Mycoplasma salivarium as a Dominant Coloniser of Fanconi Anaemia Associated Oral Carcinoma

Birgit Henrich1*, Madis Rumming2,3, Alexander Sczyryba3, Eunike Velleuer2, Ralf Dietrich4, Wolfgang Gerlach5, Michael Gombert2, Sebastian Rahn1, Jens Stoye5, Arndt Borkhardt2, Ute Fischer2

1 Institute of Medical Microbiology and Hospital Hygiene, Medical Faculty, Heinrich Heine University, Düsseldorf, Germany, 2 Department of Paediatric Oncology, Hematology and Clinical Immunology, Center for Child and Adolescent Health, Medical Faculty, Heinrich Heine University, Düsseldorf, Germany, 3 Computational Metagenomics, Faculty of Technology, Center for Biotechnology, Bielefeld University, Bielefeld, Germany, 4 German Fanconi-Anemia-Help e.V., Unna, Germany, 5 Genome Informatics, Faculty of Technology, Center for Biotechnology, Bielefeld University, Bielefeld, Germany

Abstract

Mycoplasma salivarium belongs to the class of the smallest self-replicating Tenericutes and is predominantly found in the oral cavity of humans. In general it is considered as a non-pathogenic commensal. However, some reports point to an association with human diseases. M. salivarium was found e.g. as causative agent of a submasseteric abscess, in necrotic dental pulp, in brain abscess and clogged biliary stent. Here we describe the detection of M. salivarium on the surface of a squamous cell carcinoma of the tongue of a patient with Fanconi anaemia (FA). FA is an inherited bone marrow failure syndrome based on defective DNA-repair that increases the risk of carcinomas especially oral squamous cell carcinoma. Employing high coverage, massive parallel Roche/454-next-generation-sequencing of 16S rRNA gene amplicons we analysed the oral microbiome of this FA patient in comparison to that of an FA patient with a benign leukoplakia and five healthy individuals. The microbiota of the FA patient with leukoplakia correlated well with that of the healthy controls. A dominance of Streptococcus, Veillonella and Neisseria species was typically observed. In contrast, the microbiome of the cancer bearing FA patient was dominated by Pseudomonas aeruginosa at the healthy sites, which changed to a predominance of 98% M. salivarium on the tumour surface. Quantification of the mycoplasma load in five healthy, two tumour- and two leukoplakia-FA patients by TaqMan-PCR confirmed the prevalence of M. salivarium at the tumour sites. These new findings suggest that this mycoplasma species with its reduced coding capacity found ideal breeding grounds at the tumour sites. Interestingly, the oral cavity of all FA patients and especially samples at the tumour sites were in addition positive for Candida albicans. It remains to be elucidated in further studies whether M. salivarium can be used as a predictive biomarker for tumour development in these patients.

Introduction

Human microbiomes represent complex, site-specific spectra of bacteria, fungi, and archaea, whose compositions are determined but also dependent on the state of health of the colonised individual. The microbiome of the gut is essential for food metabolism and uptake, whereas the oral microbiome preserves the physical integrity within the oral cavity and is functionally different from the gut environment [1]. Only around 50% of oral microbiomes can be cultivated and studied employing classical biochemical techniques at present. Next generation sequencing (NGS) of variable regions in the gene encoding the 16S rRNA first enabled in-depth, cultivation independent studies of the oral microbiomes [2–6]. Nine variable regions in the 16S rDNA can be used which differ in their potential to discriminate bacterial species [7,8]. For instance, to answer the question how the oral microbiome of the saliva is composed in healthy people, Roche/454-next-generation-sequencing of amplicon libraries comprising the V1-V2 variable region of the 16S rDNA was employed that had been shown to be appropriate to achieve taxonomic assignment for a wide range of bacterial genera investigated [7,8]. Besides unravelling the general composition of the microbial community, Costello and coworkers showed in 2009 that an individual’s oral microbiome is stable over time by comparing samples taken on four different occasions. The group of Zaura compared the microbiomes from intra-oral sites of three systematically and orally healthy individuals employing the V5-V6 region of the 16S rDNA [9]. They hypothesized a core oral microbiome to be present in health with the predominant taxa/phyla belonging to Firmicutes, Proteobacteria, Actinobacteria, Bacteroidetes and Fusobacteria based on their findings of a great proportion of similar amplicon reads found in all subjects. Depending on the site of colonisation, the composition of oral microbiomes differs. Diaz and coworkers analysed bacterial communities in saliva and buccal mucosa and found that intersubject variability was lower than differences between saliva and mucosal communities with high abundance of Streptococcus mitis and Gemella haemolysans predominantly found in the mucosa [10].
Variability of the oral microbiota has been demonstrated to relate to oral diseases, too. In endodontic infections, Li and coworkers found Bacteroidetes as the most prevalent bacterial phylum in infected root canal spaces [11]. The group of Hsiao characterized the site-dependent microbiomes in endodontic infections and published that Prevotella, which belongs to the Bacteroidetes, and Fusobacteria were most abundant in the oral cavity whereas the Firmicutes: Granulicatella adiacens, Eubacterium yunii, and Streptococcus mitis; and the Bacteroidetes: Prevotella melaninogenica and Prevotella salivae, were over-represented in diseased tissues of root and abscesses [12]. Porphyromonas gingivalis, Tannerella forsythia, and Actinobacillus actinomycetemcomitans were characterized as contributing pathogens in periodontitis [13].

Besides infectious diseases, variability of the oral microbiota also related to oral cancers. In the saliva of patients with oral squamous cell carcinoma (OSCC), high levels of facultative oral streptococci related to oral cancers. In the saliva of patients with oral squamous cell carcinoma (OSCC), high levels of facultative oral streptococci were reported to be associated with oral cancers in patients with oral squamous cell carcinoma (OSCC), high levels of facultative oral streptococci were found in the saliva of patients with oral squamous cell carcinoma (OSCC) [14]. and members of eight phyla of bacteria were characterised as contributing pathogens in periodontitis [13].

In a retrospective analysis, four samples of a 41 year old FA-patient were analysed, which corresponded to samples of another FA-patient who harboured a benign oral lesion (leukoplakia, sample set FA_L1) and samples of a control group of five healthy individuals (samples H1-H5). The sample sites of FA_L1 were the same as of FA_T1, the samples from the control group were all taken from site B. After preparation of genomic DNA, a high coverage, massive parallel Roche/454-next-generation-sequencing approach of the V1-V2 variable 16S rDNA region was used to obtain a total of 908,308 raw reads for 30 different samples with unique barcodes. After quality filtering in the de-multiplexing step, 272,152 sequence reads remained. From those 30 samples, thirteen samples (H1-H5, site B, FA_L1, sites A, B, E, F and FA_T1, sites A, B, E, F) belonged to this study. Table 1 gives an overview of the total number of sequence reads per sample. Using the NGS analysis software pipeline QIIME [20] 4,569 Operational Taxonomic Unit (OUT) clusters were built from all 272,152 sequences. The OTUs were subjected to further chimeric OTU filtering and taxonomical classification. 1,010 OTUs that were identified to be chimeric were excluded, which resulted only in a small change in fractions on all taxonomical levels between 0 - 0.5%. 15 taxonomic families were identified. Each accounted for at least 4% of the microbiota in one of the thirteen samples.

Bacteroidetes, Firmicutes, Proteobacteria and Tenericutes represented the most common phyla in the microbiomes of the four samples of FA patient FA_T1 (Figure 2, A), with Tenericutes as the predominant phylum found on the tumour surface (sample B). In contrast, the phyla of Actinobacteria, Bacteroidetes, Firmicutes and Proteobacteria were common in the respective samples of FA patient FA_L1 and all healthy probands (at site B), with a dominance of Firmicutes and lack of Tenericutes.

As depicted in Figure 2, B the mean species diversity at a given sample site (alpha diversity) is very low in the microbiota of FA_T1 compared to the FA_L1 microbiota and the control group. The diversity index (Shannon index) that approaches zero at lowest diversity was calculated for the control group (4.17±0.54), FA_L1 (3.75±0.2) and FA_T1 (2.43±1.42). The alpha diversity indicates that the microbiota is getting less diverse the more the probands are moving into a tumourous state. This observation is also supported by a calculation of the mean species diversity between different habitats (sample sites), the beta diversity (Figure 2, C). Samples derived from FA_T1 form a low diversity cluster on their own, completely separated from H1-H5 and FA_L1. The distribution of the data points of H1-H5 compared to FA_L1 suggests that the beta diversity is supporting our hypothesis of less diversity in an environment that is moving into a tumourous state. Taking these observations of reduced diversity into account, the dominant species, Mycoplasma salivarium can be proposed as a cofactor of the tumourous state of proband FA_T1. The significance and correlation of this finding was confirmed by the mantle test directly on the taxonomic distance matrix and the tumour weighted sample/site-specific distance matrix (with an r-value of 0.91 and a p-value of 0.001, where r stands for correlation and p for significance).

Having a closer look at the respective taxonomic orders, reads classified as Mycoplasmatales comprised 98.2% of the microbiota on the tumour surface of patient FA_T1 and 39.1% of the gingival microbiota nearby the tumour. The variety of taxonomic orders found at the healthy site of tongue and gingiva of FA_T1 was much larger with the highest values for Pseudomonadales and Mycoplasmatales at tongue (43.8% and 18.7%, respectively) and gingiva (30.3% and 26.4%).

The dominance of M. salivarium in the microbiota of healthy probands (Figure 3, A) and the FA patient FA_L1 (Figure 3, B) corresponded to one another. We found a predominance of Streptococcaceae, which belong to the Firmicutes, in
nearly all of these samples and of Neisseriaceae and Pasteurellaceae, which belong to the Proteobacteria. In contrast, the percentage of Streptococcaceae was dramatically reduced in the microbiota of the tumour patient FA_T1 (Figure 3, C). Mycoplasmataceae dominated on the tumour surface (site B) and the adjacent gingiva (site A) and Pseudomonadaceae were frequent at all sites, except the tumour site. Respective sequence reads of the predominant families were then clustered using the SeqMan program of the LaserGene software package. Highest homologous species were identified using the consensus sequences of each cluster (contig) in MegaBLAST analysis against the nucleotide collection database (nr/nt) of BLASTN (http://blast.ncbi.nlm.nih.gov/Blast.cgi [21]). Thus, it became obvious that in healthy probands and FA patient FA_L1 Rothia mucilaginosa was the dominant species of the Actinobacteria phylum, which was nearly absent in the tumour patient FA_T1 (see supplementary file of Table S1). Prevotella melaninogenica and P. nanesciens of the Bacteroidetes phylum were also mainly found in the microbiota of healthy people and the leukoplakia patient whereas P. salicue and Prevotella spp. of different oral taxa dominated the gingival microbiota of the tumour patient. Of the Firmicutes, Veillonella parvula was comparably found in the samples of healthy and diseased probands, whereas Selenomonas spp. and Megaspheura micronuformis were only increased in the gingival samples of tumour patient FA_T1. Viridans-streptococci with members of the mitis-group such as S. mitis, S. oralis and S. infantis and S. salivarius of the salivarius-group comprise the main part of streptococci in healthy individuals and the leukoplakia patient. Higher levels of mutans-streptococci, such as S. mutans and S. criceti, and of S. anginosus were found only in samples of the tumour patient. A specific feature of the microbiota of the leukoplakia patient was found in the family of Flavobacteriaceae with Capnocytophaga species being increased at all sites. Highest levels of C. sputigena and C. canimorsus were found at the tongue next to the site of leukoplakia.

With the exception of a single M. faecium sequence read detected in one of the healthy individuals, Mycoplasma salivarium was the unique species of the Mycoplasmataceae, and Pseudomonas aeruginosa the main species of the Pseudomonadaceae detected at highest levels in the microbiota of the tumour patient. Thus, it is likely that within the proteobacterial phylum, the Neisseriaceae, with N. flavescens increased in healthy and leukoplakia individuals, have been displaced by P. aeruginosa in the tumour patient.

Table 1. Total number of reads generated per sample.

| Sample set | Sample Site | Total Number of Reads | Barcode |
|------------|-------------|-----------------------|---------|
| H1         | B           | 10,595                | TACTGAGCTA |
| H2         | B           | 8,230                 | CGAGAGATAC |
| H3         | B           | 7,742                 | TCACGTACTA |
| H4         | B           | 18,730                | AGCATCCTAG |
| H5         | B           | 10,590                | TCTACGTAGC |
| FA_L1      | A           | 5,342                 | ACATACGGCT |
| FA_L1      | B           | 7,719                 | TCGTCGCTAG |
| FA_L1      | E           | 12,042                | TAGAGACGAG |
| FA_L1      | F           | 10,923                | TACTTCCTTG |
| FA_T1      | A           | 8,677                 | AGACTATACT |
| FA_T1      | B           | 7,229                 | ACTGTACAGT |
| FA_T1      | E           | 5,741                 | ACTACTATGT |
| FA_T1      | F           | 3,929                 | ACACGAGCTA |

doi:10.1371/journal.pone.0092297.t001
TaqMan PCRs were used to quantify the bacterial load of *Pseudomonas aeruginosa* and *Mycoplasma salivarium* targeting the single copy genes *gyrB* and *rpoB*, respectively. As depicted in Figure 4, A in red, the results of the 16S rDNA survey were confirmed by *M. salivarium* qPCR, with the highest load of *M. salivarium* on the tumour surface of patient FA_T1 and lower amounts at the corresponding healthy sites of tongue and gingiva. For the TaqMan PCR analyses we included samples of a second FA tumour patient (FA_T2, sample set A,B,C,D,E and F with site D corresponding to the tumour surface), and a second leukoplakia

Figure 2. Microbiota detected by NGS on the surface of distinct oral sites of the analysed cohort. DNA of samples, which derived from site B of five healthy probands (H1-H5) and sites A, B, E and F of two FA patients (FA_L1 with a leukoplakia at site F and FA_T1 with a tumour at site B), was subjected to V1-V2 16S rDNA based NGS analysis. A., phyla and families detected in the 13 samples. Families are indicated in shades of the same colour as the respective phylum; B., alpha diversity of all samples (H1-H5 in blue, FA_L1 in green and FA_T1 in red); C., beta diversity plot of all samples (H1-H5 as blue circles, FA_L1 as green squares and FA_T1 as red triangles). doi:10.1371/journal.pone.0092297.g002
For these two patients no NGS data were available. Interestingly, colonisation of both leukoplakia patients, FA_L1 and FA_L2, with *M. salivarium* was rare or of low concentration as in healthy probands (see Figure 4, B), whereas all samples of both tumour patients were *M. salivarium* positive with the highest load at the tumour surfaces. The *M. salivarium* load at the local sites of cellular alterations (tumour or leukoplakia) was compared with the respective contralateral sites of each FA-patient (Figure 4, C). The greatest differences in *M. salivarium* load were found in the tumour FA patients. Quantification of *P. aeruginosa*, which is shown in Figure 4, A, hatched, revealed that FA patient FA_T1 was the only one with a high load of this pathogen. All other FA patients and healthy probands were tested *P. aeruginosa*-negative.

As it had been known that the patient FA_T1 had have contracted oral mucositis for a long time and that oral candidiasis often impedes tumour detection at an early stage, DNA of all 23 samples were subjected to a *Candida*-specific qPCR, followed by melting curve analysis for product detection. As shown in Figure 4, A in grey, *Candida*-PCR was positive in most samples of FA patients FA_T1 and FA_T2, including the respective tumour regions B and D, and all samples of the FA patient FA_L1 with highest *Candida* load at the site F of leukoplakia. In contrast, samples of the healthy probands and samples A, B and E of FA patient L2 with leukoplakia at site B were *Candida*-negative. Products of *Candida*-PCR and of a confirmatory ITS1-ITS4-fungal PCR were sequenced by priming with ITS4 using the method of Sanger [22] and homologous species identified in BLAST analysis. All samples were positive for *Candida albicans* (90-99% homology).

**Discussion**

Fanconi anaemia is associated with an increased risk to develop OSCC which is difficult to diagnose at an early stage of development by the use of minimal invasive procedures. Currently, short-period screening of the patients is one of the preventative measures. With the findings of Mager et al., published in 2005 that the salivary microbiomes of OSCC patients differ from those of healthy people [17], a concept was formed that development of OSCC is accompanied by change of the microbiome-composition especially at the respective cell surface. To test this hypothesis, oral microbiomes were characterised by next generation sequencing and real time PCR. Four samples were of primary interest because they were collected from the surface of an oral squamous cell carcinoma of the root of the tongue, the adjacent gingiva and the contralateral (healthy) sites. These specimens were derived from a male patient born in 1970 who was diagnosed with Fanconi anaemia at the age of nine years. The advanced squamous cell carcinoma at the base of the tongue (cT3, cN2b, M0, G3) was diagnosed in July 2011. The centre line encroaching tumour (Figure 1, B) was inoperable with recidivating bleedings in pronounced thrombocytopenia. After tracheostomy, pneumonia-derived sepsis in August 2011 was treated with tacobactam/piperacillin. Palliative radiotherapy of the tumour was aborted in October 2011 due to severe complications of haemorrhage and...
mycositis. After suffering two pneumonia-derived septicemias in November with the detection of *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* in tracheostomal smear and tracheal secretion, G-CSF therapy started. Antibiogram revealed selection of a highly resistant strain of *P. aeruginosa* in the second sepsis, probably due to piperacillin/tazobactam therapy. The patient died at the end of December 2011. As stated by a family member the patient had suffered from mucositis for longer times without being medicated other than using an oral Moronal (Nystatin) suspension.

The composition of the oral microbiome in each of the four samples of this FA patient FA_T1 differed from that of the proposed oral core microbiome of healthy people, which was defined by Zaura et al. to be mainly composed of Firmicutes, Proteobacteria, Actinobacteria, Bacteriodetes and Fusobacteria [9]. Interestingly, the microbiota of the FA patient with leukoplakia, FA_L1, corresponded to that of a healthy person, as those of the healthy probands included in the present investigation, too. In contrast, the oral microbiome of the FA patient with tumour (FA_T1) was dominated by *Tenericutes* and *Proteobacteria*, contained *Bacteroidetes* and *Firmicutes* but lacked higher amounts of *Actino- and Fusobacteria*. The dominance of the opportunistic pathogenic proteobacterium *Pseudomonas aeruginosa* in the oral cavity of the patient was not surprising, as the respiratory tract of patients with tracheostomy often becomes infected with *P. aeruginosa* [23,24] and two episodes of *P. aeruginosa*-derived pneumonia and septicemia were known. A study of He and coworkers demonstrated that the salivary microbiota of healthy humans is able to prevent the integration of pathogenic bacteria such as *P. aeruginosa* [25] suggesting that the oral microbiota of the immunocompromised FA patient must have already been misbalanced before to enable *P. aeruginosa* colonisation. Unfortunately, the antibiotic regime has led to the selection of a highly resistant *P. aeruginosa* strain difficult to eradicate.

The detection of high loads of *M. salivarium* in each sample of the tumour patients FA_T1 and FA_T2 with prevalence on the surface of the oral cancers was an unexpected finding, especially as *M. salivarium* was generally considered as a non-pathogenic inhabitant of the oral cavity [26]. Only in a few cases it has been considered to participate in oral and periodontal infections [27–30] or to be the causative agent of a submasseteric abscess [31]. Nevertheless, as *M. salivarium* is difficult to culture, it has thus been rarely looked for. Incidental findings identified *M. salivarium* as causative agent of disseminated infections, such as a chronic joint infection in a patient with hypogammaglobulinaemia [32]. *M.

![Figure 4. Real time PCRs for the detection of *M. salivarium*, *P. aeruginosa*, and *Candida*. DNA samples, which were derived from healthy probands (H1-H5, site B), FA-patients with leukoplakia (FA_L1 with a leukoplakia at site F and FA_L2 with leukoplakia at site B) and FA patients with tumour (FA_T1 with a tumour at site B and FA_T2 with tumour at site D) were subjected to real time PCR in duplicates. A. TaqMan qPCR-derived copy numbers of *M. salivarium* (red), *P. aeruginosa* (hatched) and *Candida* (grey). Due to targeting a single copy gene, copy numbers correspond to genome equivalents for *M. salivarium* and *P. aeruginosa*. B. scatter plot of *M. salivarium* load in the different proband groups: Healthy (H1-H5, blue circles), samples of FA-patients with leukoplakia (FA_L1 and FA_L2, green squares) and with tumour (FA_T1 and FA_T2, red triangles). C. scatter plot of *M. salivarium* load of FA patients FA_T1 (dark red) and FA_T2 (red) at the tumour surface (Tumour) in relation to the contralateral non-tumour surface (Non-T.) and of the FA patients FA_L1 (green) and FA_L2 (dark green) at the site of leukoplakia (Leukopl.) in relation to the contralateral non-leukoplakia site (Non-L.).

doi:10.1371/journal.pone.0092297.g004
was found in brain abscesses of two patients [33], in an occluded biliary stent in polymicrobial community with Candida glabrata [34] and as causative agent of a pleural empyema [35].

Characterisation of complex microorganisms by using molecular biological approaches such as NGS nowadays offers the opportunity to detect species that are unsuitable or difficult to culture such as M. salivarium. Having a closer look at the yet published 16S rDNA surveys only a few descriptions can be found about mycoplasmas taking part in oral microbiomes. Dewhurst and coworkers published in 2010 a comprehensive human oral microbiome database (HOMID) based on yet published 16S rDNA sequences. In their clone libraries they only found relative few mycoplasmas with representatives of rDNA sequences. In their clone libraries they only found relative few mycoplasmas with representatives of Mycoplasma hominis, Mycoplasma salivarium, Mycoplasma fauclum [36] and representatives of Tenericutes [G-1] sp. oral taxon 504, which were very deeply branching within the Tenericutes. In using the Human Oral Microbe Identification Microarray (HOMIM) M. salivarium was found in subgingival plaque samples of six of 47 subjects with refractory periodontitis [30]. Mycoplasma reeds were found in the saliva of OSCC patients representing less than 0.5% of the microbiomes [15]. These findings do not suggest that M. salivarium is a predictive bio-marker of tumour development. However, one should keep in mind that analysis of the salivary microbiome is not specific for the tumour site and that M. salivarium is not a general coloniser of the oral cavity. Interestingly, all of the analysed FA patients were carriers of M. salivarium, but only two of five healthy individuals. Hooper and coworkers analysed OSCC samples by culturing and 16S-based PCR after elimination of the surface attached microbes. Within the tumour tissue they found that the majority of species were saccharolytic and aciduric, reflecting perhaps the selective nature of the acidic and hypoxic microenvironment found within tumours [37,38]. However, in their study the presence of M. salivarium within the tissue was undetectable due to culture media used, which were not suitable for mycoplasma growth, and due to a serious mismatch of the Reverse primer E94 at its 3’-end (5’-GAAGGAGGTGWTCCARCCgCA-3’) hampering amplification of the 16S rDNA-region 5’-AAAGGAGGT-GATCCATCCGCA-3’ of M. salivarium [Acc.-No: AY786574; nt 141-121].

Depending on the amplification conditions used in NGS, the ratio of each phyllum differs. Lazarevic et al. demonstrated in 2012 that the amount of Tenericutes (Mycoplasma) virtually increases in the salivary microbiome when stepping the cycle number in 16S V1-V3-PCR from 20 up to 30 [39]. Thus, quantification of bacteria using NGS analyses seems to be hindered in more than one aspect. On the one hand the amplification conditions may shift the real ratio of microorganisms in a defined environment to something unrelated in the 16S rDNA amplicon based microbiota, and on the other hand, the amount of 16S rDNA derived amplicons is biased by the number of ribosomal operon structures of a bacterium. Whereas the genome of Mycoplasma orale harbours only one ribosomal operon structure, the genome of Lactobacillus delbrueckii harbours nine 16S rDNA copies [40,41], which leads to an overestimation of L. delbrueckii. Thus, it seems to be difficult to quantify a species by a metagenome analysis. In the study presented here, the amounts of M. salivarium and P. aeruginosa were rechecked by TaqMan PCRs, each of those targeting a single copy gene. This approach confirmed a predominance of M. salivarium at the tumour sites. P. aeruginosa, exclusively found in the oral cavity of FA patient FA_T1, was not restricted to the tumour site, but detected at all analysed sites in the mouth with similar abundance. The findings of this study that the oral microbiomes of the tumour FA patients were colonised by high levels of M. salivarium and C. albicans but, as shown in NGS analysis of FA patient
were stored at the recommendations of the manufacturer. DNA preparations standard protocol for DNA preparation from tissues according to 30 min and for 15 min at 95 °C (QIAamp Blood Kit, QIAGEN) were added and incubated at 56 °C. All intact cells (bacterial, fungal and human) were collected by centrifugation (10 min at 5,000 g), resuspended in enzyme solution (20 mg/ml lysozyme; 20 mM Tris/HCl, pH 8.0; 2 mM EDTA; 1.2% Triton) and incubated for at least 30 min at 37°C. 20 µl Proteinase K (>600 µU/ml) and 200 µl Buffer AL (QIAamp Blood Kit, QIAGEN) were added and incubated at 56°C for 30 min and for 15 min at 95°C. DNA isolation then followed the standard protocol for DNA preparation from tissues according to the recommendations of the manufacturer. DNA preparations were stored at −20°C until use.

Real time PCR

Real time PCR assays for the detection of Candida DNA were carried out according to the method of Schabereiter-Gurtner et al. [47] in using Can-F (5’-CTT GGA TCT GGC AGG TCG ACC CTG GAG TAT-3’) and ITS-4 (5’-TCC TGC TTA TGG GTG-BHQ1-3’) [48] for Candida amplification. Real-time PCR was carried out in a total volume of 30 µl consisting of 1x MesaGreen qPCR MasterMix Plus for SYBR Assay (containing buffer, dNTPs (including dUTP), Meteor Taq DNA Polymerase, 4 mM MgCl₂, SYBR Green I, stabilizers and passive reference [RT-SY2X-06sWOU]; Eurogentec, Seraing, Belgium), 300 nM each forward and reverse primer and 2 µl of template DNA. Concurrent amplification of 10³ and 10² copies of pgEM-T cloned ITS1-ITS4 amplons enabled quantification of Candida load. Each positive Candida detection was verified in ITS1-ITS4-primed PCR [48] followed by sequencing [22] and BLAST analysis [21]. Cycling conditions of both PCRs were as follows: 2 min at 95°C followed by 40 cycles of 30 s at 95°C, 30 s at 55°C and 1 min at 72°C. Subsequent melting point analysis followed after 15 s at 95°C and 1 min at 60°C from 65°C to 95°C with an increment of 0.5°C for 15 s and plate read.

TaqMan PCR

In house TaqMan-PCRs were carried out in a total volume of 25 µl consisting of 1x Eurogentec qPCR MasterMix without ROX (containing buffer, dNTPs (including dUTP), HotGoldStar DNA polymerase, 5 mM MgCl₂, Uracil-N-glycosylase and stabilizers [RT-QP2X-03NR, Eurogentec], 300 nM each forward and reverse primer, 200 nM labelled probe and 2.5 µl of template DNA. The pyroB-based M. salivarium TaqMan PCR was carried out with the primers Msal-F (5’-CGG TCA AAT GAT TTC GAT TGC-3’) and Msal-R (5’-GAA CTG CTG GCT TGC TGC-3’) and probe Msal-T (5’-HEX-ATG ATG CTA ACC GTG CGC TTA TGG GTG-BHQ1-3’) [34]. The gyrB-based P. aeruginosa TaqMan-PCR was carried out with the primer pair Paer-F (5’-CTC GAC CAT CGG TCG CGG CAA CAX C-3’) and Paer-R (5’-CGC AGC AGG ATG CCG AGG CC-3’) and the probe Paer-T (5’-FAM-CCG TGG TGG TAG ACC TGT TCC CAG ACC BHQ1-3’) generating a 222 bp amplicon. Amblicon carrying plasmids (pGemT-Msal and pGemT-Paeru) were used in three concentrations (10⁷, 10⁸ and 10⁹ copies/µl) as quantification standards. Each sample was analysed in duplicates. Cycling, fluorescent data collection and analysis were carried out with an iCycler from BioRad (BioRad Laboratories, Munich, Germany) according to the manufacturer’s instructions.

Next-Generation-Sequencing (NGS)

Primers used in this study (forward primer: 5’-GCG TTC GGA GCC GCC TCG TCA AGA AGT TGT GTC TGG AGC AG-3’ and reverse primer: 5’-GCG TTC TCC CTC CGG CCA TCA GNN NNN NNN NNC NCA TGC TGC CTC CGG TAG GAG C-3’) targeted the conserved region of the 16S rDNA flanking the V1-V2 hypervariable regions [49,50]. The forward primer contained part of the Roche/454 primer B sequence (GGC TTC CGA GCC GGC), a key sequence (TCAG), the bacterial primer 27F and a two base pair (‘‘TC’’) linker sequence. The reverse primer contained part of the Roche/454 primer A sequence (GCC TCC CTG CGG CCA), a key sequence (TCAG), a barcode sequence (indicated by “NNNNNNNNNNNN”), the bacterial primer 338R and a “CA”-linker.

25 µl PCR reactions were carried out in triplicates per sample with 0.4 µM each primer, 1 µl template DNA and 1x Platinum PCR SuperMix (22 U/ml complexed recombinant Taq DNA polymerase with Platinum Taq Antibody, 22 mM Tris-HCl (pH 8.4), 55 mM KCl, 1.65 mM MgCl₂, 220 µM dGTP, 220 µM dATP, 220 µM dTTP, 220 µM dCTP, and stabilizers; Invitrogen/Life Technologies, Darmstadt, Germany). Cycling conditions consisted of an initial denaturation step at 94°C for 3 min followed by 35 cycles of 94°C for 35 s, 50°C for 30 s and 72°C for 90 s, terminated by a final extension step at 72°C for 10 min.

Triplicates were pooled and purified using the Gel Extraction Kit (Qiagen). Quality and quantity was assessed using the Quant-iT PicoGreen dsDNA kit (Invitrogen). Purified amplicons were combined in equimolar ratios into a single tube. Pyrosequencing was carried out using primer A and B on a Roche/454 Life Sciences Genome Sequencer FLX instrument according to the recommendations of the manufacturer (Roche, Mannheim, Germany).

Raw sequence data were deposited at the European Nucleotide Archive (ENA) under the study accession PRJEB5069 (http://www.ebi.ac.uk/ena/data/view/PRJEB5069).

Analysis of NGS data using QIIME

The analysis of the Roche/454 NGS data was performed with the QIIME NGS analysis pipeline, which is a general purpose collection of tools and scripts covering the needs of all necessary steps from raw data processing over data normalization, clustering, taxonomical classification to statistics and its visualization [20].

Demultiplexing and quality control. QIIME contains a ready-made workflow for 454 sequenced 16S rDNA beginning with cleaning and mapping the input data with respect to the barcode, linker primer and the reverse primer sequence (Figure S1). The demultiplexing step removed the primer sequences and
assigned reads to samples identified by the correspondent barcode. Quality filtering was also performed with the following default parameters: quality phred score of 25, min/max read length of 200/1000, no ambiguous bases and mismatches in the primer sequences.

**OTU creation and taxonomical classification.** Clustering was performed with uclust [51], a seed based clustering approach, which was run with a minimal sequence identity of 97%. The output was a set of Operational Taxonomic Units (OTUs) with each OTU representing a single species [as specified with the 97% internal sequence identity].

Subsequently, representatives from every OTU cluster were aligned with PyNAST [52] for chimera depletion with ChimeraSlayer of the Microbiome Utilities Portal of the Broad Institute, (http://microbiomeutil.sourceforge.net/ [53]). After removing the chimeric OTU clusters, the remaining representatives were used for taxonomic classification with the RDP classifier [54] and the Greengenes OTU database (http://greengenes.secondgenome.com/ ; Version: 12_10, 97% identity level [species level] [55]) with the default confidence factor of 80%. Reads belonging to a taxonomical family with a minimal total proportion of 4% in one of the four samples were clustered in using the **SeqMan** program of the **LaserGene** software package (Version 6; 97% identity level; DNA Star, Madison, WI, USA). Highest homologous species were identified in using the resulting consensus sequences in **MegaBLAST** analysis against the nucleotide collection database (nr/nt) of BLASTn (http://blast.ncbi.nlm.nih.gov/Blast.cgi [21]) with an exclusion of uncultured /environmental sample sequences. The species of the **Streptococcaceae** were identified using **vmatch** [56] by comparing the input sequences with the **Silva 115 release 16S** reference set [57].

Genome equivalents were calculated from the number of **V1-V2** sequence reads of *M. salivarium* and *P. aeruginosa* based on the copies of 16S rDNA per genome. The genome of *P. aeruginosa* is known to carry four 16S rDNA copies [40] and the *M. salivarium* genome carries one 16S rDNA gene (https://img.jgi.doe.gov/cgi-bin/m/main.cgi?taxon_oid = 2534682365).

**Computation of alpha and beta diversity**

The calculation of the alpha diversity was performed by using the QIIME alpha rarefaction workflow. The five H1-H5 samples from sample site B were pooled as control group, the four samples (sites A, B, E, F) from FA_L1 as FA_L1 group and the four samples (sites A, B, E, F) from FA_L1 as FA_T1 group. The pooled samples were then compared with the Shannon index [58]. Phyligenetic beta diversity analysis was accomplished via the weighted UniFrac metric [39] and the OTU based taxonomical representative tree. From every sample 1000 sequences were taken into computation of the beta diversity. The PCoA plot was generated from the principal coordinates 1 and 2. The correlation was tested using the mantle test comparing the weighted UniFrac matrix with a proband/sample site-specific tumour weight matrix. H1-H5 were weighted as 0, FA_L1 A, B, E as 10 and FA_L1 F as 15, FA_T1 A, E, F as 75 and FA_T1 B as 100. Principal coordinates are a generalisation of the underlying taxonomical information of sample sets with a minimal loss of information. From the input data principal coordinates are generated, which represent a transformation of the high dimensional taxonomical information into the lower dimensional (generalised) feature space of principal coordinates.

**Supporting Information**

**Figure S1** Common amplicon structure. The target sequence (V1-V2 of the 16S rDNA) is amplified using sequence specific forward (27F) and reverse (338R) primers. Amplicons belonging to a specific sample are identified by an integrated unique barcode sequence. The flanking adapter sequences A and B are sequencer –specific primer sequences, Linker sequences are introduced to provide greater flexibility. The resulting common amplicon structure is depicted. (TIF)

**Table S1** Species of the dominant families detected by NGS and their ratio in the microbiotas of FA patients and healthy individuals. (XLS)

**Acknowledgments**

We thank Dana Belick from the Institute of Medical Microbiology and Hospital Hygiene and Olga Felda from the Department of Pediatric Oncology, Hematology and Clinical Immunology, Center for Child and Adolescent Health of the Heinrich-Heine University of Düsseldorf for excellent technical assistance.

**Author Contributions**

Conceived and designed the experiments: BH UF EV. Performed the experiments: BH SR UF MG. Analyzed the data: AS BH JS MR UF WG. Contributed reagents/materials/analysis tools: AB EV RD. Wrote the paper: BH UF.

**References**

1. Belda-Ferre P, Alcaraz LD, Cabrer-Rubio R, Romero H, Simón-Soro A, et al. (2012) The oral metagenome in health and disease. ISME J 6: 46–56.

2. Ahn J, Yang L, Pastor BJ, Gandy I, Morris L, et al. (2011) Oral microbiome profiles: 16S rDNA pyrosequencing and microarray comparison. PLoS One 6: e22788.

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

4. Nasidze I, Quinque D, Li J, Li M, Tang K, et al. (2009) Comparative analysis of human saliva microbiome diversity by barcoded pyrosequencing and cloning approaches. Anal Biochem 391: 64–68.

5. Alcaraz LD, Belda-Ferre P, Cabrer-Rubio R, Romero H, Simón-Soro A, et al. (2012) Identifying a healthy oral microbiota through metagenomics. Clin Microbiol Infect 18 Suppl 4: 54–57.

6. Do T, Devine D, Marsh PD (2013) Oral biofilms: molecular analysis, challenges, and future prospects in dental diagnostics. Clin Cosmet Investig Dent 5: 11–19.

7. Chakravorty S, Helb D, Burday M, Connell N, Alland D (2007) A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. J Microbiol Methods 69: 330–339.

8. Youssef N, Sheik CS, Krumholz LR, Najar FZ, Roe BA, et al. (2009) Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rDNA gene-based environmental surveys. Appl Environ Microbiol 75: 5227–5236.

9. Zaura E, Keijser BJ, Huse SM, Crielaard W (2009) Defining the healthy "core microbiome" of oral microbial communities. BMC Microbiol 9: 259.

10. Diaz PI, Dupuy AK, Abudelo L, Reese B, Oberfeldt C, et al. (2012) Using high throughput sequencing to explore the biodiversity in oral bacterial communities. Mol Oral Microbiol 27: 182–201.

11. Li L, Hsiao WW, Nandakumar R, Barbato SM, Mongodin EF, et al. (2010) Analyzing endodontic infections by deep coverage pyrosequencing. J Dent Res 89: 980–984.

12. Hsiao WW, Li KL, Liu Z, Jones C, Fraser-Liggett CM, et al. (2012) Microbial transformation from normal oral microbiota to acute endodontic infections. BMC Genomics 13: 345.

13. Socransky SS, Haffajee AD (2005) Periodontal microbial ecology. Periodontol 2000 38: 135–187.

14. Sluga K, Tateda M, Saijo S, Horii T, Sato I, et al. (2001) Presence of *Streptococcus* infection in extra- oropharyngeal head and neck squamous cell carcinoma and its analyzing in carcinogenesis. Onkol Rep 8: 215–219.

15. Pusaskar S, Mane SP, JI X, Li Y, Evans C, et al. (2011) Microbial diversity in saliva of oral squamous cell carcinoma. FEBS Immunol Med Microbiol 61: 260–277.
33. Ørsted I, Gertsen JB, Schønheyder HC, Jensen JS, Nielsen H (2011) Mycoplasma

31. Grisold AJ, Hoenigl M, Leitner E, Jakse K, Feierl G., et al. (2008) Submasseteric

30. Colombo AP, Boches SK, Cotton SL, Goodson JM, Kent R, et al. (2009)

28. Watanabe T, Matsuura M, Seto K (1986) Enumeration, isolation, and species

Mycoplasma salivarium

26. Kundsin RB, Praznik J (1967) Pharyngeal carriage of mycoplasma species in

24. Vandecandelaere I, Matthijs N, Van Nieuwerburgh F, Deforce D, Vosters P, et

23. Lowbury EJ, Thom BT, Lilly HA, Babbi JR, Whittal K (1970) Sources of

infection with Pseudomonas aeruginosa in patients with tracheostomy. J Med Microbiol 3: 39–36.

20. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, et al. (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7: 335–336.

19. Scheckenbach K, Wagenmann M, Freund M, Schipper J, Hanenberg H (2012) Immunol Med Microbiol 22: 241–246.

17. Mager DL, Haffajee AD, Devlin PM, Norris CM, Norris CM, Posner MR, et al. (2005) The salivary microbiota as a diagnostic indicator of oral cancer: a descriptive, non-randomized study of cancer-free and oral squamous cell carcinoma patients. J Transl Med 3: 27.

16. Pushkalkar S, J X, Li Y, Estol C, Yegnanarayana R, et al. (2012) Comparison of oral microbiota in tumor and non-tumor tissues of patients with oral squamous cell carcinoma. BMC Microbiol 12: 144.

15. Kundsin RB, Praznik J (1967) Pharyngeal carriage of mycoplasma species in

20. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, et al. (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7: 335–336.

19. Scheckenbach K, Wagenmann M, Freund M, Schipper J, Hanenberg H (2012) Immunol Med Microbiol 22: 241–246.

17. Mager DL, Haffajee AD, Devlin PM, Norris CM, Norris CM, Posner MR, et al. (2005) The salivary microbiota as a diagnostic indicator of oral cancer: a descriptive, non-randomized study of cancer-free and oral squamous cell carcinoma patients. J Transl Med 3: 27.

16. Pushkalkar S, J X, Li Y, Estol C, Yegnanarayana R, et al. (2012) Comparison of oral microbiota in tumor and non-tumor tissues of patients with oral squamous cell carcinoma. BMC Microbiol 12: 144.

15. Kundsin RB, Praznik J (1967) Pharyngeal carriage of mycoplasma species in

20. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, et al. (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7: 335–336.

19. Scheckenbach K, Wagenmann M, Freund M, Schipper J, Hanenberg H (2012) Immunol Med Microbiol 22: 241–246.

17. Mager DL, Haffajee AD, Devlin PM, Norris CM, Norris CM, Posner MR, et al. (2005) The salivary microbiota as a diagnostic indicator of oral cancer: a descriptive, non-randomized study of cancer-free and oral squamous cell carcinoma patients. J Transl Med 3: 27.

16. Pushkalkar S, J X, Li Y, Estol C, Yegnanarayana R, et al. (2012) Comparison of oral microbiota in tumor and non-tumor tissues of patients with oral squamous cell carcinoma. BMC Microbiol 12: 144.

15. Kundsin RB, Praznik J (1967) Pharyngeal carriage of mycoplasma species in

20. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, et al. (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7: 335–336.

19. Scheckenbach K, Wagenmann M, Freund M, Schipper J, Hanenberg H (2012) Immunol Med Microbiol 22: 241–246.

17. Mager DL, Haffajee AD, Devlin PM, Norris CM, Norris CM, Posner MR, et al. (2005) The salivary microbiota as a diagnostic indicator of oral cancer: a descriptive, non-randomized study of cancer-free and oral squamous cell carcinoma patients. J Transl Med 3: 27.

16. Pushkalkar S, J X, Li Y, Estol C, Yegnanarayana R, et al. (2012) Comparison of oral microbiota in tumor and non-tumor tissues of patients with oral squamous cell carcinoma. BMC Microbiol 12: 144.