New microbiota found in sputum from patients with community-acquired pneumonia

Chunyan Chen¹, Tian Shen¹, Fei Tian¹, Ping Lin², Qingtian Li³, Zelin Cui¹, Yajuan Zhang⁴, Min Xue⁴, Jing Ye⁵, Xiaokui Guo¹, and Yuhua Zhou¹,⁵*

¹Department of Medical Microbiology and Parasitology/Department of Preventive Medicine, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China
²Department of Medical Laboratory, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai 200030, China
³Department of Laboratory Medicine, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China
⁴Department of Respiratory, Shanghai Minhang Central Hospital, Shanghai 201100, China
⁵Department of Emergency Medicine, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China
*Correspondence address. Tel/Fax: +86-21-64453285; E-mail: talent_zhou@sjtu.edu.cn

Community-acquired pneumonia (CAP) is a major concern in hospitals and the bacterial community of which has not been systemically discussed yet. Sputum from patients in the acute stages is a kind of accessible sample reflecting its features. In our study, we analyzed 45 sputum samples from 45 patients with CAP. Eighteen sputum samples from healthy people were chosen as the controls. Pyrosequencing of the 16s rDNA V3 hypervariable regions of all the bacteria contained in the sputum was used as a culture-independent method to disclose the community constitution. Also, our published data for hospital-acquired pneumonia (HAP) in sputum was used for comparison. By pyrosequencing, >90,000 DNA reads were detected. After being analyzed by tools in the Ribosomal Database Project, the reads were classified into five main phyla and >100 genera. At the phyla level, the reads’ distribution of CAP is similar to that of healthy people and at genera level, the occurrence of each genus possesses their feature in three categories. Genera such as Streptococcus and Neisseria showed stability in their percentages, indicating that such genera are rarely affected by exogenous bacteria or antibiotics. The role of other genera such as Moraxella and Rothia in CAP should be emphasized. According to our analysis, the bacterial communities of CAP are with slight change when compared with those of healthy people, but have a large gap between HAP. Meanwhile, Rothia might be an important endogenous pneumonia-causing factor.

Keywords 16s ribosomal DNA; microbiota; community-acquired pneumonia; pyrosequencing; sputum

Received: June 23, 2013 Accepted: August 5, 2013

Introduction

In clinical medicine, community-acquired pneumonia (CAP) is one of the most common illnesses that patients experience. CAP is characterized by clinical findings such as fever, productive cough, laboring breath, and newly emerging infiltrative or consolidative shadow on chest radiograph [1–3]. In current knowledge of pulmonary infection, the leading pathogenic agents are Streptococcus, Haemophilus, Staphylococcus, Mycoplasma, and Chlamydia [1]. Under such principle, the clinical culture-dependent or -independent methods for pathogen identification are the most wildly used passways to describe the infection status [4–6].

With the developing step of the understanding on infection, the pulmonary infections are inclined to be thought of as multi-factors-induced infections rather than a single pathogen infection. Much evidence has been put forward to prove the existence of co-infection [7]. Furthermore, Blaser and Falkow [8] revealed a new concept of ‘community as a pathogen’ in a recent review. The infection is thought to be a pathogenic state resulting from the participation of all bacteria colonizing the human body and the disease is a type of balance-losing state rather than a consequence caused by a single specific agent or a limited band of specific agents. For a deep understanding of CAP, it is necessary to disclose the constitution of bacterial community in this infection. Although the bacterial population in the healthy human respiratory tract has been discussed [9], the CAP’s bacterial population has not been systemically defined by now.

Although sputum is not the best sample reflecting the bacterial community in lungs because of contamination by the flora colonizing the upper respiratory tract and oral cavity, it is the most common sample to be taken for clinical examination because of its non-invasiveness and convenience in collecting [10]. Obviously, the commonly used traditional clinical examinations are not feasible to unveil the panorama in lungs. The method based on massive parallel pyrosequencing of bacterial 16s rDNA amplicons of V3 regions is a common method in microecology, and its feasibilities have been proven in other infections [11–15]. So we considered...
that the combination of the sputum samples and 16s V3 region pyrosequencing is an easy and non-invasive way to reveal the infective status in lungs.

The objective of our work is to analyze the bacteria community in CAP patients’ sputum sample by pyrosequencing and discuss the gap of bacteria community between healthy people (the data were also obtained by pyrosequencing using saliva samples) and the patients of HAP (published data [16]). By comparing the studies among different populations, some preliminary conclusions were drawn, which offer an insight into the causing factors of pneumonia.

Materials and Methods

Ethics statement

The study was approved by the ethics committee of Shanghai Minhang Central Hospital. The consents were obtained in verbal pattern, because all the samples were routinely collected for clinical laboratory examination under the doctor’s prescription and our laboratory work started after the examination. The study did not involve any private information of the patients or any adjustment or difference in clinical treatment. Therefore, because all the samples were not intended to be collected specially for this research, but a step after the routine clinical examination, the ethics committee of Minhang Central Hospital approved such procedures and believed that there was no requirement of the ethics committee’s ratification. The volunteers for normal control were recruited from the relatives of the members of our work group. All of them consented to support our research work and were informed of all the details of the usage of their samples.

Patients and healthy people

Forty-five CAP patients (42–70 years old) were from the In-patient Department of Shanghai Minhang Central Hospital’s respiratory diseases ward. The characteristics of the patients are presented in Supplementary Table S1. When samples were collected, the patients were in the acute stage with pneumonia symptoms such as fever, cough, yellowish sputum, and newly emerging infective shadows on chest X-ray. The patients had not taken antibiotics in the previous 3 months without special disorders such as immunocompromising which would have greatly influenced the evaluation of the health state [10]. Furthermore, one important thing that should be noted here is that aside from the routine clinical examination, the Wright’s stained sputum smear examination was also carried out; the sample with white blood cells counting ≥3 cells per oil lens field was considered to be qualified.

Because of coughing, the community originally in sputum is mainly contaminated by the normal flora in oral cavity and pharynx. In order to make comparison, samples from healthy people which were a mixture of saliva and pharynx secretions were collected during coughing. The normal people enrolled were in the same age range. All the healthy participants were free of basic pulmonary diseases, such as chronic obstructive pulmonary disease or special oral infectious disease.

The data of HAP patient are from our published article [16], which were obtained by the same methods as reported in this paper.

Barcoded primers

Polymerase chain reaction (PCR) enrichment of the 16S rDNA V3 hypervariable region was performed with forward primer 5'-CTCTTACT-TACGGGAGGCAGCAG-3' and reverse primer 5'-CTCTTACT-ATTACCCGGCTGCTGG -3'. The 5'-terminal of every primer contained an 8-base oligonucleotide tag (before the hyphen), while the sequence after the hyphen was able to pare the sequences of the V3 end region.

Samples treatment

After sample collection, 1 mol/L sodium hydroxide solution—approximately five times the volume of the sample—was added to each sample for liquefaction. The samples were then stored at 4°C for 12–18 h to ensure enough time for basic digestion.

DNA extraction from sputum samples

DNA isolation was carried out predominantly with the UltraPure™ Genome DNA Kit (Sbs, Beijing, China) in accordance with the manufacturer’s instructions. However, several crucial steps of the extraction (including digestion with lysozyme, lysostaphin, and proteinase K) were enhanced because the sputum was difficult to work with. After storage at 4°C for 12–18 h, each sample was centrifuged at 10,000 g for 10 min. All of the supernatant was removed, and the sample was resuspended in phosphate-buffered saline. Lysozyme and lysostaphin were added into each diluted sample to a final concentration of 5 mg/ml and 32 U/ml, respectively. The samples were then incubated at 37°C for 4 h. Then, proteinase K was added with a concentration of 0.1 mg/ml, and the samples were incubated at 65°C for 2 h. The samples were centrifuged at 10,000 g for 10 min. Subsequently, all of the supernatant, containing both bacterial and human DNA, was collected and transferred to sterilized 1.5 ml Eppendorf tubes for further steps and the extracted DNA from 101 samples was produced.

PCR enrichment of the V3 region

To obtain enough PCR products, a two-step PCR strategy was used. The first step was in a 25 μl reaction volume containing 2.5 μl of PCR buffer (TAKARA, Dalian, China), 0.625 U ExTaq (TAKARA), 0.1 μl of bovine serum...
albumin (BSA, TAKARA), and 2 μL of primer solution, with 100 μmol of each forward and reverse primer. Fifty nanogram of extracted DNA was added as the template. Finally, ddH2O was added to reach the final volume of reaction. Touchdown PCR conditions were as follows: 5 min at 94°C for initial denaturation, 1 min at 94°C for denaturation, 1 min at 65°C for annealing, and 1 min at 72°C for extension, with the annealing temperature decreasing by 0.5°C for each step of the last 20 cycles. The reaction volume in the second step of the PCR was 50 μL. The products from step 1 were used as the template for this reaction with 5 μL of PCR buffer (TAKARA), 1.25 U ExTaq (TAKARA), 0.2 μL of BSA (TAKARA), 24 μL of water, and 200 μmol of each bar-coded forward and reverse primer. The heating cycle lasted for 1 min at 94°C for denaturation, 1 min at 55°C for annealing, and 1 min at 72°C for extension for five cycles with the temperature kept at 20°C after the reaction.

Harvesting of purified products and establishing the pyrosequencing library
After the PCR reaction, the electrophoresis was immediately performed to isolate the enriched V3 region DNA fragments from the reaction mixture. All of the products were ~200 bp in length and were harvested using a gel extraction kit (Gel Extraction Kit, OMEGA Bio-Tek, Norcross, USA) following the manufacturer’s instructions. All of the enriched DNAs were mixed to establish the library.

Pyrosequencing using the 454 platform
Sequencing was undertaken by Roche 454 FLX in accordance with the manufacturer’s instructions, and the gross sequencing data were arranged by primer tags [17]. All the sequencing works were performed at the Chinese National Human Genome Center in Shanghai.

Submission to GenBank
For all the sequences, the removal of redundancy has been undertaken by using Jalview software at redundancy threshold 97%. Seventy-eight representative sequences were submitted to Genbank and got the accession numbers from HQ914698 to HQ914775.

Bioinformatics and statistical analysis
All of the datasets were taxonomically grouped using the RDP (the Ribosomal Database Project’s Naïve Bayesian) classifier at a confidence level of 90% [18,19]. The sequences could be assigned when the genus level in the bacteria domain was collected and screened in Microsoft Excel. The mean standard deviation and Wilcoxon signed-rank test were performed with SPSS 11.5 software (SPSS, Chicago, USA). Principal coordinate analyses and neighboring joint tree were made by using Unifrac (http://bmf2.colorado.edu/fastunifrac), and the heatmap was generated by R language [20].

Results
Total quantity of reads detected by 454 sequencing
In the samples from CAP patients and normal people, 454 sequencing detected a total of 94,951 PCR reads of ~200 bp in length. After RDP analysis, 90,152 sequences could be classified into bacterial phylum level, and in which 57,585 reads could be classified up to the genus level.

Difference of the microbiota from three categories described on the phyla and genera levels
To compare the general characteristics of the bacterial community among these three categories, the reads belonging to each bacterial phylum were counted and the average percentages of each phylum in every category of population were calculated. The feature of the reads distribution at phyla level is shown in Fig. 1. The bacterial communities belong to five main phyla, namely Firmicutes, Bacteroidetes, Proteobacteria, Fusobacteria, and Actinobacteria.

To disclose the features of each group at genera level, the percentages of every genus in each patient were calculated. Subsequently, the mean percentages of each genus were obtained. Figure 2 is a heatmap showing the reads’ genera-level distribution features of all the individuals. We
found that bacterial genera such as Rothia and Prevotella varied obviously among the individuals in different categories. Table 1 shows the most prevalent 30 genera in the three groups. Genera such as Pseudomonas, Staphylococcus, Moraxella, Acinetobacter, Klebsiella, and Stenoprohomonas are often found in patients with pulmonary infection; they are absent from the list of normal people and occupy a high position in the list of two kinds of pneumonia. Even the same genera rank differently in CAP and HAP. However, the concentrations of some bacteria, such as Streptococcus and Neisseria, are similar among all the three groups.

Statistical methods disclosing the changes of genera in different categories

To objectively compare the microflora in these three groups, we used Wilcoxon signed-rank test by SPSS to compare the percentage of genera that exist simultaneously in both CAP and HAP patients or in healthy people and CAP patients. Tables 2 and 3 show the statistical results of the genera of major concern (please see the details in the Supplementary Tables S2 and S3). These tables indicate that between healthy people and CAP patients, the percentages of most genera changed slightly and no significant difference exists ($P \geq 0.05$). Between the CAP and HAP groups, although many genera were in a crossover area, the genera’s percentages exhibited significant difference ($P < 0.05$).

To reveal the relationships among individuals, the cluster analysis including neighbor-joining tree and principle component analysis (PCoA) were undertaken; the results are shown in Fig. 3. Judged from the tree and the scatter plots figures, three categories of individuals could be clearly identified.

Discussion

The method based on barcoded primers and 454 pyrosequencing makes it possible to determine the profile for each enrolled person [21–23].

Based on the phyla level as shown in Fig. 1, we found that in the communities of the CAP patients, the proportions of Actinobacteria and Proteobacteria became higher when compared with the corresponding parts of the healthy people. The leading phylum (Firmicutes) was nearly equal to that of healthy people. In HAP, the situation was completely changed because of the obvious expansion of Proteobacteria. As a result, two leading phyla, Firmicutes and Proteobacteria, appeared, and the proportions of the other phyla were extremely suppressed. All the above results indicated that the overall communities in healthy people and CAP patients are similar, but are quite different from those in the HAP.

As the data were further analyzed to the genus level, more detailed conclusions could be drawn. According to Fig. 2, it was found that many red cells, representing the overwhelming dominance of a certain genus, appeared in the area of CAP and HAP, which suggested that some disorders result...
from the excessive growth of a certain genus and lead to the balance-losing state. When comparing the percentages of genera that existed in both healthy and CAP patients (Table 2), there was no significant difference (29 genera’s $P \geq 0.05$ out of 36 genera). However, when the genera existing both in CAP and HAP were compared, more than half of the genera showed differences (30 genera’s $P < 0.05$ out of 51 genera, Table 3). So another conclusion can be drawn that in the CAP patient, the occurrences of most natural parts of the community do not change significantly, but several obviously changing genera such as Rothia possibly act as the endogenous infecting agents. On the other hand, although there are many genera existing simultaneously in both CAP and HAP patients, most genera’s occurrences change significantly, which suggests that the situation in HAP patients has changed greatly. Therefore, the genera-distributing feature suggests that certain genera play the leading role in causing disorders in CAP patients; conversely, the infection-causing factor is shared by more genera in HAP patients.

The genera that appear simultaneously in three categories are more likely to be endogenous residents in vivo and their distribution features are very different in three categories (Fig. 4). In the Wilcoxon signed-rank test, we divided the cohort of genera into three groups. The first group includes the genera whose percentage remained stable in different populations and in which all the $P$ values were $>0.05$. Two important genera meet this standard: *Neisseria* ($N$: 1.78%; $C$: 2.15%; $H$: 1.61%) and *Streptococcus* ($N$: 21.67%; $C$: 20.58%; $H$: 20.13%). The results show that these genera are not significantly affected by exogenous bacteria or bactericidal antibiotics. The second category is the genera whose

<table>
<thead>
<tr>
<th>Ranking</th>
<th>Health</th>
<th>Mean (%)</th>
<th>CAP</th>
<th>Mean (%)</th>
<th>HAP</th>
<th>Mean (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Streptococcus</td>
<td>21.67</td>
<td>Streptococcus</td>
<td>20.58</td>
<td>Streptococcus</td>
<td>20.13</td>
</tr>
<tr>
<td>2</td>
<td>Prevotella</td>
<td>10.11</td>
<td>Rothia</td>
<td>11.35</td>
<td>Staphylococcus</td>
<td>8.10</td>
</tr>
<tr>
<td>3</td>
<td>Haemophilus</td>
<td>4.00</td>
<td>Prevoetella</td>
<td>9.90</td>
<td>Pseudomonas</td>
<td>7.64</td>
</tr>
<tr>
<td>4</td>
<td>Veillonella</td>
<td>3.62</td>
<td>Veillonella</td>
<td>3.60</td>
<td>Acinetobacter</td>
<td>6.21</td>
</tr>
<tr>
<td>5</td>
<td>Fusobacterium</td>
<td>3.39</td>
<td>Pseudomonas</td>
<td>3.49</td>
<td>Rothia</td>
<td>4.84</td>
</tr>
<tr>
<td>6</td>
<td>Capnocytophaga</td>
<td>3.24</td>
<td>Gemella</td>
<td>2.98</td>
<td>Gemella</td>
<td>3.92</td>
</tr>
<tr>
<td>7</td>
<td>Granulicatella</td>
<td>3.14</td>
<td>Granulicatella</td>
<td>2.70</td>
<td>Prevotella</td>
<td>3.83</td>
</tr>
<tr>
<td>8</td>
<td>Gemella</td>
<td>2.81</td>
<td>Neisseria</td>
<td>2.15</td>
<td>Granulicatella</td>
<td>3.01</td>
</tr>
<tr>
<td>9</td>
<td>Neisseria</td>
<td>1.78</td>
<td>Capnocytophaga</td>
<td>1.91</td>
<td>Stenotrophomonas</td>
<td>2.70</td>
</tr>
<tr>
<td>10</td>
<td>Leptotrichia</td>
<td>1.75</td>
<td>Fusobacterium</td>
<td>1.66</td>
<td>Ensifer</td>
<td>2.37</td>
</tr>
<tr>
<td>11</td>
<td>Actinomyces</td>
<td>1.47</td>
<td>Haemophilus</td>
<td>1.65</td>
<td>Corynebacterium</td>
<td>1.81</td>
</tr>
<tr>
<td>12</td>
<td>Rothia</td>
<td>1.04</td>
<td>Actinomyces</td>
<td>1.06</td>
<td>Neisseria</td>
<td>1.61</td>
</tr>
<tr>
<td>13</td>
<td>Aggregatibacter</td>
<td>0.87</td>
<td>Aeromonas</td>
<td>1.05</td>
<td>Enterococcus</td>
<td>1.40</td>
</tr>
<tr>
<td>14</td>
<td>Campylobacter</td>
<td>0.80</td>
<td>Staphylococcus</td>
<td>1.01</td>
<td>Streptococcus</td>
<td>1.30</td>
</tr>
<tr>
<td>15</td>
<td>Porphyromonas</td>
<td>0.72</td>
<td>Leptotrichia</td>
<td>0.97</td>
<td>Actinomyces</td>
<td>1.20</td>
</tr>
<tr>
<td>16</td>
<td>Selenomonas</td>
<td>0.72</td>
<td>Porphyromonas</td>
<td>0.96</td>
<td>Limmobacter</td>
<td>1.01</td>
</tr>
<tr>
<td>17</td>
<td>Peptostreptococcus</td>
<td>0.66</td>
<td>Serratia</td>
<td>0.88</td>
<td>Veillonella</td>
<td>0.86</td>
</tr>
<tr>
<td>18</td>
<td>Parvimonas</td>
<td>0.54</td>
<td>Moraxella</td>
<td>0.82</td>
<td>Leptotrichia</td>
<td>0.69</td>
</tr>
<tr>
<td>19</td>
<td>Treponema</td>
<td>0.49</td>
<td>Streptophyta</td>
<td>0.61</td>
<td>Fusobacterium</td>
<td>0.55</td>
</tr>
<tr>
<td>20</td>
<td>Cardiobacterium</td>
<td>0.18</td>
<td>Peptostreptococcus</td>
<td>0.56</td>
<td>Sphingomonas</td>
<td>0.50</td>
</tr>
<tr>
<td>21</td>
<td>Atopobium</td>
<td>0.18</td>
<td>Mycoplasma</td>
<td>0.48</td>
<td>Elizabethkingia</td>
<td>0.49</td>
</tr>
<tr>
<td>22</td>
<td>Tannerella</td>
<td>0.06</td>
<td>Campylobacter</td>
<td>0.40</td>
<td>Opitutus</td>
<td>0.48</td>
</tr>
<tr>
<td>23</td>
<td>Mycoplasma</td>
<td>0.06</td>
<td>Treponema</td>
<td>0.29</td>
<td>Haemophilus</td>
<td>0.43</td>
</tr>
<tr>
<td>24</td>
<td>Peptococcus</td>
<td>0.04</td>
<td>Atoobium</td>
<td>0.19</td>
<td>Hydrogenophagia</td>
<td>0.41</td>
</tr>
<tr>
<td>25</td>
<td>Kingella</td>
<td>0.04</td>
<td>Acinetobacter</td>
<td>0.17</td>
<td>Aggregatibacter</td>
<td>0.40</td>
</tr>
<tr>
<td>26</td>
<td>Oribacterium</td>
<td>0.04</td>
<td>Bifidobacterium</td>
<td>0.14</td>
<td>Peptostreptococcus</td>
<td>0.33</td>
</tr>
<tr>
<td>27</td>
<td>Dialister</td>
<td>0.04</td>
<td>Klebsiella</td>
<td>0.14</td>
<td>Capnocytophaga</td>
<td>0.33</td>
</tr>
<tr>
<td>28</td>
<td>Megasphaera</td>
<td>0.03</td>
<td>Thermus</td>
<td>0.14</td>
<td>Symbiobacterium</td>
<td>0.28</td>
</tr>
<tr>
<td>29</td>
<td>Mogibacterium</td>
<td>0.03</td>
<td>Tannerella</td>
<td>0.14</td>
<td>Phyllobacterium</td>
<td>0.28</td>
</tr>
<tr>
<td>30</td>
<td>Corynebacterium</td>
<td>0.02</td>
<td>Lactobacillus</td>
<td>0.11</td>
<td>Porphyromonas</td>
<td>0.18</td>
</tr>
</tbody>
</table>
quantities had a tendency to be increased in CAP patients compared with healthy people. Gemella (N: 2.81%; C: 2.98%; \(P < 0.05\)) and Rothia (from 1.04–11.35%, \(P < 0.05\)) are such genera. Therefore, these two genera seem to be the important endogenous source of infection. Most other genera existing simultaneously in the three groups such as Prevotella are classified into the third category because of descending inclination in percentage from the healthy people to HAP patients. These genera tend to yield the territories to the invaders and be easily suppressed by antibiotics.

Contrastingly, the genera belonging to both CAP and HAP patients or to HAP patients only were more likely to be the exogenous invaders. In addition, in accordance with the percentage of each genera statistically compared, many genera had a higher percentage in HAP and significant differences between CAP and HAP patients. The genera Acinetobacter, Stenotrophomonas, Limnobacter, and Atopobium, which were dispersed throughout CAP data, occupy the highest place in the HAP table (Table 1, right column). This result suggests that, as exogenous agents, these genera play more important roles in HAP than in CAP, and they prefer the antibiotic-selected environment.

In addition, contributions in causing the infection made by Rothia and Moraxella should be emphasized in CAP. To our knowledge, Rothia is a natural part of normal flora in the oral cavity and upper respiratory tract; Moraxella is significant in HAP and minor in CAP for adults [24,25]. However, based on our results, Rothia is an important endogenous infective agent in CAP because of its significant increase from healthy people to CAP patients and decrease from CAP to HAP patients (N: 1.04%; C: 11.84%; H: 4.04%; \(P < 0.05\); Tables 1, 2, and 3). Meanwhile, the yellow cells of Rothia frequently appear in the CAP area in Fig. 2, which indicates that it is more prevalent in patients with CAP. Otherwise,
Moraxella shows no difference in the incidence between CAP and HAP patients ($P = 0.559$), but Moraxella is absent in healthy people. So, as an invader, Moraxella is equally worth being emphasized in both CAP and HAP patients as an infection agent.

The viewpoint that Rothia may play important roles in CAP is a new finding, which is proposed on the basis of microbiota quantitative analysis. In order to confirm it, further experiments to prove its virulence are surely needed. In our research work, non-small carcinoma cell (respiratory tract epithelial origin) and the bacteria named Rohia mucilaginosa were chosen in the study, and the interaction between cells and R. mucilaginosa could be used to evaluate the bacteria’s virulence on cell level. As the controls, non-virulent Escherichia coli and virulent Streptococcus pneumoniae acted as negative and positive controls. The results showed that R. mucilaginosa had stronger effects on interference with cell metabolic, apoptosis-inducing, and cytoskeletal damage, when compared with the effects made by E. coli and quite similar to that of S. pneumoniae, which indicates that towards the respiratory epithelial cells, R. mucilaginosa may possess similar ability to destroy the cells with S. pneumoniae. All these contents will be discussed in detail in another article separately.

Figure 3 shows the relationship between the individuals. The distance between CAP and HAP is larger than the distance between CAP and healthy people in the tree, since the CAP and the healthy people are on the same subtree, but
Figure 3  Cluster analysis of bacterial community of enrolled people  
(A) Neighbor-joining tree representing the relationship among the people enrolled. Branches of samples from healthy people are in green, branches of CAP patients are in blue, and HAPs are in red. The black branches are linking the subtree that contains the sample from different groups. (B) PCoA of the samples. Green triangles represent the sample from healthy people, blue squares represent the CAP, and red rounds represent the HAP.
HAP is on the other subtree [Fig. 3(A)]. Correspondingly, the scatter plots of CAP and healthy people are in adjacent quadrants, but the HAP is very far from them. Although there are three branches of CAP mixing in HAP subtree, their scatter plots are on the edge of the HAP region in [Fig. 3(B)] instead of being surrounded by HAP plots. These data indicate that even if very few CAP communities possess features similar to HAP, it is hard for them to represent a very typical HAP-like community.

Overall, from our result, the following conclusion could be drawn: first, although the whole bacterial community is considered to be an integrated pathogen according to the principle of ‘community as a pathogen’, the responsibility of each genus is surely distinguishable. Secondly, the extreme expansion of one certain genus has a strong inclination to result in the disorder. Thirdly, Rothia is a very important endogenous infective factor in CAP development. Finally, as the pyrosequencing-based method is able to clarify different individuals from different categories; it might become a promising approach in pulmonary infection diagnosis or evaluation in the future.

**Supplementary Data**

Supplementary data is available at ABBS online.

**Funding**

This work was supported by grants from the National Natural Science Foundation of China (30770111, 30900051, 30970125), the National Key Program for Infectious Diseases of China (2008ZX10004-002, 2008ZX10004-009, and 2009ZX10004-712), and the Program of Shanghai Subject Chief Scientist (09XD1402700).

**References**