UC Irvine
UC Irvine Previously Published Works

Title

Bacteria in the airways of patients with cystic fibrosis are genetically capable of producing VOCs in breath.

Permalink

Journal

ISSN

Authors

Bos, Lieuwe DJ
Meinardi, Simone
Blake, Donald
et al.

Publication Date

DOI

Copyright Information

Peer reviewed

Journal of Breath Research

PAPER

Bacteria in the airways of patients with cystic fibrosis are genetically capable of producing VOCs in breath

Lieuwe D J Bos[1], Simone Meinardi[2], Donald Blake[2] and Katrine Whiteson[3]

1 Department of Respiratory Medicine and Intensive Care, Universiteit van Amsterdam—Academic Medical Center, Amsterdam, the Netherlands
2 Department of Chemistry, University of California, Irvine, CA, USA
3 Department of Molecular Biology and Biochemistry, University of California, Irvine, USA

E-mail: katrine@uci.edu

Abstract

Breath contains hundreds of volatile organic compounds (VOCs), the composition of which is altered in a wide variety of diseases. Bacteria are implicated in the formation of VOCs, but the biochemical mechanisms that lead to the formation of breath VOCs remain largely hypothetical. We hypothesized that bacterial DNA fragments in sputum of CF patients could be sequenced to identify whether the bacteria present were capable of producing VOCs found in the breath of these patients.

Breath from seven patients with cystic fibrosis was sampled and analyzed by gas-chromatography and mass-spectrometry. Sputum samples were also collected and microbial DNA was isolated. Metagenomic sequencing was performed and the DNA fragments were compared to a reference database with genes that are linked to the metabolism of acetaldehyde, ethanol and methanol in the KEGG database.

Bacteria in the genera *Escherichia, Lactococcus, Pseudomonas, Rothia* and *Streptococcus* were found to have the genetic potential to produce acetaldehyde and ethanol. Only DNA sequences from *Lactococcus* were implicated in the formation of acetaldehyde from acetate through aldehyde dehydrogenase family 9 member A1 (K00149). *Escherichia* was found to be genetically capable of producing ethanol in all patients, whilst there was considerable heterogeneity between patients for the other genera. The ethanol concentration in breath positively correlated with the amount of *Escherichia* found in sputum (Spearman rho = 0.85, $P = 0.015$). *Rothia* showed the most versatile genetic potential for producing methanol.

To conclude, bacterial DNA fragments in sputum of CF patients can be linked to enzymes implicated in the production of ethanol, acetaldehyde and methanol, which are VOCs that are predictive of respiratory tract colonization and/or infection. This supports that the lung microbiome can produce VOCs directly.

Introduction

Breath contains hundreds of volatile organic compounds (VOCs), the composition of which is altered in a wide variety of diseases [1]. The biochemical mechanisms that lead to the formation of breath VOCs remain largely hypothetical [2–6]. Although a mechanistic understanding of VOC origin is not required to use VOCs in the exhaled breath as a diagnostic tool [1], it would enable identification of the most appropriate patient populations to target with the test and would facilitate the acceptance of breath tests by clinicians.

Bacteria have been implicated in the production of VOCs for a long period of time. For centuries, lung abscesses have been reported to have a strong odor; breath odors were described as a diagnostic test in the ancient writings of Hippocrates. VOC profiles can distinguish between particular pathogens *in vitro* and in animal models of pneumonia [7–9]. This suggests that bacteria create the volatile compounds responsible for the odor associated with infection [10], and that particular bacterial species produce characteristic VOC profiles.

Recent developments in genomics allow for unbiased assessment of bacteria present in clinical samples

through direct sequencing of DNA in each sample [11]. The corresponding gene function and taxonomic origin can be predicted by comparing the sequenced DNA fragments to a reference database containing sequences of known taxonomy and/or function [12]. This 'metagenomic' approach has been especially useful for characterizing the polymicrobial communities that exist in human-associated environments, e.g. lungs of individuals with cystic fibrosis (CF), where there is a high abundance of different types of bacteria due to decreased mucociliary clearance [13].

In a related study, 2,3-butanedione was implicated as a marker for microbial fermentation processes in the lungs [12]. In this study, we focused on the presence and quantification of ethanol, acetaldehyde and methanol in breath from CF patients. Ethanol and acetaldehyde were described as an important marker of positive sputum and broncho-alveolar lavage fluid cultures with any bacterium in several studies [19–21], and they are found in the breath of healthy people [22]. Ethanol and acetaldehyde were both described as general markers of bacterial growth in a systematic review of all *in vitro* studies [9]. Methanol is found in the breath of healthy people [23], and could have microbial or human origins. Methanol has been shown to regulate the production of detoxification genes in mice [24].

We hypothesized that bacterial DNA fragments in sputum of CF patients could be sequenced to identify whether the bacteria present were capable of producing VOCs found in the breath of these patients. For this purpose, we used gas-chromatography and mass-spectrometry to detect VOCs in breath and genomics and metagenomics to characterize bacterial genes in sputum.

Methods

Inclusion and obtained samples
This was an observational cohort study that included 7 patients with CF recruited at the University of California San Diego Adult CF clinic as described in previous studies [12, 14]. IRB approval was obtained from the University of California Institutional Review Board (HRPP 081500) and San Diego State University Institutional Review Board (SDSU IRB#2121). After informed consent was obtained we collected sputum and breath samples. Species identification (with 16S) and per species genetic potential (with metagenomic shotgun analysis) were studied on sputum samples. The 16S data have been published previously, along with some of the metagenomic data [12, 15, 16].

Breath analysis
From each CF patient and a healthy volunteer (in the same room), three breath samples were collected within 5 min of each other. Volunteers were instructed to eat and drink normally until 1 h before sample collection;

food and drink in the hours before sampling were recorded. Prior to sample collection, volunteers used saline to rinse their mouths, and volunteers were instructed to discard the first third of their breath (approximately 2 s), collecting the last two-thirds of the breath sample which is less influenced by the contents of the oral cavity. Samples were collected in 1.9L stainless steel canisters [17], along with a simultaneous background room air sample. Sample collection and analysis is described in Whiteson [12], using the approach detailed in Colman [18]. Briefly, breath samples were concentrated in a stainless steel loop submerged in liquid nitrogen, and then heated to ~80 C and split into six different columns. The data shown here were obtained from the DB-5 ms column (J&W; 60 m, 0.25 mm I.D., 0.5 mm film thickness) output to a quadrupole mass selective detector (MSD, HP5973). The absolute intensity information is not the major objective of the results shown here, but rather the relative amounts of the VOCs of interest in relation to the microbial sequence information obtained from parallel samples.
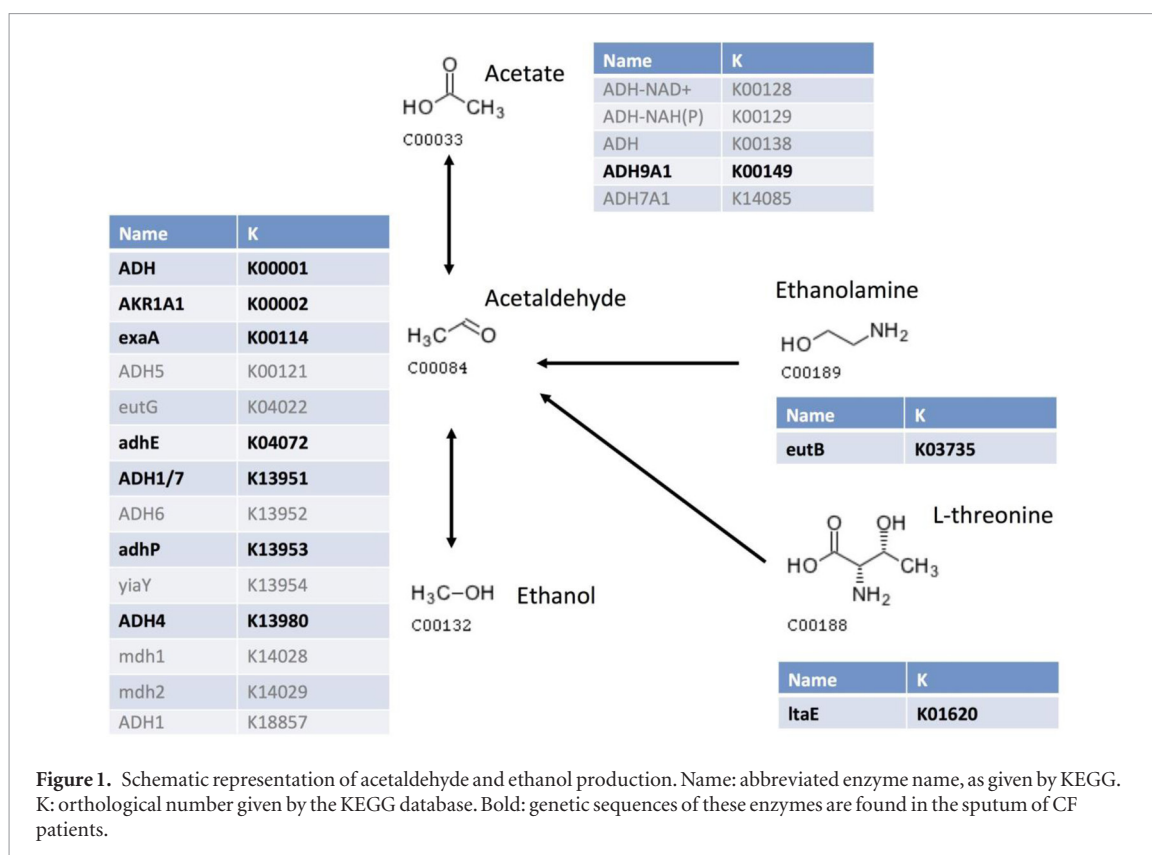
KEGG
The KEGG compound database was searched for ethanol (C00469), acetaldehyde (C00084) and methanol (C00132). The enzyme numbers (*ec:x.x.x.x*) associated with these compounds were extracted and coupled to the orthology database (*K*-numbers). For each of these *K*-numbers, the corresponding gene sequences in bacterial species that were previously described in the CF respiratory tract were extracted and then used to create a database.

16S rDNA analysis
The 16S rRNA gene amplicon data from these samples have been published, along with the associated methods [15]. Briefly, frozen sputum samples were thawed in the presence of Trizol reagent, DNA was extracted using a Life Technologies protocol, and amplicon sequencing of the V4 region (515F/860R) of the bacterial 16S rRNA gene was carried out at the RTSF Genomics Core at Michigan State University using the MiSeq v2 reagent kit for paired-end 250bp sequencing. Sequence data quality control with Prinseq [25] and a Mothur SOP for Miseq data [26] were employed to obtain 97% operational taxonimic units (OTUs), a proxy for bacterial species. Data processing and OTU clustering are described in further detail in Quinn *et al* [15].

Metagenomic analysis
Fresh sputum samples were processed using hypotonic lysis and washing to remove eukaryotic cells, damaged bacterial cells and extracellular DNA, as described previously [12, 14, 27]. Ion Torrent sequencing yielded 19.8 million reads, and 13.6 million reads were retained

Figure 1. Schematic representation of acetaldehyde and ethanol production. Name: abbreviated enzyme name, as given by KEGG. K: orthological number given by the KEGG database. Bold: genetic sequences of these enzymes are found in the sputum of CF patients.

after removal of low quality reads or human sequences [25, 28]. Some of these data were presented in previous publications that did not focus on breath analysis [15]. These metagenomic data were then compared to the previously constructed database (using BLASTn searches) to assist with assigning function to these data from CF patients. Metagenomic sequences were selected if they matched ethanol, acetaldehyde or methanol pathway genes from the KEGG with a minimum length of 40 bp, sequence identity of 40% and a BLAST *e*-value cutoff of 1×10^{-10}.

Data analysis

The number of sequences in metagenomic data from CF patients that matched sequences in the database was calculated per patient (normalized to the number of sequence reads for that sample), per microbe (on the genus level) and per orthological unit (*K*-number from KEGG database). Heatmaps were constructed using the *ggplot2* package within the R environment via the R-studio interface [29].

Results

Patients

Seven patients with CF were included from the UCSD Adult CF Clinic. Their ages ranged from 27 to 52 years. All patients were on chronic antibiotic therapy. Three patients received additional antibiotics in the week prior to sampling. The predominant microorganism cultured in the clinical microbiology lab did not often

reflect the dominant microbe in the sequence data [12, 15, 30].

Acetaldehyde

Acetaldehyde was present in the breath of all included patients. Figure 1 shows the enzymes that are involved in the metabolism of acetaldehyde. The enzymes that convert between acetate and acetaldehyde, and between ethanol and acetaldehyde are bi-directional. Bacteria in the genera *Escherichia*, *Lactococcus*, *Pseudomonas*, *Rothia* and *Streptococcus* were found to have the genetic potential to produce acetaldehyde (figure 2). Only DNA sequences from *Lactococcus* were implicated in the formation of acetaldehyde from acetate through aldehyde dehydrogenase family 9 member A1 (K00149). Most reads associated with threonine aldolase were associated with *Rothia*.
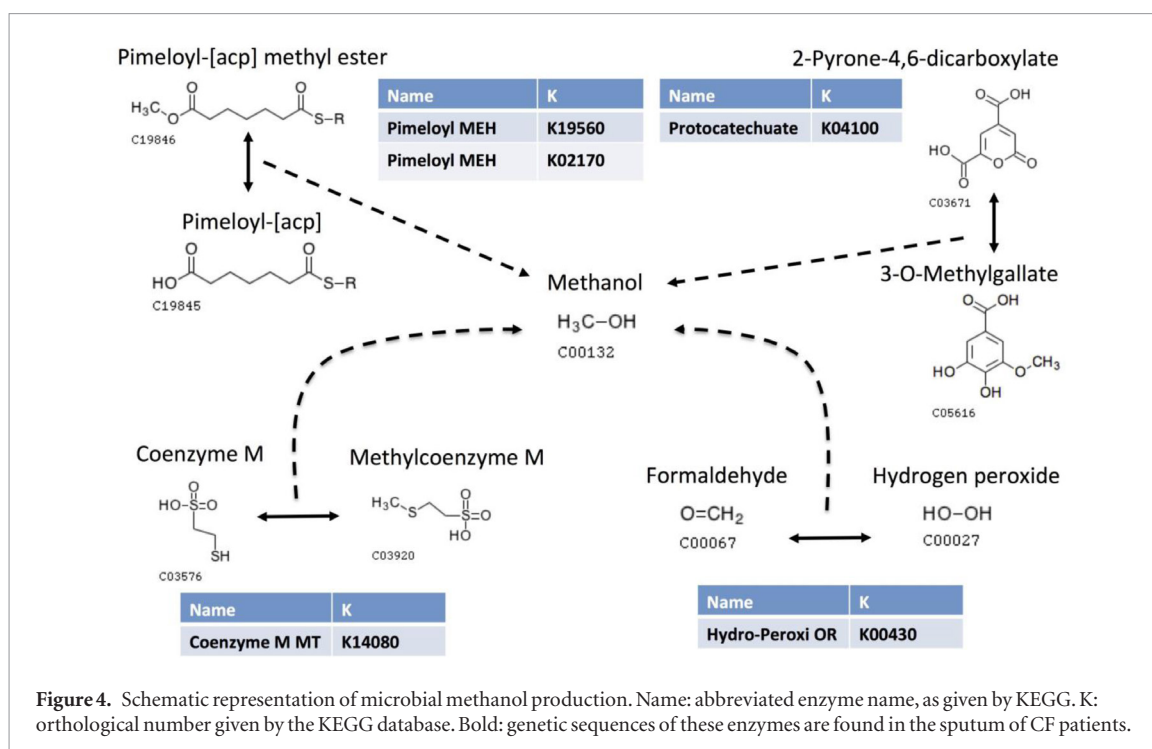
Ethanol

Ethanol was also present in the breath of all patients. Figure 1 also shows the enzymes that are associated with ethanol. Since ethanol is synthesized from acetaldehyde (and vice versa), most enzymes are similar to those described in the section on acetaldehyde. *Escherichia*, *Pseudomonas, Rothia* and *Streptococcus* were found to have the genetic potential to produce ethanol (figure 3). *Escherichia* was found to be genetically capable of producing ethanol in all patients, whilst there was considerable heterogeneity between patients for the other genera. The ethanol concentration in breath positively correlated with the amount of 16S rRNA
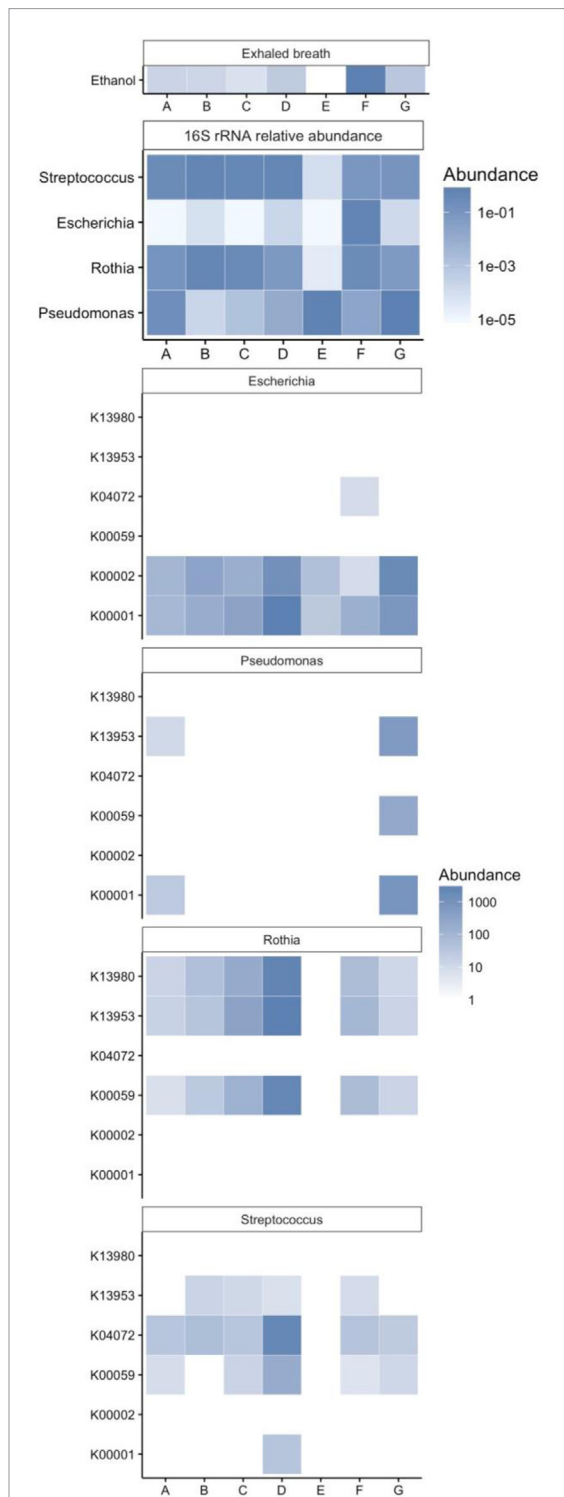
Figure 2. Genetic capabilities for ethanol production by different bacteria in the sputum samples of CF patients. Heatmaps of VOC concentrations and DNA fragment abundance. The intensity of blue represents the relative concentration/abundance. Upper panel: ethanol concentration in breath. Middle panel: relative abundance of *Streptococcus* species, *Escherichia* species, *Rothia* species and Pseudomonas species. Lower panel: number of DNA fragment copies per microbe, per enzyme (see figure 1), normalized to the number of reads for that sample.
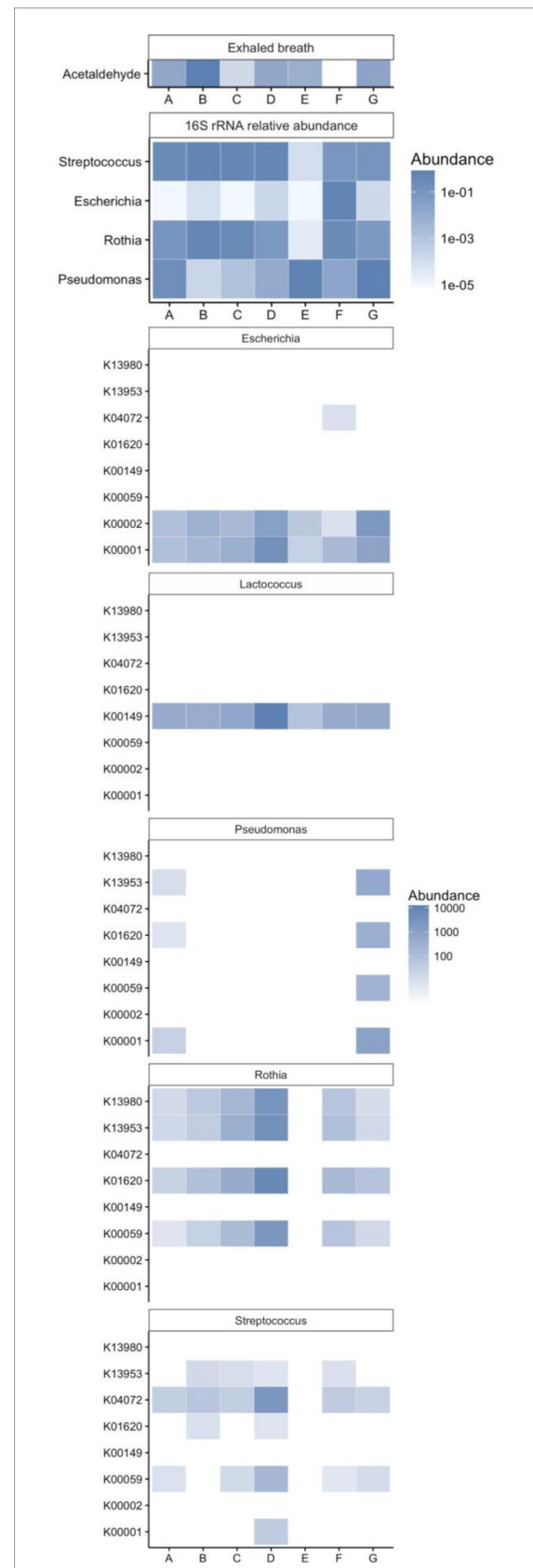
Figure 3. Genetic capabilities for acetaldehyde production by different bacteria in the sputum samples of CF patients. Heatmaps of VOC concentrations and DNA fragment abundance. The intensity of blue represents the relative concentration/abundance. Upper panel: acetaldehyde concentration in breath. Middle panel: relative abundance of *Streptococcus* species, *Escherichia* species, *Rothia* species and *Pseudomonas* species. Lower panel: number of DNA fragment copies per microbe, per enzyme (see figure 1), normalized to the number of reads for that sample.

Figure 4. Schematic representation of microbial methanol production. Name: abbreviated enzyme name, as given by KEGG. K: orthological number given by the KEGG database. Bold: genetic sequences of these enzymes are found in the sputum of CF patients.

copies of *Escherichia* found in sputum (Spearman rho = 0.85, P = 0.015).

Methanol

Figure 4 displays the enzymes that are linked to the production of methanol in the KEGG database for which matching sequences were found in the CF sputum samples. *Rothia* showed the most versatile genetic potential for producing methanol. DNA sequences that matched to *Pseudomonas*, *Streptococcus* and *Escherichia* were also linked to methanol production (figure 5).
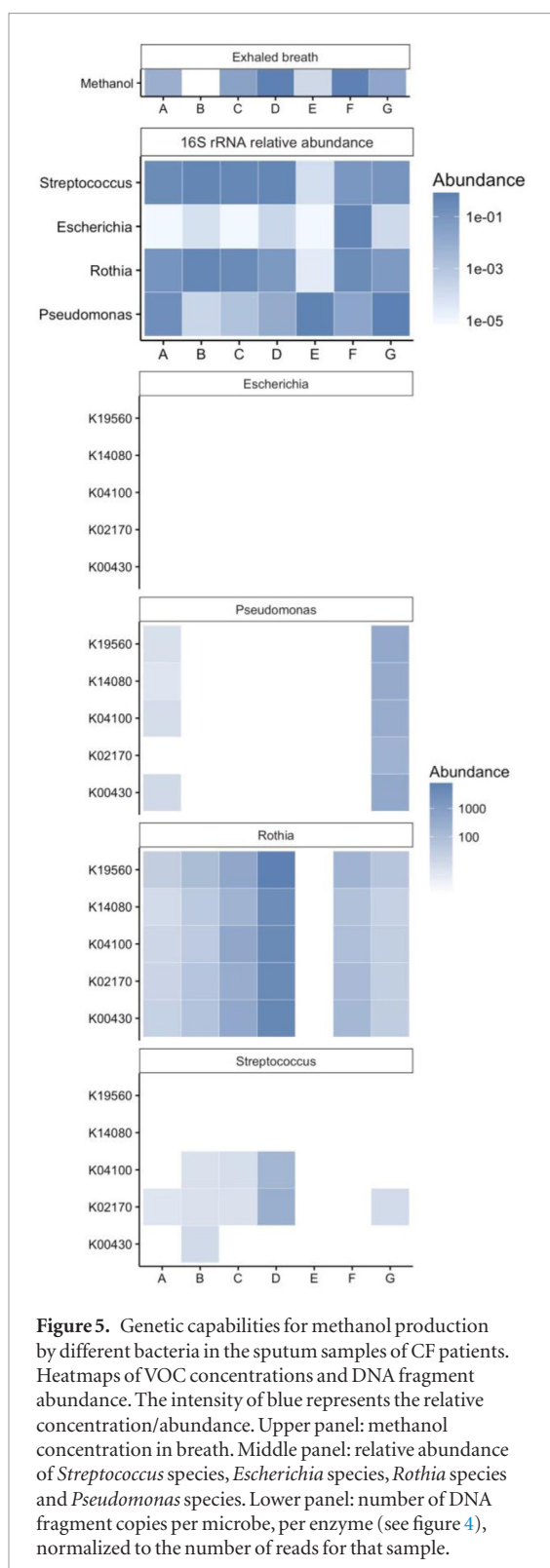
Discussion

The reported data support the hypothesis that bacterial DNA fragments in sputum of CF patients can be linked to gene functionality that would result in the production of VOCs that were also found in the breath of these patients. We found that the microbes present in these samples have the potential to produce ethanol, acetaldehyde and methanol, three VOCs that have been previously linked with airway colonization and pneumonia [19–21]. As a breath research community we should look at the pulmonary microbiome as a potential source for the VOCs we find in exhaled breath.

To our best knowledge, this is the second study to link VOCs in exhaled breath with their microbial producers via metagenomic analysis of the airway microbiome. In an earlier study drawing from the same samples, 2,3-butanedione in breath was linked to anaerobic metabolism of *Streptococcus* spp., *Rothia mucilaginosa*, and other acetoin metabolizers. 2,3-butanedione was more abundant during exacerbation, and decreased dramatically after antibiotic treatment; it may be a useful marker of exacerbation and of successful treatment in patients with CF [12]. This study further extends

those findings to include other VOCs that have been studied more widely as markers of bacterial colonization and infection of the respiratory tract. The role of microbes in the production of these compounds is not new; it has been studied particularly well in the gut and in biogas systems [31, 32]. Excess production of ethanol by dysbiosis of the microbiome in the gut is even recognized as a clinical syndrome called auto-brewery syndrome, which is highly disabling as patients are constantly intoxicated [33]. Local acetaldehyde production by bacteria in the gut has been implicated in the pathogenesis of colon carcinoma [34]. However, no research has been performed on the potential harmful effects of ethanol and acetaldehyde locally in the lung.

This study has several limitations. First, we focused only on three VOCs; acetaldehyde, ethanol and methanol, as a lot is known about their metabolism and therefore we could obtain gene sequences from the bacteria in the respiratory microbiome. These VOCs are of great interest as they are implicated as markers of colonization and/or infection of the respiratory tract [19–21]. However, other markers may be just as, or even more, clinically relevant to study in the future. For that, more knowledge on the metabolic pathways leading to these markers is required. There are also likely to be other, unknown, metabolic pathways that lead to the production of acetaldehyde, ethanol and methanol that are currently not listed in the KEGG database.

Another limitation is our use of breath and sputum from seven patients with CF. The generalizability of the results has to be tested in patients with other respiratory diseases and the validity has to be replicated in larger cohorts of patients. With increasing numbers of included patients, more quantitative analyses might be possible. Furthermore, we focused on bacterial production of VOCs. We cannot exclude that fungi are also

Figure 5. Genetic capabilities for methanol production by different bacteria in the sputum samples of CF patients. Heatmaps of VOC concentrations and DNA fragment abundance. The intensity of blue represents the relative concentration/abundance. Upper panel: methanol concentration in breath. Middle panel: relative abundance of *Streptococcus* species, *Escherichia* species, *Rothia* species and *Pseudomonas* species. Lower panel: number of DNA fragment copies per microbe, per enzyme (see figure 4), normalized to the number of reads for that sample.

The strengths of these studies are the prospective data collection and the meticulous analysis of the samples. The metagenomic analysis of the in tact bacterial DNA allowed for a comprehensive view on the pathways that are involved in the production of the VOCs. Furthermore, while sequencing DNA from a sputum sample in itself does not support the viability of the bacteria identified, as the DNA could be extracellular or come from dead cells, most bacteria identified in the metagenomes from the same samples were found to be viable in a capillary culture model, supporting the possibility that they are capable of producing the VOCs identified here [16].

While *Rothia* and *Streptococcus* may be converting ethanol to acetaldehyde, the presented data suggest that *Lactococcus*, from the order of *Lactobacilales* and a close cousin to *Streptococcus*, was the only bacterium in the lung of CF patients that contains annotated genes to metabolize acetaldehyde from acetate, and vice versa. Acetaldehyde leads to a pro-inflammatory, oxidative stress and carcinogenic response. The conversion to acetate protects against this [32, 40]. Acetate is a short chain fatty acid whose production from mucins by oral anaerobes may enable *Pseudomonas* to derive nutrients from mucins that it otherwise cannot access [47]. *Lactococcus* is also part of the healthy, normal oral and airways microbiome; the conversion of a toxic molecule (acetaldehyde) to an anti-inflammatory short chain fatty acid (acetate) could be an important role, complicated by the fact that acetate is a nutrient for *Pseudomonas*. The fact that volatile molecules produced throughout the airways and oropharyngeal cavity can travel means that they can impact the physiology of host and bacterial cells throughout the airways, even coordinating physiological events the way that hormones do in multicellular organisms, with interesting implications for both microbiology and human health.

An important caveat to determining the origin of a particular VOC found in breath is that many molecules can be produced by humans or bacteria, therefore it is not possible to determine whether the molecule has human or bacterial origin. Furthermore, bacterial physiology is heavily influenced by interactions with the host immune system [41]. For example, *Pseudomonas* and some other anaerobes can use nitrate as an alternative electron acceptor [42], and inflammation results in the recruitment of immune cells that produce nitrate [43–45]. In addition, neutrophils are thought to consume a large fraction of the oxygen available in the CF lung [46], forcing the colonizing microbes to rely on fermentation or anaerobic respiration, which is likely to be a large part of why growth rates in the CF lung are very slow—*Staphylococcus aureus* doubling times are slow and heterogeneous, averaging about two weeks in a recent study using the incorporation of stable isotope labels into bacteria isolated from CF sputum samples [18]. VOC production profiles of slow growing microbes is likely to differ from typical culture conditions, and emphasizes the importance of considering

capable of producing the studied metabolites. Indeed, it is suggested in the literature that ethanol and acetaldehyde is produced *in vitro* by *Candida albicans* [35–38]. We have also not studied the influence of substrates for the production of ethanol, methanol and acetaldehyde. Ingested ethanol is an obvious source; but this bias was excluded as the subjects did not drink alcohol around sample collection. However, non-alcohol beverages may also contribute [39].

physiologically relevant conditions and combining *in vitro* and *in vivo* approaches.

To conclude, bacterial DNA fragments in sputum of CF patients can be linked to enzymes implicated in the production of ethanol, acetaldehyde and methanol, which are VOCs that are predictive of respiratory tract colonization and/or infection. This confirms that the lung microbiome can produce VOCs directly.

Acknowledgement

References

[1] Boots A W, Bos L D, van der Schee M P, van Schooten F-J and Sterk P J 2015 Exhaled molecular fingerprinting in diagnosis and monitoring: validating volatile promises *Trends Mol. Med.* **21** 633–44

[2] Risby T H 2006 Current status of clinical breath analysis *Appl. Phys.* B **85** 421–6

[3] Frank Kneepkens C M, Lepage G and Roy C C 1994 The potential of the hydrocarbon breath test as a measure of lipid peroxidation *Free Radic. Biol. Med.* **17** 127–60

[4] Van Gossum A and Decuyper J 1989 Breath alkanes as an index of lipid peroxidation *Eur. Respir. J.* **2** 787–91

[5] Scholpp J, Schubert J K, Miekisch W and Geiger K 2002 Breath markers and soluble lipid peroxidation markers in critically ill patients *Clin. Chem. Lab Med.* **40** 587–94

[6] Leopold J H, van Hooijdonk R T, Sterk P J, Abu-Hanna A, Schultz M J and Bos L D 2014 Glucose prediction by analysis of exhaled metabolites—a systematic review *BMC Anesthesiol.* **14** 46

[7] Bean H D, Jiménez-Díaz J, Zhu J and Hill J E 2014 Breathprints of model murine bacterial lung infections are linked with immune response *Eur. Respir. J.* **45** 181–90

[8] Jiangjiang Z, Jaime J-D, Heather D B, Nirav A D, Minara I A, Lennart K A L and Jane E H 2013 Robust detection of *P. aeruginosa* and *S. aureus* acute lung infections by secondary electrospray ionization-mass spectrometry (SESI-MS) breathprinting: from initial infection to clearance *J. Breath Res.* **7** 37106

[9] Bos L D J, Sterk P J and Schultz M J 2013 Volatile metabolites of pathogens: a systematic review *PLoS Pathog.* **9** e1003311

[10] Lemfack M C, Nickel J, Dunkel M, Preissner R and Piechulla B 2014 mVOC: a database of microbial volatiles *Nucl. Acids Res.* **42** D744–8

[11] Garza D R and Dutilh B E 2015 From cultured to uncultured genome sequences: metagenomics and modeling microbial ecosystems *Cell. Mol. Life Sci.* **72** 4287–308

[12] Whiteson K L, Meinardi S, Lim Y W, Schmieder R, Maughan H, Quinn R, Blake D R, Conrad D and Rohwer F 2014 Breath gas metabolites and bacterial metagenomes from cystic fibrosis airways indicate active pH neutral 2,3-butanedione fermentation *ISME J.* **8** 1247–58

[13] Surette M G 2014 The cystic fibrosis lung microbiome *Ann. Am. Thorac. Soc.* **11** S61–5

[14] Lim Y W, Evangelista J S, Schmieder R, Bailey B, Haynes M, Furlan M, Maughan H, Edwards R, Rohwer F and Conrad D 2014

[15] Clinical insights from metagenomic analysis of sputum samples from patients with cystic fibrosis *J. Clin. Microbiol.* **52** 425–37

[15] Quinn R A, Phelan V V, Whiteson K L, Garg N, Bailey B A, Lim Y W, Conrad D J, Dorrestein P C and Rohwer F L 2015 Microbial, host and xenobiotic diversity in the cystic fibrosis sputum metabolome *ISME J.* **10** 1483–98

[16] Quinn R A, Whiteson K, Lim Y-W, Salamon P, Bailey B, Mienardi S, Sanchez S E, Blake D, Conrad D and Rohwer F 2014 A Winogradsky-based culture system shows an association between microbial fermentation and cystic fibrosis exacerbation *ISME J.* **9** 1024–38

[17] Kamboures M A, Blake D R, Cooper D M, Newcomb R L, Barker M, Larson J K, Meinardi S, Nussbaum E and Rowland F S 2005 Breath sulfides and pulmonary function in cystic fibrosis *Proc. Natl Acad. Sci.* **102** 15762–7

[18] Colman J J, Swanson A L, Meinardi S, Sive B C, Blake D R and Rowland F S 2001 Description of the analysis of a wide range of volatile organic compounds in whole air samples collected during PEM-tropics A and B *Anal. Chem.* **73** 3723–31

[19] Schnabel R, Fijten R, Smolinska A, Dallinga J, Boumans M-L, Stobberingh E, Boots A, Roekaerts P, Bergmans D and van Schooten F J 2015 Analysis of volatile organic compounds in exhaled breath to diagnose ventilator-associated pneumonia *Sci. Rep.* **5** 17179

[20] Fowler S J, Basanta-Sanchez M, Xu Y, Goodacre R and Dark P M 2015 Surveillance for lower airway pathogens in mechanically ventilated patients by metabolomic analysis of exhaled breath: a case-control study *Thorax* **70** 320–5

[21] Filipiak W *et al* 2015 Breath analysis for *in vivo* detection of pathogens related to ventilator-associated pneumonia in intensive care patients: a prospective pilot study *J. Breath Res.* **9** 016004

[22] Turner C, Spanel P, Smith D, Španěl P and Smith D 2006 A longitudinal study of ethanol and acetaldehyde in the exhaled breath of healthy volunteers using selected-ion flow-tube mass spectrometry *Rapid Commun. Mass Spectrom.* **20** 61–8

[23] Eriksen S P and Kulkarni A B 1963 Methanol in normal human breath *Science* **141** 639–40

[24] Komarova T V, Petrunia I V, Shindyapina A V, Silachev D N, Sheshukova E V, Kiryanov G I and Dorokhov Y L 2014 Endogenous methanol regulates mammalian gene activity *PLoS One* **9** e90239

[25] Schmieder R and Edwards R 2011 Quality control and preprocessing of metagenomic datasets *Bioinformatics* **27** 863–4

[26] Kozich J J, Westcott S L, Baxter N T, Highlander S K and Schloss P D 2013 Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform *Appl. Environ. Microbiol.* **79** 5112–20

[27] Lim Y W, Haynes M, Furlan M, Robertson C E, Harris J K and Rohwer F 2014 Purifying the impure: sequencing metagenomes and metatranscriptomes from complex animal-associated samples *J. Vis. Exp.* **94** e52117

[28] Schmieder R, Lim Y W and Edwards R 2012 Identification and removal of ribosomal RNA sequences from metatranscriptomes *Bioinformatics* **28** 433–5

[29] R Core Team 2013 *R: A Language and Environment for Statistical Computing* (Vienna: R Foundation for Statistical Computing)

[30] Quinn R A, Lim Y W, Maughan H, Conrad D, Rohwer F and Whiteson K L 2014 Biogeochemical forces shape the composition and physiology of polymicrobial communities in the cystic fibrosis lung *MBio* **5** e00956

[31] Campanaro S, Treu L, Kougias P G, De Francisci D, Valle G and Angelidaki I 2016 Metagenomic analysis and functional characterization of the biogas microbiome using high throughput shotgun sequencing and a novel binning strategy *Biotechnol. Biofuels* **9** 26

[32] Jokelainen K, Matysiak-Budnik T, Mäkisalo H, Höckerstedt K and Salaspuro M 1996 High intracolonic acetaldehyde values produced by a bacteriocolonic pathway for ethanol oxidation in piglets *Gut* **39** 100–4

[33] Dahshan A and Donovan K 2001 Auto-brewery syndrome in a child with short gut syndrome: case report and review of the literature *J. Pediatr. Gastroenterol. Nutr.* **33** 214–5

[34] Homann N, Tillonen J and Salaspuro M 2000 Microbially produced acetaldehyde from ethanol may increase the risk of colon cancer via folate deficiency *Int. J. Cancer* **86** 169–73

[35] Chen A I *et al* 2014 Candida albicans ethanol stimulates *Pseudomonas aeruginosa* WspR-controlled biofilm formation as part of a cyclic relationship involving phenazines *PLoS Pathog.* **10** e1004480

[36] Gainza-Cirauqui M L, Nieminen M T, Novak Frazer L, Aguirre-Urizar J M, Moragues M D and Rautemaa R 2013 Production of carcinogenic acetaldehyde by Candida albicans from patients with potentially malignant oral mucosal disorders *J. Oral Pathol. Med.* **42** 243–9

[37] Nieminen M T *et al* 2009 Acetaldehyde production from ethanol and glucose by non-*Candidaalbicans* yeasts *in vitro Oral Oncol.* **45** e245–8

[38] Filipiak W, Sponring A, Filipiak A, Baur M, Ager C, Wiesenhofer H, Margesin R, Nagl M, Troppmair J and Amann A 2013 Chapter 23—volatile organic compounds (VOCs) released by pathogenic microorganisms *in vitro*: potential breath biomarkers for early-stage diagnosis of disease *Volatile Biomarkers* (Oxford: Elsevier) pp 463–512

[39] Paiano V, Bianchi G, Davoli E, Negri E, Fanelli R and Fattore E 2014 Risk assessment for the Italian population of acetaldehyde in alcoholic and non-alcoholic beverages *Food Chem.* **154** 26–31

[40] Sapkota M and Wyatt T 2015 Alcohol, aldehydes, adducts and airways *Biomolecules* **5** 2987–3008

[41] Scales B S, Dickson R P and Huffnagle G B 2016 A tale of two sites: how inflammation can reshape the microbiomes of the gut and lungs *J. Leukoc. Biol.* **100** 1–8

[42] Line L *et al* 2014 Physiological levels of nitrate support anoxic growth by denitrification of *Pseudomonas aeruginosa* at growth rates reported in cystic fibrosis lungs and sputum *Front. Microbiol.* **5** 554

[43] Kolpen M *et al* 2014 Nitrous oxide production in sputum from cystic fibrosis patients with chronic *Pseudomonas aeruginosa* lung infection *PLoS One* **9** 1066–73

[44] Palmer K L, Aye L M and Whiteley M 2007 Nutritional cues control *Pseudomonas aeruginosa* multicellular behavior in cystic fibrosis sputum *J. Bacteriol.* **189** 8079–87

[45] Winter S E *et al* 2013 Host-derived nitrate boosts growth of *E. coli* in the inflamed gut *Science* **339** 708–11

[46] Kolpen M *et al* 2010 Polymorphonuclear leucocytes consume oxygen in sputum from chronic *Pseudomonas aeruginosa* pneumonia in cystic fibrosis *Thorax* **65** 57–62

[47] Flynn J M, Niccum D, Dunitz J M and Hunter R C 2016 Evidence and role for bacterial mucin degradation in cystic fibrosis airway disease *PLoS Pathogens* **12** e1005846