Analysis of microbiota in stable patients with chronic obstructive pulmonary disease

ESTEFANIA AGUIRRE,1 ANTONIO GALIANA,1 ALEJANDRO MIRA,2 RAMÓN GUARDIOLA,3 LAURA SÁNCHEZ-GUILLÉN,1 EDUARDO GARCIA-PACHON,4 MIGUEL SANTIBÁNEZ,5 GLORIA ROYO1,6 and JUAN CARLOS RODRÍGUEZ3,6

1Section of Microbiology, Hospital General Universitario de Elche, Alicante; 2Centro Superior de Investigación en Salud Pública (CSISP), Valencia; 3Section of Microbiology, Hospital General Universitario de Alicante, Alicante; 4Section of Respiratory Medicine, Hospital General Universitario de Elche, Alicante; 5Universidad de Cantabria, IFIMAV, Santander; and 6Universidad Miguel Hernández, Elche, Spain


To identify the bacterial diversity (microbiota) in expectorated sputum, a pyrosequencing method that investigates complex microbial communities of expectorated sputum was done in 19 stable chronic obstructive pulmonary disease patients (mean (SD) FEV1: 47 (18%) of predicted value). Using conventional culture, 3 phyla and 20 bacterial genera were identified, whereas the pyrosequencing approach detected 9 phyla and 43 genera (p < 0.001). In sputum the prevalent genera with pyrosequencing approach were Streptococcus, Actinomyces, Neisseria, Haemophilus, Rothia, Fusobacterium, Gemella, Granulicatella, Porphyromonas, Prevotella and Veillonella. Enterobacteriaceae, detected frequently in conventional culture, were not significantly detected with pyrosequencing methods. In addition, we found that important pathogens such as Haemophilus and Moraxella were detected more frequently with the new genetic procedures. The presence of Enterobacteriaceae is probably overestimated with conventional culture, whereas other difficult cultivable pathogens are underestimated. These studies open a new perspective for evaluating the role of bacterial colonization in chronic obstructive pulmonary disease pathogenesis and progression.

Key words: COPD; sputum microbiota.

Juan Carlos Rodríguez, S. Microbiology, Hospital General Universitario de Alicante, Alicante 03010, Spain. e-mail: rodriguez_juadia@gva.es

Chronic obstructive pulmonary disease (COPD), the fourth leading cause of death in the world, represents an important public health challenge that is both preventable and treatable. COPD is a major cause of chronic morbidity and mortality worldwide (http://www.goldcopd.org) (1). According to the EPI-SCAN study, The Epidemiologic Study of COPD in Spain (2), the prevalence of COPD in adults aged 40–80 years in Spain is estimated at 10.2%, although it varies widely across different geographical areas. An estimated 2 185 764 Spanish COPD patients are present among a population of 21.4 million aged between 40 and 80 years.

Microorganisms are one of the main aetiological factors involved in exacerbations of COPD. In contrast, understanding of their role during stable phases of the disease is still incomplete, although some studies have suggested that they actively contribute to chronic airway inflammation leading to the progression of COPD (3).

We performed a prospective observational study to identify the microbiota in the expectorated sputum of stable patients with COPD by pyrosequencing PCR-amplified bacterial 16S rRNA gene, a technique that provides a culture-independent analysis of all microbes in the sample. The aims of the present study were to describe the bacteria present in expectorated sputum using a pyrosequencing approach and to compare these findings with those obtained with conventional culture.

METHODS

Patient selection

Nineteen consecutive COPD patients attending the pulmonary clinic in a stable condition (at least 3 months without...
exacerbation or use of antibiotics for any other reason) capable of spontaneous expectoration were recruited. Patients with current or recurrent symptomatic ischaemic heart disease, congestive heart disease, cerebrovascular disease, dementia, lung cancer, known psychiatric illness, maintenance treatment with systemic corticosteroids (oral or parenteral), active tuberculosis, inflammatory bowel syndrome or insulin-dependent diabetes mellitus were excluded. Post-bronchodilator spirometry was performed in clinically stable patients (MasterLab Pro, Jaeger, Würzburg, Germany). Smoking habit (current smoking status, years of smoking, and number of packs of 20 cigarettes smoked per day) was noted. Patients who had stopped smoking for more than 6 months before recruitment were considered ex-smokers. Diagnosis and classification of COPD was established accordingly to GOLD recommendations (1).

All subjects were informed of the study objectives and gave their informed consent prior to their inclusion in the study. The research protocol was approved by the local ethics and research committees of the participating hospital.

**Sputum sampling**

Patients were asked to expectorate spontaneously; one-third of the good quality sample was used for quantitative bacterial culture, one-third was used for pyrosequencing and the remaining sample was kept in a biobank.

**Conventional microbiological studies**

For quantitative analysis (bacterial load), sputa were processed after liquefying with N-acetylcysteine. Serial dilutions (1/10, 1/100, 1/1000 and 1/10 000) were made and cultured at 37 °C for 48 h in Columbia blood agar, chocolate agar and MacConkey agar (Biomerieux, Marcy l’Etoile, France). BCYE agar (Biomerieux, Marcy l’Etoile, France) cultures were incubated for 20 days at 37 °C under humid atmosphere; all cultures were incubated under CO2 enriched atmosphere. No anaerobic cultures were performed. Bacterial isolates were identified by phenotypic methods and by partial sequencing of the 16S rRNA gene. Genus-level identification was obtained by genetic methods.

**DNA isolation**

For isolating bacterial DNA we used 300 μL of sample sputum and 100 μL of lysis buffer (10 mg/mL of lysozyme, 10% SDS, 10% CTAB and 25 μL of proteinase K) were added to each sample, then the mixes were incubated for 1 h at 37 °C and 20 min at 56 °C. The purification of nucleic acids was made with phenol-chloroform and absolute ethanol precipitation. DNA extraction quality was verified by measuring with Nanodrop (Thermo Scientific). The PCR program designed was; 94 °C for 5 min followed by 30 cycles of; 94 °C 30 s, 55 °C 30 s, 68 °C 30 s and finally 7 min at 68 °C. Before pyrosequencing, the amplicons were purified using a Minielute PCR purification Kit (Qiagen, Valencia, CA, USA) and after that, AMPure beads (Agencourt, Brea, CA, USA). The amplicons were pyrosequenced using a Roche GS-FLX Titanium with Lib-L type microspheres (Roche, Basel, Switzerland).

**Statistical analysis and assignment of OTU’s**

The pyrosequencing results were analysed bioinformatically using QIIME v1.6.0 software (4) to filter by quality and size the readings of the sequences obtained by the Roche 454 GS-FLX Titanium. After eliminating chimeric sequences from these readings, only sequences measuring between 250 bp and 500 bp with an end-trimming quality greater than 25 analysed in windows of 50 bases were included in the study.

The Operational Taxonomic Units (OTU’s) were assigned using the Ribosomal Database Project pyrosequencing pipeline (http://pyro.cme.msu.edu/) for the 16S rRNA gene with a bootstrap cutoff of 80% and only the OTU’s representing over 0.5% of the total sequences of each sample were considered. The tables were drawn up using the data obtained from the RDP classifier. Statistical analysis was performed using ANOVA test.

**RESULTS**

After filtering the sequences by quality and size and eliminating chimeric sequences, 20 648 sequences passed the quality filters. A mean of 1086 sequences per sample were analysed.

The individual characteristics of the 19 patients are described in supplementary Table 1. There were nine patients (48%) with moderate COPD and ten patients (52%) with severe COPD, seventeen were men and two women with a 71 years [SD 7.5] age mean. Nine patients were current smokers and ten were past smokers, with a smoking history average of 50.5 pack-years [SD 20.7].

With conventional cultures, 3 phyla and 20 bacterial genera were identified, while 9 phyla and 43 genera (p < 0.001) were identified by pyrosequencing. On average (SD) the number of genera in sputum was 4.3 (1.0) with conventional culture and 19 (8.0) with pyrosequencing approach (p < 0.001). After taxonomic assignment of the sequences obtained by pyrosequencing (using the RDP classifier), the best represented phyla were Firmicutes and Proteobacteria present in 19/19.

(5’-TTACC GGCGKGC CGTGGC A-3’). This region comprises the regions of gene hyper-variability from V1 to V3 of the 16S ribosomal gene used to classify bacterial species taxonomically. Each sample was assigned a unique multiplex identifier (MID). These amplicons were synthesized with the high-fidelity Extensor Long Range PCR Enzyme (Thermo Scientific). The PCR program designed was; 94 °C for 5 min followed by 30 cycles of; 94 °C 30 s, 55 °C 30 s, 68 °C 30 s and finally 7 min at 68 °C. Before pyrosequencing, the amplicons were purified using a Minielute PCR purification Kit (Qiagen, Valencia, CA, USA) and after that, AMPure beads (Agencourt, Brea, CA, USA). The amplicons were pyrosequenced using a Roche GS-FLX Titanium with Lib-L type microspheres (Roche, Basel, Switzerland).
samples, followed by Actinobacteria, Bacteroidetes and Fusobacteria present in 18/19, 16/19 and 15/19 samples respectively. Phyla as TM7, Tenericutes, Spirochaetes and Synergistetes were found only marginally and in less than 10/19 samples. Members of Streptococcus, Fusobacterium, Prevotella, Rothia, Veillonella, Porphyromonas and Haemophilus genera dominated individual sputum samples. Other members as Haemophilus, Pseudomonas or Corynebacterium were dominant in samples 15S and 17M; 8S; and 20S respectively. Actinomyces genus was mainly found in samples of patients from moderate group (8/9), while in samples of patients from severe group was only found in 2/10 (p < 0.01). The distribution of the genera in the different samples is shown in Fig. 1.

Microbial culturing analysis from sputum samples showed the presence of $10^6$ to $10^{11}$ cfu/ml cultivable microorganisms per sample. The phenotypic identification and partial sequencing of the 16S rRNA gene showed the best represented phylum was Proteobacteria with 11 different genera, followed by Actinobacteria with 5 different genera and the less represented phylum was Firmicutes with 4 genera although the Streptococcus genus was present in all samples. Streptococcus and Neisseria were the most frequent genera found in 19/19 and 11/19 samples, respectively, followed by both Actinomyces and Rothia found in 9/19 samples. Other genera were predominant in one sample as Haemophilus (sample 17M), Pseudomonas (sample 8S) and Corynebacterium (sample 20S) (data shown in Table 1).

When comparing the two methodologies we noted that genera such as Actinomyces or Rothia were more frequently identified by the pyrosequencing assay than by conventional methods. Difficult cultivable genera as Treponema, Mycoplasma, Eikenella, Kingella or Capnocytophaga, were detected by pyrosequencing and not detected by conventional culture. Anaerobic genera as Veillonella, Prevotella, Leptotrichia, Fusobacterium, Porphyromonas, Pascardiavia and Bulleidia were only detected by pyrosequencing approach due to conventional culture were carried only under aerobic conditions according to standard clinical microbiological protocols.

**DISCUSSION**

Sputum culture has been considered a reliable method of obtaining microbiological data (5) and, despite its limitations, remains the predominant microbiological investigation performed in the
clinical setting in patients with chronic obstructive pulmonary disease (COPD) (6). Bacterial colonization in COPD has been shown to be an important factor in disease progression and exacerbations (7, 8).

However, conventional culture methods are not capable of isolating fastidious or emerging bacteria. It would be desirable to identify all the bacterial diversity (microbiota) in expectorated sputum so as to determine its role in inflammation, disease progression or clinical characteristics.

Thanks to recent advances, next-generation DNA-sequencing technologies have greatly enhanced capabilities for sequencing large meta-data sets (9), and provide an unprecedented opportunity to investigate the complex microbial communities associated with the human body (10). This method has been applied in several studies in sputa and other respiratory samples. In patients with asthma, microbiota composition and diversity were significantly correlated with bronchial hyperresponsiveness (11). In patients with cystic fibrosis, metagenomic studies demonstrated the presence of multiple new pathogens (12). Sze et al. (13) compared the COPD lung to the cystic fibrosis lung to demonstrate microbiome differences between these two lung diseases. These findings are also consistent with the previously published individual culture-independent studies of the cystic fibrosis lung microbiome and COPD microbiome, not to mention consistency with older culture-based studies that highlight differences in the microbiology of these two respiratory diseases (14). Few data are available on COPD patients. Hilty et al. (15), reported some differences in the microbiota of five COPD patients compared with controls. In eight patients intubated due to COPD exacerbation, Huang et al. (16) described a great diversity of bacteria. In a study with a small sample of COPD patients, Erb-Downward et al. (17) reported limited community diversity in moderate to severe patients compared with controls.

This study using next-generation sequencing provides novel information on the microbiota of sputum in COPD patients in a stable clinical situation. As expected, metagenomics was capable of detecting a significantly higher number of bacteria than was conventional culture (Table 1 and Fig. 1). It thus reflects more accurately the wide bacterial biodiversity present in the respiratory mucosa of these patients since over 70% of the bacteria present in the respiratory mucosa cannot be cultured (18, 19). Our data confirm the strong evidence that culture-based methods detect only a small proportion of bacteria present in the respiratory tracts, as has been previously reported in patients with cystic fibrosis (12).

Of the seven patients in whom Enterobacteriaceae were found by conventional culture, in only one was the isolation confirmed by pyrosequencing (Table 1, Fig. 1). A similar discordance was seen when evaluating patients in whom the genus Actinomyces was present by conventional culture compared with pyrosequencing results. This is probably due to the fact that these genera are overestimated by culture even though the patients exhibit very low bacterial loads since the primers used are capable of amplifying them (20). This is of clinical interest since Enterobacteriaceae are considered to be

Fig. 1. Relative bacterial composition at the genus level in moderate and severe (M or S) COPD patients.
important pathogens in COPD (15) and are frequently isolated in conventional culture, as in our sample. Since this may modify patterns of antibiotic treatment, further studies may explain the role of this genus in COPD patients.

On the other hand, pathogens such as *Haemophilus* and *Moraxella*, whose pathogenicity in these patients has been demonstrated in multiple studies, are more often detected by molecular methods due to the difficulties involved in isolating these microorganisms using conventional culture-based methods. This higher prevalence should also be analysed in greater depth in order to optimize the antibiotic treatment normally administered empirically in these patients. With this new methodology such optimization would be possible (21).

Metagenomic methods allowed us to identify anaerobes in expectorated sputum. In our series, all patients had anaerobes, which accounted for about 15% of the bacterial load. The clinical significance of this finding needs to be established in the light of these new tools that make it possible to obtain a much better understanding of this complex group of bacteria which to date are little understood due to the serious limitations of culture-based methods for isolating these microorganisms. Very few metagenomic analysis data are available in patients with COPD. Compared with healthy individuals and using bronchoscopy samples, Hilty et al. (15) reported in five COPD patients a higher isolation of phylum *Proteobacteria* (including *Haemophilus*) and lower proportion of phylum *Bacteroidetes* (mainly *Prevotella*). Interestingly, a decrease in total bacterial load and a diminished quantity of Bacteroidetes has also been described in inflammatory bowel disease (22). It is known that *Prevotella* spp. inhibit the growth of other bacteria (commensal flora) that may play a role in immune response (15). In our series, *Prevotella* spp was present in 90% of COPD patients, but at a very low proportion (about 4% of bacterial load). Huang et al. (16) studied with metagenomic technology eight intubated COPD patients with exacerbation and reported a greater diversity of bacteria than previously appreciated in the airways of these patients.

Our sample of patients probably does not represent all COPD patients because obviously we only included those able to expectorate spontaneously. However, in a large study, the frequency of sputum expectoration was similar in all proposed subtypes of COPD (23). Nevertheless, it would be very interesting to investigate the lung microbiota with other methods (bronchoalveolar lavage, surgical samples) in patients with and without expectoration or other clinically relevant symptoms or characteristics (24).

Microbiota in stable COPD patients shows a great diversity of bacteria that differ from bacteria detected in conventional cultures. Thus, metagenomic studies open a new perspective for evaluating the role of bacterial colonization in COPD pathogenesis and progression and may contribute to a more comprehensive understanding of this serious disease (9, 25) since the lung is a complex structure composed of cells belonging to all three domains of life on Earth, Eukarya, Bacteria and Archaea, as well as their viruses. The microbiome constitutes the last human organ under active research. Like any other organ, the microbiome has its own pathophysiology, and individual health might be damaged when its collective population structure is altered (26).

This study was supported by grants ‘Ayudas para la realización de trabajos de investigación en materia de atención de media y larga estancia. Agencia Valenciana de Salud. 2010’ (MLE 2/10), ‘Fundación de la Comunidad Valenciana para la investigación biomédica, la docencia y la cooperación internacional y para el desarrollo del Hospital General Universitario de Elche’ (FIBELX-C011/02), ‘Consejería de Sanitat. Generalitat Valenciana’ (AP-168/11), ‘Ayuda de investigación de la Fundación Bienvenida Navarro Luciano Trijodi 2010–2011 y 2011–2012’ and from the Spanish Ministry of Health (Instituto Salud Carlos III P112/02209).

REFERENCES


SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table S1. Characteristics of patients as a function of COPD status.