clone libraries (2 volunteers), totaling to 59 community structures was explored by ordination using non-metric multidimensional scaling (50 random restarts and 999 iterations). A stress value below 0.2 corresponds to a useful ordination (Clarke and Warwick, 2001). Superimposing clusters at different similarity levels (calculated by the Bray-Curtis algorithm) are used to show which samples group together at certain % similarity levels (as determined by group-average agglomerative hierarchical clustering). Bubble plots visualize the relative abundance of a particular species among sample points on the ordination. Vector overlays visualize potential linear and monotonic relationships between a given set of species. Each vector was determined using multiple partial correlation algorithm, which takes into account the correlation of all other species. Vector length and direction indicates the strength and sign of the relationship between species and the orientation of samples.

Phylotype diversity and richness of the six volunteers used for 454-pyrosequencing were further explored using *k*-dominance plots and univariate measures using PRIMER. *k*-dominance curves are cumulative ranked abundances plotted against species rank. Total phylotypes (S) refers to the number of phylotypes present within each sample of a total of 65 phylotypes classified using the 454-pyrosequencing method. *Phylotype richness*, also known as *Margalef's index* (d), is a measure of the number of phylotypes present for a given number of individuals and is calculated as $d=(S-1)/\log N$. *Pielou's evenness* (J') is an equitability measure, which expresses how evenly the individuals are distributed among the different phylotypes and is calculated as $J'=H'/H'_{\max}=H'/\log S$. *Shannon diversity* is calculated as $H'=-\sum_i p_i \ ln(p_i)$, where p_i is the relative abundance of the ith phylotypes.

Non-random associations between pairs of phylotypes were assessed using the PAIRS program on presence/absence data (Ulrich, 2008). Only relationships between phylotypes that were present in at least 10% and no more than 80% of samples were evaluated (yielding a list of 11 available phylotypes and thus 55 pairwise comparisons). Random matrices for generating standardized scores (C-score) (Stone and Roberts, 1990) and significance levels (P-values) were obtained using the fixed row (phylotypes)—fixed column (anterior nare samples) constraints algorithm, and 100 random matrices were computed from the binary data matrix. Significant species underdispersion or overdispersion results in Z-transformed scores (observed C-score—expected C-score/standard deviation) above 1.96 or below −1.96 (at the 5% error level) (Ulrich and Zalewski, 2006). False detection error rates were corrected for by a sequential Bonferroni correction (BY criterion) (Benjamini and Yekutieli, 2001).

Results

Optimizing SSCP to fingerprint anterior nare bacterial communities

SSCP analysis of 16S rRNA gene fragments is typically performed using an approximately 410 bp fragment of the eubacterial 16S rRNA gene enclosing the variable regions IV and V. However, this fragment does not differentiate between different staphylococcal species, particularly *S. epidermidis* and *S. aureus* (Supplementary Figure 1B), known to be highly prevalent in the anterior nares (as also confirmed by cultivation of 20 nasal swabs, data not shown). In contrast, targeting the variable regions I and II showed high discriminatory power (Supplementary Figure 1C), and *S. aureus* and *S. epidermidis* could be successfully separated from mixed communities using SSCP fragments generated with Primers 64F and 521RP (Figure 1). Some 16S rRNA gene fragments generated multiple bands, such as those of organisms related to *Peptoniphilus* sp. (triple band) or *Moraxella lacunata* (quintuple band). 16S rRNA gene fragments originating from corynebacterial species resulted in rather broad bands difficult to quantify when present in low abundance. Quantification of these species was optimized by using specific primers targeting *C. accolens, C. macginleyi, C. tuberculostearicum*, *C. propinquum* and C. *pseudodiphtheriticum* (indicated by SSCP as present in the anterior nares) and a nested PCR approach. Only fragments originating from *C. accolens* and *C. macginleyi* could not be

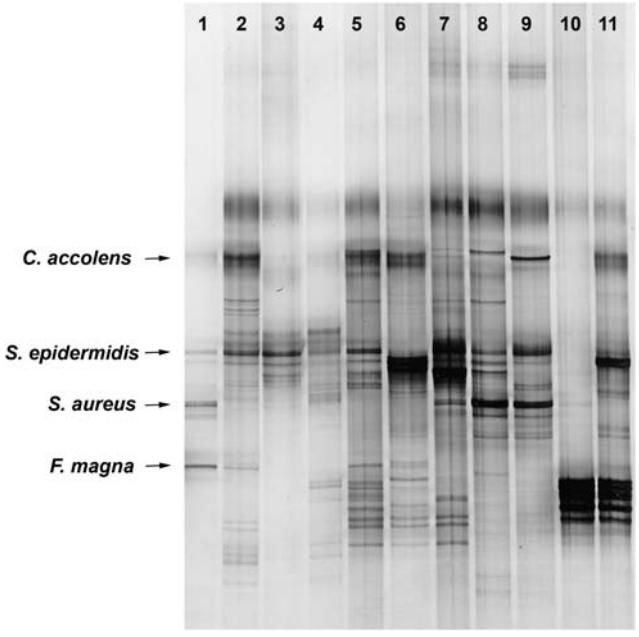

Figure 1 SSCP analysis of 16S rRNA gene segments generated from microbial communities from the anterior nares. Lane 1 shows separation of an artificial mixture of fragments obtained from *C. accolens, S. epidermidis, S. aureus* and *F. magna*. Lanes 2–11 show profiles obtained from 10 selected volunteers. An empty lane between lane 2 and lane 3 was eliminated by Adobe Photoshop CS3.

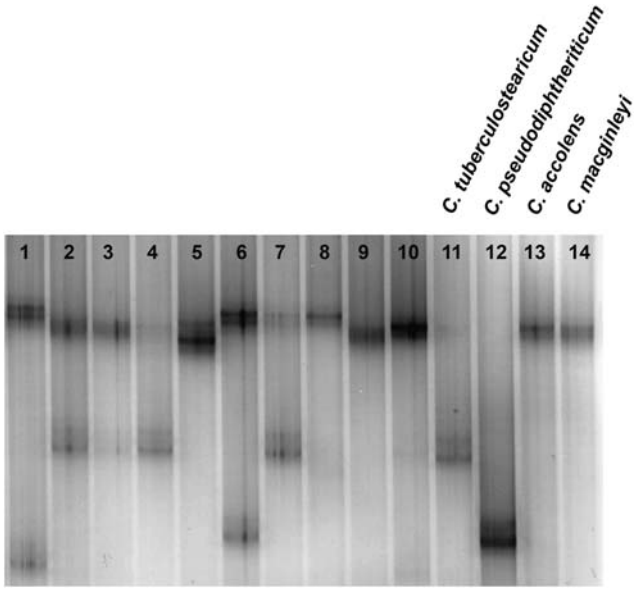

Figure 2 SSCP analysis of 16S rRNA gene segments generated from microbial communities from the anterior nares after pre-enrichment of corynebacterial 16S rRNA genes through PCR. Lanes 1–10 show profiles obtained from 10 selected volunteers. Lanes 11, 12, 13 and 14 show electrophoresis of PCR-SSCP fragments generated from *C. tuberculostearicum, C. pseudodiphtheriticum, C accolens* and *C. macginleyi,* respectively. Slight differences in running behaviors of fragments indicating the presence of organisms related to *C. accolens/C. macginleyi* are due to 1–2 base differences compared with the *C. accolens* segment separated in lane 13. The lower band observed in lane 1 is due to the presence of organisms related to *C. propinquum* as verified by sequencing.

separated by this approach (Figure 2). However, sequencing of SSCP bands indicated bands with such electrophoretic behavior to be typically originating from organisms related to *C. accolens*, as verified by 454-pyrosequencing (see below). The nested PCR approach could be validated through 454-pyrosequencing, in which the abundance of corynebacterial phylotypes quantified in six volunteers matched with that deduced after the nested PCR approach (see below).

Anterior nare microbial community structure as revealed by SSCP

The microbial communities of the 40 volunteers, revealed by SSCP analysis, comprised 30 bacterial phylotypes (Figure 3 and Supplementary Table 2) belonging to 4 bacterial phyla and 13 families (Figure 4). The most frequently occurring phylotypes (>85% of volunteers) were those related to *C. accolens/C. macginleyi*, *S. epidermidis* and *Propionibacterium acnes*, belonging to genera previously reported as abundant in the anterior nares (Costello *et al.*, 2009). Phylotypes related to *Dolosigranulum pigrum*, *Finegoldia magna*, *Peptoniphilus* sp. gpac 121, *Propionibacterium granulosum* and a hitherto uncultured actinomycete were observed in 40–60% of volunteers, thus suggesting that this set of species constitute the core nasal community. *S. aureus* was detected in 25% of volunteers.

Corynebacterial and propionibacterial species were prevalent in high abundance, whereas members of the proteobacteria typically comprised <10% of the whole community (Figure 4). Three volunteers, however, were inundated by an organism related to *M. lacunata*. The high abundance of *Lactobacillales* was mainly due to the presence of *Dolosigranulum* sp. (family *Carnobacteriaceae*). Interestingly, *Clostridiales,* belonging to the family Incertae sedis XI, were observed in 55% of the volunteers revealing, as previously indicated (Costello *et al.*, 2009), a highly ubiquitous presence. Comparison of the communities of all 40 volunteers (Figure 5) revealed 10 distinct clusters at 50% similarity, with the majority of volunteers (36 of 40) grouping within 5 of these clusters. The community composition of individuals at different time points showed high similarity in most cases.

The resolution power of SSCP as an appropriate means for evaluating diversity was validated against two small clone libraries, comprising 94 or 87 clones, prepared from volunteer 22 and 28, respectively. Eighty-five and 76 clones, respectively, could be assigned to the phylotypes detected during SSCP analysis. Further clones typically represented single sequence types (Figure 3) and can be assumed to be only of low abundance within the total community. The global community structures as assessed by both methods showed >85% similarity (Figure 5).

Coaggregation of species

Analysis of phylotype distribution across anterior nare communities indicated that species were not completely randomly distributed. Species vectors presented in Figure 5b indicate whether species share similarities/differences in their distribution and abundance. Those vectors representing *F. magna* and *Peptoniphilus* sp. share the same direction, indicating that their distribution and abundance across all samples was similar and possibly symbiotic. This relationship was confirmed by null hypothesis testing methods assessing the joint presence of both species across samples other than by chance using PAIRS, where *F. magna and Peptoniphilus sp.* showed a statistically significant co-occurrence ($P<0.05$, Table 1). The vectors representing *C. pseudodiphtheriticum* and *S. aureus* pointed in opposing directions to those vectors of *F. magna* and *Peptoniphilus* sp., indicating that these species have a substantially different distribution and abundance pattern and thus reflect their inability to share the same niche (Figure 5b). This was further confirmed when the two pairs; *F. magna* and *S. aureus*, and *F. magna* and *C. pseudodiphtheriticum* showed a statistically significant avoidance to each other in the PAIRS analysis ($P<0.05$,

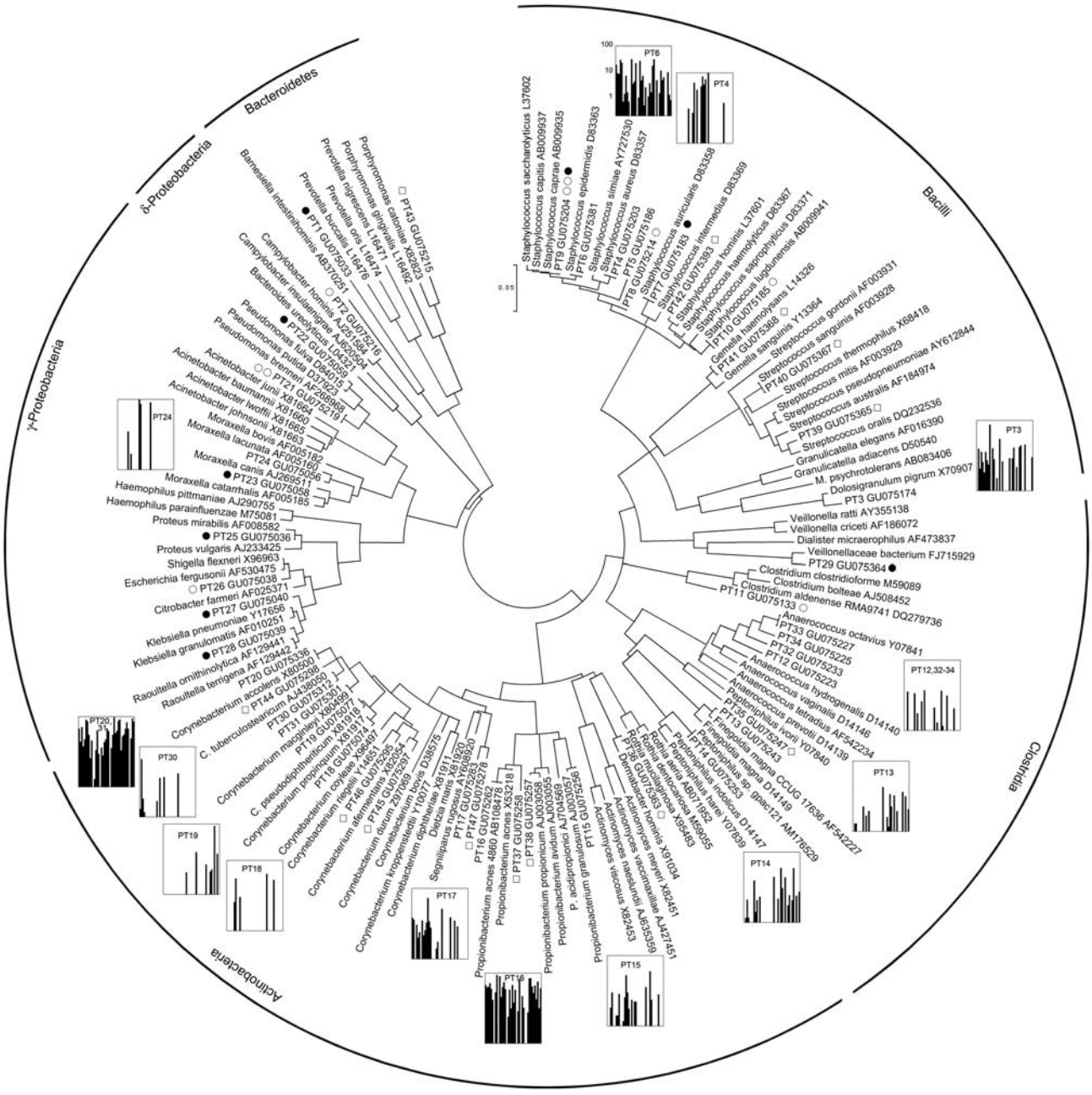

Figure 3 Taxonomic distribution of the 16S rRNA gene sequences retrieved by SSCP profiling of nasal microbial communities. 16S rRNA gene sequences retrieved by the preparation of clone libraries are also indicated. Sequences showing >99% sequence identity were grouped as 'phylotypes' and representative phylotypes are indicated. Phylotypes observed by SSCP profiling in only up to three volunteers are indicated by circles. An open circle indicates an abundance of the phylotype of <5%, and a filled circle of >5%. If a phylotype was observed in more than three volunteers, its abundance in each of the 40 volunteers is indicated by bar graphs. Abundance is given at a logarithmic scale from 0.2% to 100% (see phylotype 6). Phylotypes observed in the clone libraries are indicated by squares. The evolutionary history was inferred using the Neighbor-Joining method. The evolutionary distances were computed using the Jukes-Cantor method and are in the units of the number of base substitutions per site. Sequences of closely related bacterial type strains and isolates were included in the analysis. The different bacterial families identified are separated by dashed lines. GenBank accession numbers are given after the strain name or the phylotype.

Table 1). The avoidance pattern of *S. aureus* and *F. magna* is further illustrated by bubble plots in Figure 5c. There are some additional pairs of species that may share a niche at times but show a converse relationship in their abundance, such as *C. accolens* to *S. aureus, P. acnes* and *S. epidermidis* (Figure 5b and Supplementary Figure 2).

Comparing the active community to the overall community

As the prevalence and abundance of strictly anaerobic bacteria in the anterior nare microbial communities was rather astonishing, we analyzed if those species were active by comparing the 16S rRNA gene diversity with the 16S rRNA diversity

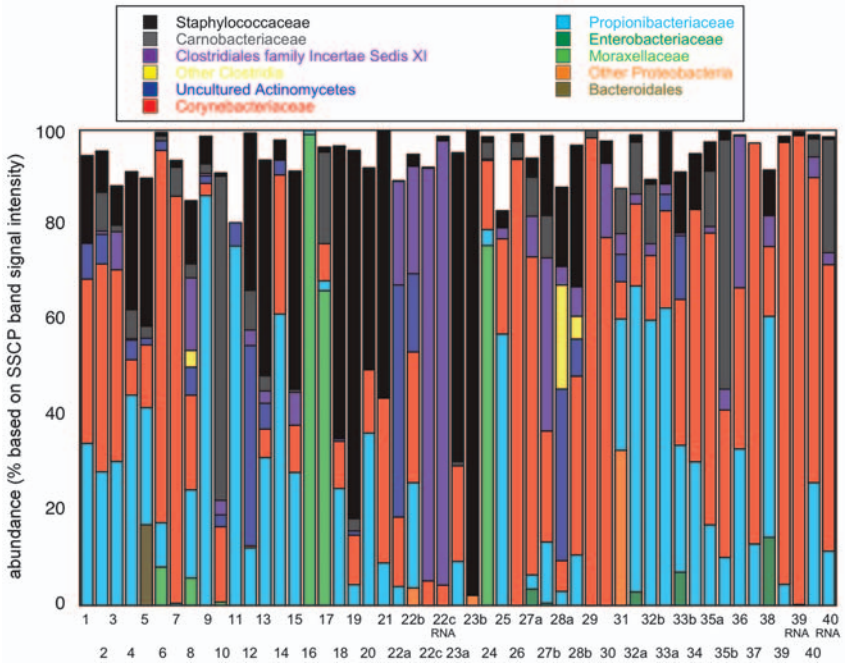

Figure 4 Bacterial community composition of the anterior nares from 40 volunteers as assessed by PCR-SSCP. Phylotypes were grouped into families and the relative abundance of the family members is given. If multiple samples were taken over time from one volunteer, this is indicated by a, b or c. RNA indicates the community composition as assessed by profiling of 16S rRNA instead of the 16S rRNA genes.

in three selected volunteers. The community structures defined by comparing the RNA profile (representing the active community) with the DNA (representing the overall community) profile was almost identical (Figures 4 and 5a). Specifically, *F. magna, Peptoniphilus* sp., *Anaerococcus* sp. and *Dolosigranulum* sp. were all observed to be in an equivalent or greater abundance in the active community.

Anterior nare microbial community structure as revealed by 454-pyrosequencing

Six anterior nare microbial communities were selected for 454-pyrosequencing to reveal the fine-scale community structure. In total, 13 007 reads (96.7%) were assigned to the different samples by their correct barcodes. After removal of low-quality sequences and sequences of a length <200 bases, a total of 9953 sequences remained, of which an additional 15 were discarded from further analyses as they showed the presence of chloroplast DNA. For a detailed analysis of the community structure, only sequence types observed at least twice were further analyzed. By this procedure, 93 sequences (0.9%) were discarded. The remaining 9845 sequences (1725; 2142; 1442; 2079; 1075; and 1478 sequences, from volunteers number 3; 10; 13; 15; 17; and 18, respectively) were classified into 65 phylotypes (Figure 6 and Supplementary Table 2) belonging to 23 families and 4 phyla (Figure 7).

Phylotypes observed in an abundance of $>1\%$ by 454-pyrosequencing were also typically observed by SSCP profiling. Only three phylotypes (phylotype 72, 73 and 77) observed by 454-pyrosequencing once or twice in an abundance of 1.5–1.7% had not been identified by SSCP profiling, indicating that sequence types present in abundance of $>2\%$ can be reproducibly identified by SSCP profiling. Accordingly, the community structures as revealed by SSCP versus 454-pyrosequencing typically showed similarities $>80\%$ (Figure 5a).

Using 454-pyrosequencing with higher resolution power to detect low abundance species, hitherto uncultured members of the *Neisseriales* order were observed in five of the six volunteers in an abundance of 0.1–1.7%, indicating that they may be putative core community members. Furthermore, microorganisms related to *S. capitis/S. caprae* or *S. lugdunensis* were found to occur typically in low abundance of 0.1–0.2%. Overall, all six communities were highly dominated by just a few species and have relatively low diversity. This is reflected by the phylotype richness (d) ranging from 2.60 to 7.62, an evenness index (J') ranging between 0.620 and 0.476 and a Shannon diversity index (H') between 2.224 and 1.222 (Supplementary Figure 3).

Discussion

In recent years, we have seen particular focus towards extending our knowledge on the bacterial

diversity within humans using new culture-independent technologies. In particular, various reports are now available on the microbiota inhabiting the human skin (Dekio *et al.*, 2005; Gao *et al.*, 2007; Fierer *et al.*, 2008; Grice *et al.*, 2008), in which the most recent reports gave a broad overview on the diversity of distinct skin sites of up to 10 healthy humans (Costello *et al.*, 2009; Grice *et al.*, 2009). However, an in-depth characterization of our anterior nares is still lacking.

Our detailed characterization of the nares revealed the fine-scale diversity of the nares and unraveled novel associations between species. There was a significant occurrence of members of the Gram-positive anaerobic cocci (GPAC), originally classified in the genus *Peptostreptococcus*. Members of the currently termed family Insertae sedis XI, such as *Anaerococcus* sp., *F. magna* and *Peptoniphilus* sp. could be identified as core members of anterior nare microbial communities, however, only a subset of volunteers harbored these species in the anterior nares. The dominant phylotype 14 representing *Peptoniphilus* was most closely related to a yet to be described genospecies designated *Peptoniphilus* sp. gpac 121, isolated from the pus of a human neck injury (Wildeboer-Veloo *et al.*, 2007). In contrast to previous analyses of a small set of volunteers (Costello *et al.*, 2009), our study showed also a high abundance of organisms related to *D. pigrum* of the *Carnobacteriaceae,* recently reported in human respiratory diseases (Lécuyer *et al.*, 2007) suggesting that they are part of the core nasal community. Although most of the abundant phylotypes observed in anterior nare microbial communities correspond to members of genera that have already been isolated, a group of uncultured actinomycetales (phylotypes 17 and 47) were observed to be abundant. In addition, we describe for the first time the actual distribution of corynebacterial species in anterior nare microbial communities. Only sequences indicating the presence of organisms related to *C. accolens, C. macginleyi, C. propinquum, C. pseudodiphtheriticum* and *C. tuberculostearicum* were observed at high abundances, with *C. accolens* to be ubiquitous in anterior nare microbial communities.

A range of interactions between anterior nare microorganisms have been previously suggested based on culture-dependent studies. Epidemiological investigations have indicated that co-colonization by

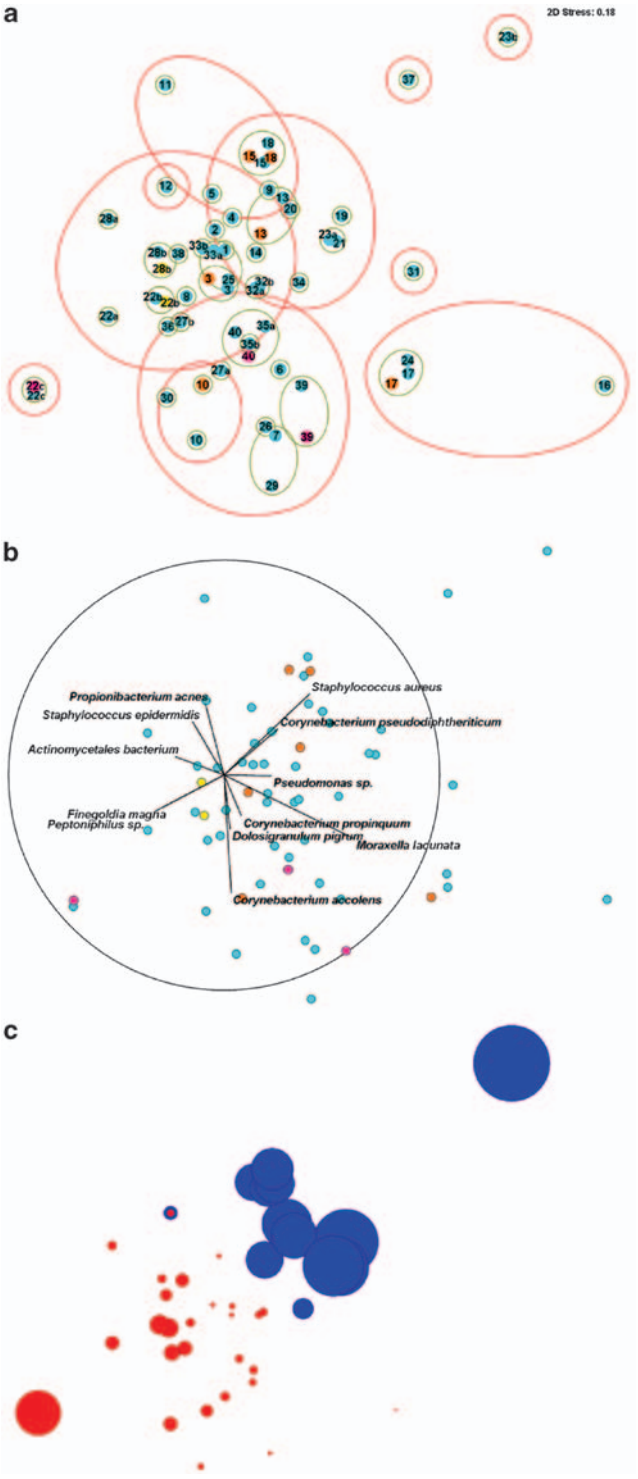

Figure 5 Non-metric multidimensional scaling plot illustrating similarities in the global bacterial community structure of the anterior nares of 40 volunteers. Numbers 1–40 correspond to individual volunteers using SSCP data from extracted DNA (shown in blue). RNA was simultaneously extracted with the DNA of three samples (shown in pink). A further six volunteers were sampled twice and one volunteer was sampled thrice (denoted after the volunteer number as a, b or c). DNA from six volunteers was profiled by 454-pyrosequencing (shown in orange) and that of two volunteers by clone libraries (shown in yellow). The stress value of 0.18 corresponds to a useful two-dimensional interpretation. (**a**) Clusters outlined in red represent samples that group together at 50% similarity, whereas clusters outlined in gray represent 80% similarity. (**b**) Vectors represent the direction of species distribution across the samples, taking into account the correlation of all other species present. Only vectors with a length >0.2 are shown. (**c**) Bubble plots represent the relative abundance of *S. aureus* and *F. magna* (shown as blue and red circles, respectively).

Table 1 Co-occurrence of phylotypes in anterior nare microbial communities

Phylotype 1–phylotype 2	*Abundance* (n = *40*)				
	Phylotype 1	*Phylotype 2*	*Co-occurrence*[a]	*Z-Score*[b]	P-*value*
Peptoniphilus sp.–*F. magna*	18	15	14	−2.19	0.0282
F. magna–*S. aureus*	15	10	1	2.66	0.0076
F. magna–*C. pseudodiphtheriticum*	15	8	0	2.65	0.0078

[a]The number of samples in which presence of both phylotypes was observed is given.
[b]A Z-score below −1.96 indicates significant co-occurrence of the respective two phylotypes, whereas a Z-score above 1.96 indicated significant avoidance.

S. aureus and *S. pneumoniae* is negatively correlated in the nasopharynx (Bogaert *et al.*, 2004). However, whereas *Streptococcus* species are dominant in the nasopharynx (Pettigrew *et al.*, 2008) and various skin habitats (Gao *et al.*, 2007; Fierer *et al.*, 2008), this study shows that they are only in low abundance in the anterior nares, a feature also supported by previous culture-dependent studies (Lina *et al.*, 2003).

There is some dispute regarding the interactions between *S. aureus* and corynebacterial species and *S. aureus* with coagulase-negative staphylococci. While one study reported on a reduced incidence of *S. aureus* in individuals heavily colonized by *Corynebacterium* spp. (Uehara *et al.*, 2000), another study (Lina *et al.*, 2003) observed a high incidence of members of the *Corynebacterium* genus even in *S. aureus* carriers. The findings of the latter study are reflected here where a positive correlation was observed between *C. pseudodiphtheriticum* and *S. aureus*. However, although *C. accolens* can share a niche with *S. aureus, P. acnes* and *S. epidermidis*, we observed a negative correlation in their abundances, suggesting that a higher prevalence of *C. accolens* controls the abundance of these other species. Thus, potential interactions clearly need to be analyzed with a higher level of resolution, particularly as an interaction has only been suggested between corynebacteria and *S. aureus* strains harboring the *agr-1*$_{sa}$ allele (Lina *et al.*, 2003).

In contrast, a statistically significant converse association between *S. aureus* and *F. magna* was found. As culture-dependent studies usually focus on the aerobic community (Lina *et al.*, 2003), the contribution of anaerobic bacteria is usually underestimated. Despite the fact that *F. magna* is one of the most ubiquitous anaerobic species on the skin (Higaki and Morohashi, 2003; Gao *et al.*, 2007), only limited information is available concerning the molecular mechanisms that promote their colonization and survival. Only recently, a *F. magna* adhesion factor was identified and the bacteria were shown to be localized on the epidermis adhering to basement membranes (Frick *et al.*, 2008). Whether *F. magna* and *S. aureus* in fact interfere with attachment sites or whether other factors are responsible for the observed negative correlation remains to be elucidated.

The anterior nares are an open environment, permanently exposed to microorganisms present in the surroundings through the 10 000 l of air we inhale per day (Wilson, 2005). It cannot be excluded that some of the phylotypes reported here are due to the transient microorganisms. Recent culture-independent studies on indoor environments, specifically house dust (Rintala *et al.*, 2008) have indicated a complex community dominated by Gram-positive taxa, including members of the genera *Corynebacterium, Propionibacterium, Staphylococcus* and *Streptococcus* (Täubel *et al.*, 2009), supporting a permanent source of inoculation for the nose. A total of 15 sequences from five of six volunteers not included in our bacterial diversity analysis supported the presence of chloroplasts. Chloroplasts have recently been shown to be ubiquitous in house dust, accounting for up to 20–25% of sequences (Pakarinen *et al.*, 2008; Täubel *et al.*, 2009). It is thus reasonable to assume that the chloroplast sequences are indicative of those organisms that only pass through the anterior nares. Interestingly, an analysis of biofilms present on shower curtains (Kelley *et al.*, 2004) showed sequences representing *Sphingomonas* spp. and *Methylobacterium* spp. similar to phylotypes observed in low abundance in one and five volunteers, respectively. Consequently, those species may also be only transients and not residents. A further possible source of anterior nare transient microorganisms is through hand-to-nose transfer. In a study on hand surface bacteria (Fierer *et al.* (2008)), *Enterobacteriacea* were more abundant on the dominant hand, a feature argued to be due to differences in skin environmental conditions or to the dominant hand coming into contact with a greater variety of environmental surfaces. The abundance of *Enterobacteriacea* also increased with time after hand washing. The extent to which fecal contact by hands results in the transfer of microorganisms to the anterior nares and the subsequent detection of transient low abundant phylotypes, such as members of the *Enterobacteriacae* family (Figure 7), remains to be elucidated.

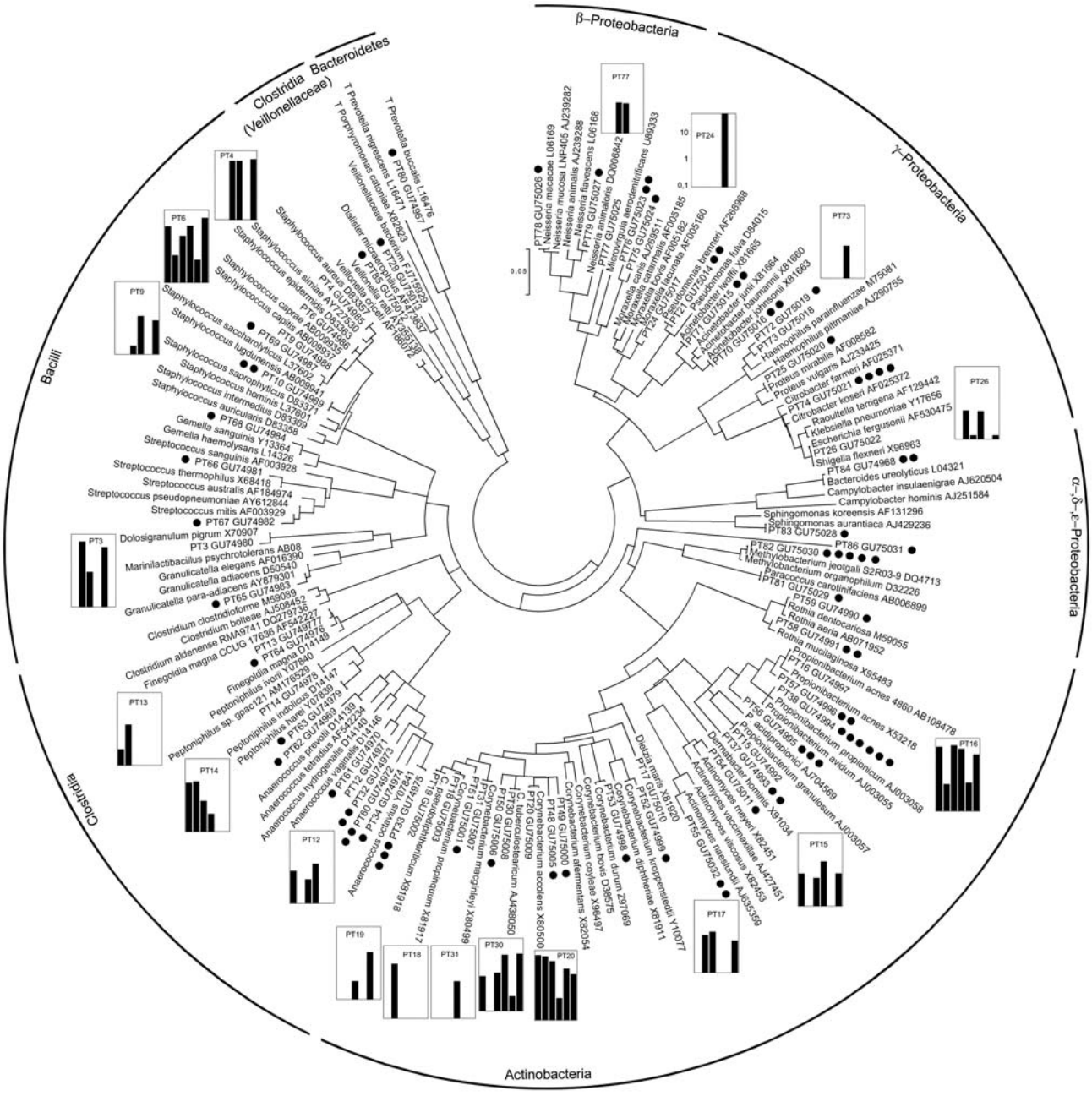

Figure 6 Taxonomical distribution of the 16S rRNA gene sequences retrieved by pyrosequencing analysis of six selected nasal microbial communities. Sequences showing >99% sequence identity were grouped as 'phylotypes' and representative phylotypes are indicated. Phylotypes never observed at an abundance >1% are indicated by filled circles, with the number of circle indicating the number of volunteers, which carried the respective phylotype. If a phylotype was observed at least once at an abundance >1%, its abundance in each of the six volunteers is indicated by bar diagram. Abundance is given at a logarithmic scale from 0.1% to 60% (see phylotype 24). Sequences of closely related bacterial type strains and isolates were included in the analysis. GenBank accession numbers are given after the strain name or the phylotype.

Although overall our ecological indices of species diversity and evenness were similar to other published reports (Grice *et al.*, 2009) and a gross overview of the anterior nare microbial community has recently been given (Costello *et al.*, 2009), our detailed characterization of the nares revealed its fine-scale diversity and novel associations between species. In particular, we identified members belonging to either the core resident or transient communities. Further insights into species–species interactions and their relationships with the host may result in better intervention strategies to control pathogenic strains. Future work should be geared towards unraveling and exploring the

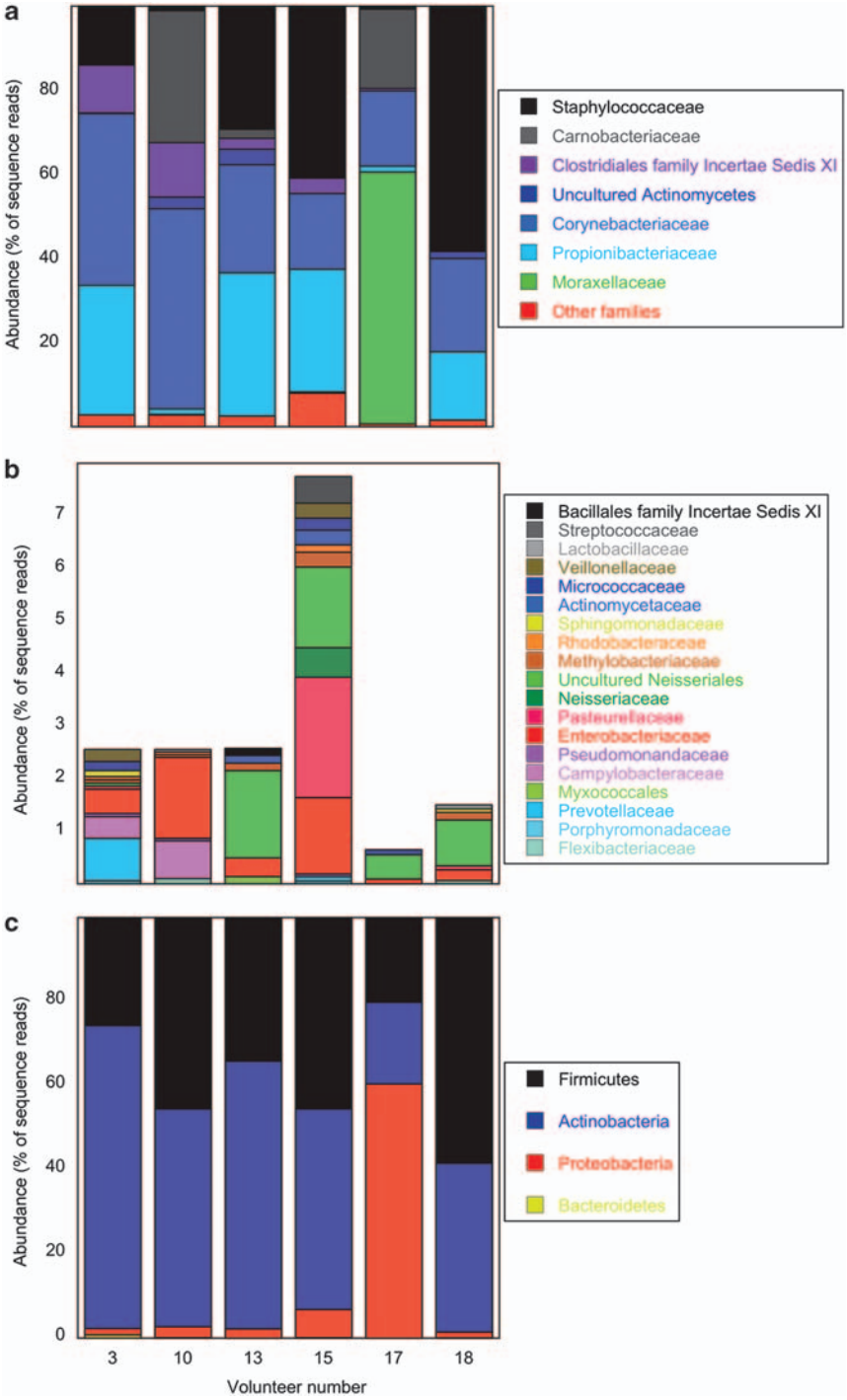

Figure 7 Bacterial community composition of the anterior nares of six volunteers as assessed by 454-pyrosequencing. Phylotypes were grouped into families and phyla, and the relative abundance of the families or phyla is given. (**a**) Predominant families, (**b**) families observed in minor abundance and (**c**) phylum level diversity.

mechanisms involved in such relationships. To this end, we believe the significant converse relationship between *S. aureus* and *F. magna* further highlights the associations among community members, where such knowledge may equip us with strategies to better control *S. aureus* in health-care settings in the future.

Acknowledgements

We thank C Luge, S Edelmann, M Schulte and S Kahl for meticulous technical assistance. We further appreciate Andrew Oxley for critically reading our manuscript and scrupulous data evaluation. This work was supported in part by grants to KB and CvE from the BMBF, Germany

(01KI07100 and PTJ-BIO/0313801B). Nucleotide sequences reported in this study have been deposited into the GenBank database under accession numbers GU074963–GU075395.

References

Becker K, Pagnier I, Schuhen B, Wenzelburger F, Friedrich AW, Kipp F *et al.* (2006). Does nasal cocolonization by methicillin-resistant coagulase-negative staphylococci and methicillin-susceptible *Staphylococcus aureus* strains occur frequently enough to represent a risk of false-positive methicillin-resistant *S. aureus* determinations by molecular methods? *J Clin Microbiol* **44**: 229–231.

Benjamini Y, Yekutieli D. (2001). The control of false discovery rate in multiple testing under dependency. *Annals of Statistics* **29**: 1165–1188.

Bogaert D, van Belkum A, Sluijter M, Luijendijk A, de Groot R, Rümke HC *et al.* (2004). Colonisation by *Streptococcus pneumoniae* and *Staphylococcus aureus* in healthy children. *Lancet* **363**: 1871–1872.

Clarke KR. (1993). Non-parametric multivariate analyses of changes in community structure. *Austral Ecol* **18**: 117–143.

Clarke KR, Warwick RM. (2001). *Change in marine communities: An approach to statistical analysis and interpretation*. PRIMER-E Ltd: Plymouth, UK.

Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R, Science. (2009). Bacterial community variation in human body habitats across space and time. *Science* **326**: 1694–1697.

Dekio I, Hayashi H, Sakamoto M, Kitahara M, Nishikawa T, Suematsu M *et al.* (2005). Detection of potentially novel bacterial components of the human skin microbiota using culture-independent molecular profiling. *J Med Microbiol* **54**: 1231–1238.

Edgar RC. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**: 113.

Fierer N, Hamady M, Lauber CL, Knight R. (2008). The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proc Natl Acad Sci USA* **105**: 17994–17999.

Frick I, Karlsson C, Mörgelin M, Olin AI, Janjusevic R, Hammarström C *et al.* (2008). Identification of a novel protein promoting the colonization and survival of *Finegoldia magna*, a bacterial commensal and opportunistic pathogen. *Mol Microbiol* **70**: 695–708.

Galtier N, Gouy M, Gautier C. (1996). SeaView and Phylo_win, two graphic tools for sequence alignment and molecular phylogeny. *Comput Applic Biosci* **12**: 543–548.

Gao Z, Tseng CH, Pei Z, Blaser MJ. (2007). Molecular analysis of human forearm superficial skin bacterial biota. *Proc Natl Acad Sci USA* **104**: 2927–2932.

García-Rodríguez JA, Fresnadillo Martínez MJ. (2002). Dynamics of nasopharyngeal colonization by potential respiratory pathogens. *J Antimicrob Chemother* **50**: S2:59–S2:73.

Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC *et al.* (2009). Topographical and temporal diversity of the human skin microbiome. *Science* **324**: 1190–1192.

Grice EA, Kong HH, Renaud G, Young AC, Program NCS, Bouffard GG *et al.* (2008). A diversity profile of the human skin microbiota. *Genome Res* **18**: 1043–1050.

Hamady M, Walker JJ, Harris JK, Gold NJ, Knight R. (2008). Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nat Methods* **5**: 235–237.

Higaki S, Morohashi M. (2003). Characteristics of anaerobes from skin specimens. *Drugs Exp Clin Res* **29**: 153–155.

Kelley ST, Theisen U, Angenent LT, St Amand A, Pace NR. (2004). Molecular analysis of shower curtain biofilm microbes. *Appl Environ Microbiol* **70**: 4187–4192.

Lécuyer H, Audibert J, Bobigny A, Eckert C, Jannière-Nartey C, Buu-Hoï A *et al.* (2007). *Dolosigranulum pigrum* causing nosocomial pneumonia and septicemia. *J Clin Microbiol* **45**: 3474–3475.

Lina G, Boutite F, Tristan A, Bes M, Etienne J, Vandenesch F. (2003). Bacterial competition for human nasal cavity colonization: role of Staphylococcal agr alleles. *Appl Environ Microbiol* **69**: 18–23.

Marchesi JR, Sato T, Weightman AJ, Martin TA, Fry JC, Hiom SJ *et al.* (1998). Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding for bacterial 16S rRNA. *Appl Environ Microbiol* **64**: 795–799.

Nawrocki EP, Eddy SR. (2007). Query-dependent banding (QDB) for faster RNA similarity searches. *PLoS Comput Biol* **3**: e56.

Pakarinen J, Hyvärinen A, Salkinoja-Salonen M, Laitinen S, Nevalainen A, Mäkelä MJ *et al.* (2008). Predominance of Gram-positive bacteria in house dust in the low-allergy risk Russian Karelia. *Environ Microbiol* **10**: 3317–3325.

Pettigrew MM, Gent JF, Revai K, Patel JA, Chonmaitree T. (2008). Microbial interactions during upper respiratory tract infections. *Emerg Infect Dis* **14**: 1584–1591.

Regev-Yochay G, Dagan R, Raz M, Carmeli Y, Shainberg B, Derazne E *et al.* (2004). Association between carriage of *Streptococcus pneumoniae* and *Staphylococcus aureus* in children. *JAMA* **292**: 716–720.

Rintala H, Pitkäranta M, Toivola M, Paulin L, Nevalainen A. (2008). Diversity and seasonal dynamics of bacterial community in indoor environment. *BMC Microbiol* **8**: 56.

Schwieger F, Tebbe CC. (1998). A new approach to utilize PCR-single-strand-conformation polymorphism for 16s rRNA gene-based microbial community analysis. *Appl Environ Microbiol* **64**: 4870–4876.

Stone L, Roberts A. (1990). The checkerborad score and species distribution. *Oecologia* **85**: 74–79.

Tamura K, Dudley J, Nei M, Kumar S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**: 1596–1599.

Täubel M, Rintala H, Pitkäranta M, Paulin L, Laitinen S, Pekkanen J *et al.* (2009). The occupant as a source of house dust bacteria. *J Allergy Clin Immunol* **124**: 834–840.

Uehara Y, Nakama H, Agematsu K, Uchida M, Kawakami Y, Abdul Fattah AS *et al.* (2000). Bacterial interference among nasal inhabitants: eradication of *Staphylococcus aureus* from nasal cavities by artificial implantation of *Corynebacterium* sp. *J Hosp Infect* **44**: 127–133.

Ulrich W. (2008). Pairs—a FORTRAN program for studying pair-wise species associations in ecological matrices. www.uni.torun.pl/~ulrichw.

Ulrich W, Zalewski M. (2006). Abundance and co-occurrence patterns of core and satellite species of ground beetles on small lake islands. *OIKOS* **114**: 338–348.

van Belkum A, Verkaik NJ, de Vogel CP, Boelens HA, Verveer J, Nouwen JL *et al.* (2009). Reclassification of *Staphylococcus aureus* nasal carriage types. *J Infect Dis* **199**: 1820–1826.

van Rijen MM, Bonten M, Wenzel RP, Kluytmans JA. (2008). Intranasal mupirocin for reduction of *Staphylococcus aureus* infections in surgical patients with nasal carriage: a systematic review. *J Antimicrob Chemother* **61**: 254–261.

von Eiff C, Becker K, Machka K, Stammer H, Peters G. (2001). Nasal carriage as a source of *Staphylococcus aureus* bacteremia Study Group. *N Engl J Med* **344**: 11–16.

Wildeboer-Veloo ACM, Harmsen HJM, Welling GW, Degener JE. (2007). Development of 16S rRNA-based probes for the identification of Gram-positive anaerobic cocci isolated from human clinical specimens. *Clin Microbiol Infect* **13**: 985–992.

Williams RE. (1963). Healthy carriage of *Staphylococcus aureus*: its prevalence and importance. *Bacteriol Rev* **27**: 56–71.

Wilson M. (2005). *Microbial inhabitants of humans: their ecology and role in health and disease*. Cambridge University Press: New York, USA.

Witzig R, Junca H, Hecht HJ, Pieper DH. (2006). Assessment of toluene/biphenyl dioxygenase gene diversity in benzene-polluted soils: links between benzene biodegradation and genes similar to those encoding isopropylbenzene dioxygenases. *Appl Environ Microbiol* **72**: 3504–3514.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)