

## Analysis of the Bacterial Communities Present in Lungs of Patients with Cystic Fibrosis from American and British Centers<sup>∇</sup>

Franziska A. Stressmann,<sup>1</sup> Geraint B. Rogers,<sup>1</sup> Erich R. Klem,<sup>2</sup> Andrew K. Lilley,<sup>1</sup>  
Scott H. Donaldson,<sup>2</sup> Thomas W. Daniels,<sup>3</sup> Mary P. Carroll,<sup>3</sup> Nilesh Patel,<sup>4</sup>  
Benjamin Forbes,<sup>4</sup> Richard C. Boucher,<sup>2</sup> Matthew C. Wolfgang,<sup>2,5</sup>  
and Kenneth D. Bruce<sup>1\*</sup>

Molecular Microbiology Research Laboratory, Pharmaceutical Science Research Division, King's College London, 150 Stamford Street, Franklin-Wilkins Building, London, SE1 9NH, United Kingdom<sup>1</sup>; Cystic Fibrosis/Pulmonary Research and Treatment Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599<sup>2</sup>; Cystic Fibrosis Unit, Southampton University Hospitals NHS Trust, Tremona Road, Southampton, SO16 6YD, United Kingdom<sup>3</sup>; Pharmaceutical Science Division, King's College London, 150 Stamford Street, Franklin-Wilkins Building, London, SE1 9NH, United Kingdom<sup>4</sup>; and Department of Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599<sup>5</sup>

Received 16 August 2010/Returned for modification 24 September 2010/Accepted 1 November 2010

**The aim of this study was to determine whether geographical differences impact the composition of bacterial communities present in the airways of cystic fibrosis (CF) patients attending CF centers in the United States or United Kingdom. Thirty-eight patients were matched on the basis of clinical parameters into 19 pairs comprised of one U.S. and one United Kingdom patient. Analysis was performed to determine what, if any, bacterial correlates could be identified. Two culture-independent strategies were used: terminal restriction fragment length polymorphism (T-RFLP) profiling and 16S rRNA clone sequencing. Overall, 73 different terminal restriction fragment lengths were detected, ranging from 2 to 10 for U.S. and 2 to 15 for United Kingdom patients. The statistical analysis of T-RFLP data indicated that patient pairing was successful and revealed substantial transatlantic similarities in the bacterial communities. A small number of bands was present in the vast majority of patients in both locations, indicating that these are species common to the CF lung. Clone sequence analysis also revealed that a number of species not traditionally associated with the CF lung were present in both sample groups. The species number per sample was similar, but differences in species presence were observed between sample groups. Cluster analysis revealed geographical differences in bacterial presence and relative species abundance. Overall, the U.S. samples showed tighter clustering with each other compared to that of United Kingdom samples, which may reflect the lower diversity detected in the U.S. sample group. The impact of cross-infection and biogeography is considered, and the implications for treating CF lung infections also are discussed.**

Cystic fibrosis (CF) is one of the most common genetic diseases in Europe and North America, with an incidence of around 1 in 2,500 (7). Due to the altered lung physiology, CF patients suffer chronic bacterial lung infections (25). The host immune response to these infections leads to an irreversible loss of lung function, with 85 to 95% of CF patients ultimately succumbing to respiratory failure (11, 20, 25, 34). Therefore, gaining a better understanding of the characteristics and dynamics of these infections is of the greatest importance.

A limited number of bacterial species have been traditionally considered key pathogens in CF lung disease progression, including *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Haemophilus influenzae*, *Burkholderia cepacia* complex, and *Stenotrophomonas maltophilia* (18, 20, 22, 25). Conventional diagnostic microbiology focuses on the detection of a limited

group of species, including those listed above (16). However, more recent studies have indicated that the range of bacterial species harbored in sputa from adult CF patients is much wider (28, 29, 31). These findings emerged through the application of culture-independent analytical strategies that detect bacterial species signatures in nucleic acids extracted directly from clinical samples. In particular, the combination of 16S ribosomal RNA (rRNA) gene terminal restriction fragment length polymorphism (T-RFLP) profiling and 16S rRNA clone sequence analysis has revealed the widespread presence of anaerobic bacterial species within the lower CF airways (15, 29).

Patient characteristics, such as cystic fibrosis transmembrane conductance regulator (CFTR) genotype, the presence of disease manifestations (e.g., CF-related diabetes), and patient demographics, vary according to the country of treatment (5, 11, 17). Although median survival for the United Kingdom and the United States is similar, with approximately 35 years for both (2, 9), patient and disease characteristics have been shown to vary, e.g., 40% of patients over 40 years in a United Kingdom CF center were pancreatic sufficient, compared to only 16% of patients in an equivalent U.S. center (17). There

\* Corresponding author. Mailing address: King's College London, FWB, Pharmacy, 150 Stamford St., London, SE1 9NH, United Kingdom. Phone: 207 848 4670. Fax: 207 848 4500. E-mail: kenneth.bruce@kcl.ac.uk.

<sup>∇</sup> Published ahead of print on 10 November 2010.

TABLE 1. Patient and sample information<sup>a</sup>

Patient	Age	Sex	Lung obstruction (FEV <sub>1</sub> % predicted)	Genotype I	Genotype II	Distance from home to hospital (km)	Days since last exacerbation	Last i.v. antibiotics given	Last inhaled antibiotics given	Last oral antibiotics given	BMI	Deceased since start of study	Mucolytics (pulmozyme/HS)
UK1	30	M	Mild	phe508del	Rare unidentified	20.5	>28	Flucloxacillin-tobramycin	Nil	Ciprofloxacin-azithromycin	29	N	Pulmozyme
US1	38	M	Normal	phe508del	phe508del	131	>28	NK	Colistin	Azithromycin	26.4	N	Pulmozyme
UK2	45	F	Mild	phe508del	Rare unidentified	25.6	14	Colistin-tobramycin	Colomycin	Azithromycin	18.5	N	Pulmozyme
US2	39	F	Mild	phe508del	INTRON 17 POL'	160	12	Cefepime-tobramycin	Nil	Azithromycin	29.4	N	3% HS
UK3	47	M	Severe	phe508del	Rare unidentified	30.7	>28	Ceftazidime-gentamicin	Colomycin	Azithromycin	20.7	N	Pulmozyme
US3	40	M	Severe	phe508del	G542X	44	>28	Tobramycin-aztreonam	Nil	Azithromycin	20.6	N	10% HS
UK4	30	F	Severe	phe508del	711 + 3A7G	6	13	Teicoplanin-gentamicin	Colistin-tobramycin	Azithromycin	25	N	Nil
US4	47	F	Severe	phe508del	Rare unidentified	82	EXAC	Meropenem-tobramycin	Colistin-tobramycin	Sulfamethoxazole-trimethoprim	20.9	N	Nil
UK5	22	F	Severe	phe508del	phe508del	51.3	>28	Meropenem-piperacillin-tazobactam	Colomycin	Azithromycin	19	N	Pulmozyme
US5	21	F	Severe	phe508del	phe508del	24	>28	Tobramycin-ceftazidime	Tobramycin	Azithromycin	17.9	N	Nil
UK6	40	F	Moderate	phe508del	phe508del	1.7	>28	Ceftazidime-gentamicin	Colomycin	Azithromycin	19.4	N	Nil
US6	46	F	Moderate	phe508del	phe508del	60	>28	Piperacillin-tazobactam-tobramycin	Tobramycin	Nil	27.8	N	Nil
UK7	21	F	Moderate	phe508del	phe508del	29.8	>28	Meropenem-colistin	Colomycin	Azithromycin	19	N	Pulmozyme
US7	21	F	Mild	phe508del	3849 + 10Kb->T	5	>28	NK	Tobramycin course just completed	Azithromycin	24.4	N	Nil
UK8	22	M	Severe	phe508del	phe508del	6	21	Meropenem-tobramycin	Colomycin	Erythromycin	17.9	N	Pulmozyme
US8	23	M	Severe	Rare unidentified	Rare unidentified	82	>28	Vancomycin-impipenem	Tobramycin	Azithromycin-linezolid-ciprofloxacin	16.1	N	Pulmozyme/7% HS
UK9	24	F	Normal	phe508del	G542X	28.6	>28	Teicoplanin	Colomycin	Azithromycin	23.4	N	Pulmozyme
US9	33	F	Mild	phe508del	DeltaF508	229	>28	Meropenem	Nil	Azithromycin	21.4	N	Pulmozyme/7% HS
UK10	20	M	Severe	phe508del	phe508del	4.5	>28	Ceftazidime-colistin	Tobramycin	Azithromycin	21	N	Nil
US10	26	M	Severe	phe508del	phe508del	54	>28	Imipenem-tobramycin-itraconazole-impipenem	Tobramycin	Azithromycin	17	Y	Pulmozyme/HS
UK11	35	M	Severe	phe508del	phe508del	28.9	>28	Meropenem-tobramycin	Colomycin	Azithromycin	nk	Y	Nil
US11	36	M	Severe	phe508del	phe508del	21	>28	Vancomycin-ceftazidime	Nil	Azithromycin	27.5	N	Nil
UK12	22	F	Moderate	phe508del	phe508del	47.2	>28	Nil	Colistin	Nil	nk	N	Nil
US12	18	F	Moderate	G542X	Rare unidentified	540	>28	Imipenem-piperacillin-tazobactam	Tobramycin	Ciprofloxacin	15.9	N	Pulmozyme/10% HS
UK13	24	F	Severe	phe508del	Rare unidentified	2	>28	Ceftazidime-gentamicin	Nil	Nil	14	N	NK
US13	28	F	Severe	phe508del	phe508del	103	>28	Tobramycin-cefepime	Colistin	Chronic azithromycin	15.7	N	Nil
UK14	27	M	Moderate	phe508del	G551D	47	>28	Ceftazidime-gentamicin	Tobramycin-colistin	Nil	22.7	Y	NK
US14	28	M	Normal	phe508del	2183delAAG	2	>28	Ceftazidime-tobramycin	Tobramycin	Azithromycin	23	N	Pulmozyme/10% HS
UK15	36	M	Severe	Rare unidentified	Rare unidentified	37	>28	Nil	Vancomycin	Clarithromycin	18.5	Y	NK
US15	28	M	Severe	Rare unidentified	Rare unidentified	308	>28	Piperacillin-tazobactam-levofloxacin	Nil	Nil	17.2	N	Nil
UK16	25	F	Moderate	phe508del	Rare unidentified	16	>28	Nil	Nil	Nil	nk	N	Nil
US16	23	F	Moderate	phe508del	phe508del	246	27	Tobramycin-piperacillin-tazobactam-vancomycin-meropenem-gentamicin	Tobramycin every other month	Tobramycin every other month	18.1	N	Nil
UK17	28	F	Severe	phe508del	Rare unidentified	30	>28	Vancomycin-meropenem-gentamicin	Nil	Nil	21	N	NK
US17	36	F	Severe	phe508del	phe508del	34	>28	Tobramycin-cefepime	Colistin due to hemoptysis	Azithromycin	19.2	N	Pulmozyme/3% HS
UK18	22	F	Severe	Rare unidentified	Rare unidentified	53	>28	Nil	Nil	Nil	19.7	Y	Nil

US18	29	F	Severe	phe508del	phe508del	137	>28	Meropenem- vancomycin- tobramycin	Nil	Azithromycin	26.9	N	3% HS with albuterol
UK19	20	F	Severe	phe508del	phe508del	43	EXAC	Meropenem-ceftazidime	Azithromycin	Nil	NK	N	NK
US19	26	F	Severe	phe508del	phe508del	36	EXAC	Tobramycin-levofloxacin	Alternating tobramycin	Ciprofloxacin	18	Y	Pulmozyme alternating with tobramycin

<sup>a</sup> Patients are numbered in pairs. Lung obstruction categories: normal, > 80%; mild, 60 to 80%; moderate, 40 to 59%; severe, <40%. NK, not known; HS, hypertonic saline.

are many possible reasons why this may be the case, including socioeconomic factors, neonatal screening practices, and differences in treatment practices (14). Recommendations for the latter differ to some extent between the United Kingdom and United States (6, 10, 39). In addition, how these recommendations are implemented depends on individual treatment centers and can vary considerably.

However, as CF lung disease involves infection by bacteria of an apparently wide range of origins, it may be important to consider the environment as a source of these species and thus as another disease-modifying factor. The concept of early microbiological studies that “everything is everywhere, but the environment selects” infers that despite the purging effect of selection in different habitats, microbial dispersal is widespread, and as such it offers little opportunity for the spatial differentiation of bacterial species (8, 26).

Evidence is emerging that the composition of bacterial communities in different habitats can vary from region to region (4, 12, 13, 19). The degree to which endemism (localized populations) is important in terms of its impact on species presence in the CF lung is not clear. This could, however, be of clinical importance, given that many of the species regarded as pathogens are present in a range of natural environments, e.g., *P. aeruginosa*, *B. cepacia* complex, and *S. maltophilia*. For example, *P. aeruginosa* has been shown to occur worldwide, but clones with increased virulence and enhanced antibiotic resistance occasionally arise, leading to localized epidemic spread in the environment and CF centers, such as the Liverpool epidemic strain (27, 33).

In this study, bacterial community profiles, generated from adult CF patients in two distinct geographical regions (the United States and the United Kingdom), were compared to investigate geographical differences in the bacterial CF airway community. Culture-independent approaches of terminal restriction fragment length polymorphism (T-RFLP) profiling and clone sequence analysis were used to analyze the bacterial content of sputum samples. Patients were grouped in pairs, with one United Kingdom and one U.S. patient matched on parameters considered clinically relevant. The analysis of these data showed that, both in terms of species presence/absence as well as the relative prevalence of species in the lung, no clear transatlantic divide could be defined between bacterial communities. More-subtle differences in community composition and structure were, however, observed.

**MATERIALS AND METHODS**

**Clinical samples.** Spontaneously expectorated sputum samples were collected from 19 adult CF patients attending the Adult Cystic Fibrosis Clinic at Southampton General Hospital, United Kingdom, and 19 patients attending the University of North Carolina Cystic Fibrosis/Pulmonary Research and Treatment Center, Chapel Hill, NC, under full ethical approval. After collection, samples were stored at -80°C prior to nucleic acid extraction. All nucleic acid extractions, as well as T-RFLP analysis, were carried out by one person in the laboratory in the United Kingdom. Clone sequence analysis was performed on these same nucleic acid extracts by one person in the United States. Patient age, sex, lung disease severity, CFTR genotype, body mass index (BMI), and antibiotic therapy details relating to these samples are shown in Table 1. Patients recruited to this study represented a random cross-section of CF patients. All were between 18 and 47 years old, with 12 females and 7 males per group. Enrolled patients from the United Kingdom and United States were paired according to age, gender, and lung obstruction category (forced expiratory volume in 1 second [FEV<sub>1</sub>%] predicted), as assessed clinically (Table 1). An age

difference of less than 10 years together with a difference of 15% or less in FEV<sub>1</sub>% predicted but in the same lung obstruction category (i.e., normal, mild, moderate, severe) was considered acceptable. Mean age difference was 5.4 years (standard deviation [SD], 4.1) and mean difference in FEV<sub>1</sub>% predicted was 10.1% (SD, 7.7%). Exceptions where only one parameter could be matched acceptably were UK/US1, UK/US3, UK/US4, and UK/US9, with an age difference and difference in FEV<sub>1</sub>% predicted of 8 years and 30%, 7 years and 18%, 17 years, and 2% and 9 years and 20%, respectively. Additional criteria, such as distance to hospital, were recorded. Patients were clinically stable at the time of sampling (at least 21 days had passed before or after a treatment for pulmonary exacerbation), with eight exceptions (U.S. patients 4 and 19 and United Kingdom patient 19, who were receiving intravenous [i.v.] antibiotics for pulmonary exacerbation; U.S. patient 2, United Kingdom patient 2, and United Kingdom patient 4, who had finished i.v. antibiotics 12, 14, and 13 days previously, respectively; and U.S. patients 7 and 17, who were clinically stable but were receiving antibiotic therapy for reasons other than pulmonary exacerbation). In all cases, clinical status was given priority over antibiotic therapy for patient matching.

**DNA extraction.** Prior to DNA extraction, sputum samples were washed three times in phosphate-buffered saline (PBS; Fisher Scientific, Loughborough, United Kingdom) to remove adherent saliva. DNA extraction from clinical samples was adapted from a previously described procedure (29). Briefly, 200  $\mu$ l of each sputum sample was resuspended in 800  $\mu$ l of 200 mM PBS (pH 8.0) and 300  $\mu$ l guanidium thiocyanate-EDTA-Sarkosyl. After the addition of 0.2 g of 0.18-mm-diameter glass beads (B. Braun Biotech International GmbH, Melsungen, Germany), samples were homogenized for 60 s at 30 Hz in a Mixer Mill 300 (Qiagen, Crawley, United Kingdom). Samples were heated at 70°C for 20 min and placed on ice for 20 min, and beads and cell debris were removed by centrifugation (13,000  $\times$  g for 5 min) at room temperature. Supernatants were transferred to fresh microcentrifuge tubes, followed by the addition of NaCl (to a final concentration of 0.5 mM) and polyethylene glycol (to a final concentration of 15%). Nucleic acids were precipitated at 4°C for 1 h.

DNA was pelleted at room temperature by centrifugation at 13,000  $\times$  g for 10 min and resuspended in 300  $\mu$ l nuclease-free water. Supernatants were transferred to fresh microcentrifuge tubes, and 300  $\mu$ l Tris-acetate-EDTA (TAE)-saturated phenol (pH 8.0) (Sigma-Aldrich, Gillingham, United Kingdom) was added. After the mixture was vortexed vigorously, the phases were separated by centrifugation at 13,000  $\times$  g for 5 min. The upper phase was transferred to a fresh microcentrifuge tube, 300  $\mu$ l phenol-chloroform-isoamylalcohol (25:24:1 ratio) (Sigma-Aldrich) was added, and samples were vortexed vigorously and centrifuged at 13,000  $\times$  g for 10 min. Supernatants were transferred to fresh microcentrifuge tubes, and DNA was precipitated at -20°C for 1 h after the addition of an equal volume of isopropanol (Sigma-Aldrich) and 0.1 volume of 10 M ammonium acetate. The DNA was pelleted by centrifugation at 13,000  $\times$  g for 10 min and washed three times in 200  $\mu$ l 75% ethanol. Pellets were briefly air dried, resuspended in 50  $\mu$ l of nuclease-free water, and stored at -20°C.

**DNA verification and quantification.** Extracted genomic DNA and PCR products were verified by TAE-agarose gel electrophoresis stained with GelRed (Biotium, Hayward, CA) and visualized on a UV transilluminator (Herolab, Wiesloch, Germany). Images were captured by using a Herolab image analyzer with E.A.S.Y. Stop win 32 software (Herolab). DNA and PCR products were quantified by spectrophotometry by applying 1.5  $\mu$ l directly to a NanoDrop ND-1000 spectrophotometer (LabTech International, Ringmer, United Kingdom).

**PCR and T-RFLP profiling.** The universal oligonucleotide primers for the amplification of a region of the 16S rRNA gene specific for the domain bacteria 926r (5'-CCG TCA ATT CAT TTG AGT TT-3') and 8f700IR (5'-AGA GTT TGA TCC TGG CTC AG-3') were used as described previously (24). Primer 926r was unlabeled, and primer 8f700 was labeled with the dye IRD700 at the 5' end. Both primers were synthesized by TAGN, Newcastle, United Kingdom. The constituents of the PCR mixture per reaction were the following: 25  $\mu$ l of the Sigma readymix REDTaq (Sigma-Aldrich), each primer at a final concentration of 0.2 mM, 50 ng of DNA template, made up to a final volume of 50  $\mu$ l with nuclease-free water.

Cycling conditions comprised an initial denaturation at 94°C for 2 min, followed by 32 cycles of denaturation at 94°C for 1 min, annealing at 56°C for 1 min, and extension at 72°C for 2 min, with a final extension step at 72°C for 10 min. Amplifications were carried out in a Gene Amp PCR system 9700 (Applied Biosystems, United Kingdom). PCR products were verified on TAE-agarose gels as described above and stored at -20°C for T-RFLP analysis.

Approximately 20 ng of each PCR product was digested to completion with 1 U of the restriction endonuclease CfoI (Sigma-Aldrich) for 5 h at 37°C, in accordance with the manufacturer's instructions. The restriction enzyme was inactivated by heating to 90°C for 20 min. Approximately 10 ng of digested PCR

product was denatured at 95°C for 1 min and separated by length using a 25-cm SequagelXR denaturing polyacrylamide gel (National Diagnostics, Hesse, United Kingdom), with the addition of 8.3 M urea and formamide (to a final concentration of 10%). Electrophoresis was performed at 55°C and 1,200 V on an IR2 automated DNA sequencer (LI-COR Biosciences, Lincoln, NE).

**T-RFLP profile analysis.** T-RFLP gel images were analyzed using Phoretix one-dimensional advanced software (version 5.10; Nonlinear Dynamics, Newcastle upon Tyne, United Kingdom). The lengths of the bands detected on the gel were determined by comparison to the positions of the size marker microSTEP 15a (700 nm) (Microzone, Lewes, United Kingdom). Additionally, Phoretix software was used to determine the band volume (the product of the area over which the band was detected and the intensity of signal recorded over that area). The band volume was expressed as a percentage of the total volume of bands resolved in one particular electrophoretic profile. The resolution of T-RFLP bands was over the region of 50 to 950 bases. T-RF bands shorter than 50 bases were not included in the analysis, as this region is susceptible to high levels of background signal. The threshold of band detection used in this study was 0.1% of total signal in one profile in the specified region.

**In silico sequence analysis and band identification.** Band identification was performed as described previously by Rogers and coworkers (28, 29). Published bacterial 16S rRNA gene sequences were retrieved from GenBank (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide>). The software MapSort (Wisconsin Package, version 10.3; Accelrys, Cambridge, United Kingdom) was used to predict T-RFLP band lengths (in bases) from the 5' end of primer 8f700IR to the first site of cleavage for CfoI in each recovered 16S rRNA gene sequence.

**Generation of 16S rRNA clone libraries and sequence analysis.** 16S rRNA gene fragments were generated by PCR using primers 8f700 and 926r and the high-fidelity proofreading enzyme PFU-Ultra (Stratagene, La Jolla, CA) as described above. The pool of 16S rRNA gene fragments was cloned into the pCR-Blunt II-TOPO plasmid vector (Invitrogen, Carlsbad, CA) according to manufacturer's instructions. Plasmids then were used to transform competent *Escherichia coli* DH5 $\alpha$  by heat shock at 42°C and plated on selective LB agar plates with 50  $\mu$ g/ml kanamycin. Individual colonies were used to inoculate 2 ml of LB broth containing 50  $\mu$ g/ml kanamycin in 96-deep-well blocks and grown for 20 h with vigorous shaking at 37°C. One hundred microliters of culture was saved as a glycerol stock, and the remaining 1.9 ml of culture was centrifuged at 500  $\times$  g for 10 min. Supernatants were removed and plasmid DNA isolated from the bacterial pellet with the Wizard SV 96 plasmid purification system (Promega Corporation, Madison, Wisconsin). The quality of resulting cloned 16S rRNA gene fragments containing plasmid DNA was assessed by TAE-agarose gel electrophoresis. Aliquots of purified plasmid were sequenced with M13 vector primer (MWG Biotech, High Point, NC). A total of 871 clones derived from samples taken from five patients from each geographical location were sequenced.

**Phylogenetic tree construction.** For each patient, sequences were vector trimmed with Sequencher v4.8 (Gene Codes Corporation, Ann Arbor, Michigan) and aligned at 100% identity and clustered. The longest sequence of each cluster was conserved and shorter redundant sequences removed. Sequences were aligned with the Ribosomal Database Project (RDP; Michigan State University, Michigan) Pyrosequencing Alignment and Clustering Program (<http://pyro.cme.msu.edu/>) and clustered into bins of 98% similarity with the "farthest nearest neighbor" complete linkage clustering algorithm. To reduce bin complexity, bins containing three or more sequences were aligned using the Greengenes alignment tool ([http://greengenes.lbl.gov/cgi-bin/nph-NAST\\_align.cgi](http://greengenes.lbl.gov/cgi-bin/nph-NAST_align.cgi)) using a batch size of 100, a minimum length of 100, and a minimum percent identity of 75%. The bins of three or more aligned sequences were transferred to ClustalX v2.0.12 (Conway Institute, UCD, Dublin, Ireland), and a multiple alignment was carried out with an output in the form of a phylogenetic tree. For each bin, the longest most-similar branch was chosen as the representative sequence, labeled to represent the number of sequences it contained (at 100 and 98% sequence identity), and the remaining sequences removed. Overall, this reduced sequence complexity from 871 to 136. To construct the phylogenetic tree for all patients, these 136 binned sequences were put into the Greengenes alignment tool and ClustalX as described above, together with the sequences of type strains that all sequences were most similar to, including CF pathogens, and the phylogenetic tree was constructed with the same parameters as those described above. Clusters were named according to the type strains that the sequences clustered most closely with.

**Statistical analysis.** Correlations, *t* tests, and one-way analysis of variance with tests of normality and homogeneity of variance, as well as Bonferroni and Tukey post hoc tests, were performed using SPSS software version 15.0.1 (Chicago, IL).

Raup and Crick as well as Morisita similarity coefficients were used to

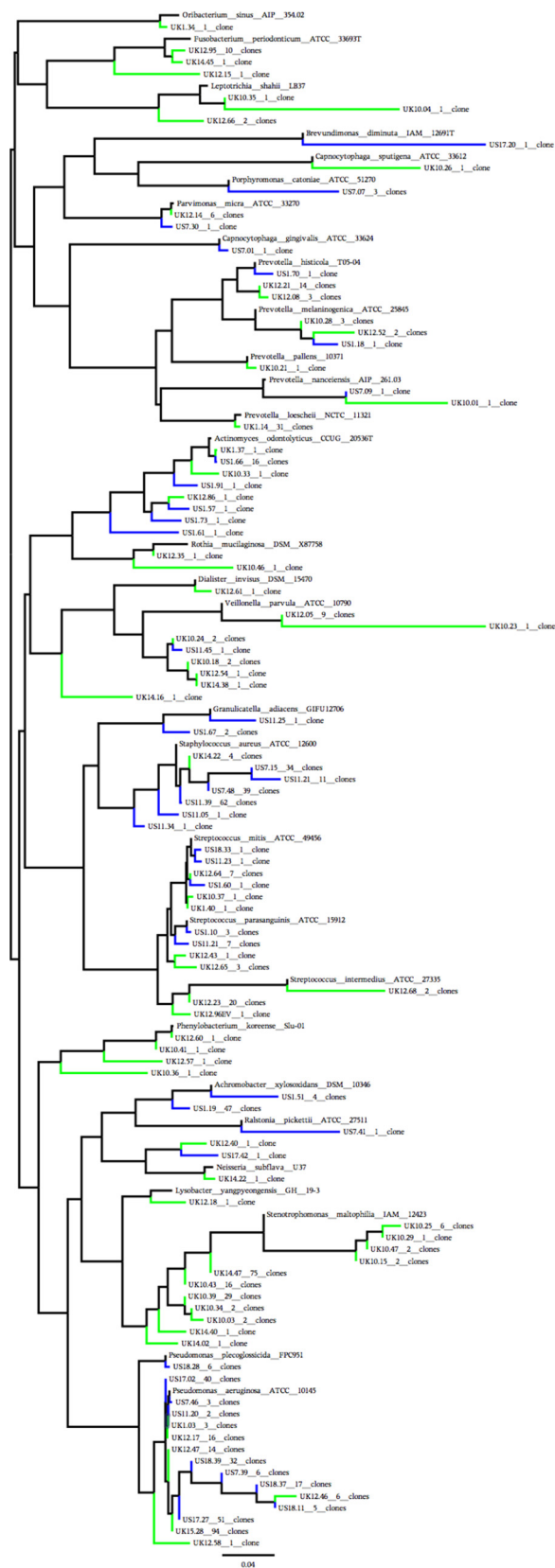


FIG. 1. Phylogenetic tree of clones from United Kingdom and U.S. patients. Clone sequences are clustered with the bacterial type strains to which they showed highest sequence similarity (minimum of 98%

evaluate similarities in community composition. Similarity coefficients were calculated and dendrograms plotted using PAST software version 1.74 (<http://folk.uio.no/ohammer/past/>; University of Oslo, Norway). Raup and Crick, a probability-based similarity index ( $S_{RC}$ ), was calculated from binary matrices of OTU presences and absences. The probabilities were calculated using 2,000 Monte Carlo simulations to compare the number of bands shared by two samples and the number of bands predicted if two samples were randomly selected from an amalgamation of all the samples studied.

**Nucleotide sequence accession numbers.** Sequences determined in the course of this work are available through the EMBL nucleotide sequence database (<http://www.ebi.ac.uk/embl/>) under accession numbers FN825913 to FN826783.

## RESULTS

**Community composition: species absence/presence.** In the 38 sputum samples analyzed from both United Kingdom and U.S. patients by T-RFLP, a total of 173 bands representing 73 different T-RF lengths were detected. The number of individual bands resolved in each sample ranged from 2 to 10 for the U.S. patients and 2 to 15 for the United Kingdom patients. The mean number of T-RF bands per patient of 4.6 (standard deviation [SD], 3.6) in the United Kingdom sample set was similar to that of the U.S. sample set at 4.5 (SD, 2.3) and not significantly different [ $P = 0.87$ ;  $t_{(df\ 31)} = 0.159$ ]. A modest significant correlation was observed between the number of bands detected for the paired U.S. and United Kingdom patient sets [ $P = 0.03$ ;  $r_{(df\ 17)} = 0.496$ ].

Of the 173 T-RF bands detected, 32.4% were “singletons,” defined here as bands that occurred in only one sample. Singletons were detected in nine United Kingdom and eight U.S. samples. The mean number of singletons in these patients was 1.7 (SD, 3.0) and 1.2 (SD, 1.8) for United Kingdom and U.S. patients, respectively. This difference was not significant [ $P = 0.51$ ;  $t_{(df\ 30)} = 0.66$ ].

T-RF band lengths that occurred more than once were analyzed in detail. All band lengths that occurred more than four times in the overall sample set were present in at least one United Kingdom and U.S. patient. Certain band lengths were commonly found in both sample sets. Ten band lengths accounted for 57% of all bands and were found in at least one United Kingdom and U.S. patient. Species consistent with these band lengths included *Pseudomonas aeruginosa*, *Staphylococcus aureus*, and *Stenotrophomonas maltophilia*, certain anaerobic species of the *Prevotella* genus, as well as two band lengths for which no species prediction could be made.

Clone sequence analysis was carried out for five patients from the United Kingdom (UK1, UK10, UK12, UK14, and UK15) and the United States (US1, US7, US11, US17, and US18). Samples for this analysis were chosen on the basis of T-RFLP results to include two patients from each group with high diversity (at least 10 different T-RF band lengths), one patient with low diversity (5 or fewer bands), and two patients for which the band at 155 bases, consistent with *P. aeruginosa*, was not detected. Clone sequence analysis showed that clones

similarity for cluster membership). Clones were grouped into unique sequences; each labeled sequence indicates the patient it was isolated from and the number of redundant clone sequences it comprises. Clones isolated from United Kingdom patients are shown in green, clones isolated from U.S. patients are shown in blue.

TABLE 2. Type strain cluster details from Fig. 1<sup>a</sup>

Type strain cluster	No. of unique sequences from United Kingdom present in type strain cluster	No. of United Kingdom patients in type strain cluster	No. of unique sequences from U.S. present in type strain cluster	No. of U.S. patients in type strain cluster	Total no. of unique sequences in type strain cluster
<i>Achromobacter xylosoxidans</i>			2	1	2
<b><i>Actinomyces odontolyticus</i></b>	<b>3</b>	<b>3</b>	<b>5</b>	<b>1</b>	<b>8</b>
<i>Brevundimonas diminuta</i>			1	1	1
<i>Capnocytophaga gingivalis</i>			1	1	1
<i>Capnocytophaga sputigena</i>	1	1			1
<i>Dialister invisus</i>	2	2			2
<i>Fusobacterium periodonticum</i>	3	2			3
<i>Granulicatella adiacens</i>			2	2	2
<i>Leptotrichia shahii</i>	3	2			3
<i>Lysobacter yangyeongensis</i>	1	1			1
<b><i>Neisseria subflava</i></b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>1</b>	<b>3</b>
<i>Oribacterium sinus</i>	1	1			1
<b><i>Parvimonas micra</i></b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>2</b>
<i>Phenylobacterium koreense</i>	4	2			4
<i>Porphyromonas catoniae</i>			1	1	1
<b><i>Prevotella histicola</i></b>	<b>2</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>3</b>
<i>Prevotella loescheii</i>	1	1			1
<b><i>Prevotella melaninogenica</i></b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>1</b>	<b>3</b>
<b><i>Prevotella nanceiensis</i></b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>2</b>
<i>Prevotella pallens</i>	1	1			1
<b><i>Pseudomonas aeruginosa</i></b>	<b>6</b>	<b>3</b>	<b>7</b>	<b>3</b>	<b>13</b>
<i>Pseudomonas plecoglossicida</i>			1	1	1
<i>Ralstonia pickettii</i>			1	1	1
<i>Rothia mucilaginosa</i>	2	2			2
<b><i>Staphylococcus aureus</i></b>	<b>1</b>	<b>1</b>	<b>6</b>	<b>2</b>	<b>7</b>
<i>Stenotrophomonas maltophilia</i>	11	3			11
<i>Streptococcus intermedius</i>	3	1			3
<b><i>Streptococcus mitis</i></b>	<b>3</b>	<b>3</b>	<b>3</b>	<b>3</b>	<b>6</b>
<b><i>Streptococcus parasanguinis</i></b>	<b>2</b>	<b>1</b>	<b>2</b>	<b>2</b>	<b>4</b>
<b><i>Veillonella parvula</i></b>	<b>6</b>	<b>3</b>	<b>1</b>	<b>1</b>	<b>7</b>

<sup>a</sup> The number of unique sequences for each cluster, as well as the number of patients from which unique sequences were identified for United Kingdom and U.S. samples, are shown. Clusters that contained sequences from both United Kingdom (UK1, UK10, UK12, UK14, and UK15) and U.S. (US1, US7, US11, US17, and US18) patients are shown in boldface.

clustered with 30 different typed strains. These typed strains were used to name the clusters of the phylogenetic tree constructed (Fig. 1, Table 2) and here will be referred to as species. A wide range of species was identified from 22 different genera. These included one species, *Phenylobacterium koreense*, which had not previously been associated with the CF lung, as well as six species (20%) of obligate anaerobes. Overall, 9 genera of 22 were detected in both patient groups (Table 2, Fig. 1). Individual species within eight bacterial genera detected, namely, *Pseudomonas*, *Streptococcus*, *Actinomyces*, *Staphylococcus*, *Parvimonas*, and *Neisseria*, as well as two genera of the obligate anaerobes *Prevotella* and *Veillonella*, were present in at least one individual from both the United Kingdom and United States as analyzed by this approach (Table 2, Fig. 1). In the United Kingdom patient group, the most abundantly identified clones were *P. aeruginosa* and *S. maltophilia* (31% each), and in the U.S. patient group they were *P. aeruginosa* (30%) and *S. aureus* (33%). Overall, 12 (40%) species were found for the United Kingdom patient group alone, 7 (23%) species were found for the U.S. patient group alone, and 11 (37%) species occurred in both patient groups (Fig. 1, Table 2). Type strain clusters that comprised four or more unique sequences all contained sequences from both United Kingdom and U.S. patients, with the exceptions of *Phenylobac-*

*terium koreense* and *S. maltophilia*, which both were detected in the United Kingdom patient group only (Table 2, Fig. 1).

**Similarities in community composition.** The similarity of the bacterial content of samples was statistically assessed on the basis of the T-RFLP data. The patterns in shared species composition between the two sample sets were assessed through the generation of a Venn diagram (Fig. 2). From the data set as a whole, the four most commonly occurring band lengths, consistent with bands produced by *P. aeruginosa* (155 bases), *Prevotella* spp. (103 bases), *Achromobacter/Bordetella/Pseudomonas* spp. (565 bases), and an unassigned band at 580 bases, were examined in detail here. The two largest sets (band lengths of 155 and 565 bases) contained high numbers of both U.S. and United Kingdom patients. In total, this represented 79% of both patient groups. Overall, all four of the most commonly detected band lengths were found in at least one patient from each geographical group. In two of the samples studied (UK14 and US16), none of the four most commonly occurring band lengths were detected, so they are presented as outliers in Fig. 2.

To assess the significance of similarities and differences in species composition between United Kingdom and U.S. samples, Raup and Crick similarity indices ( $S_{RC}$ ) were calculated for the 38 T-RFLP profiles from United Kingdom and U.S.

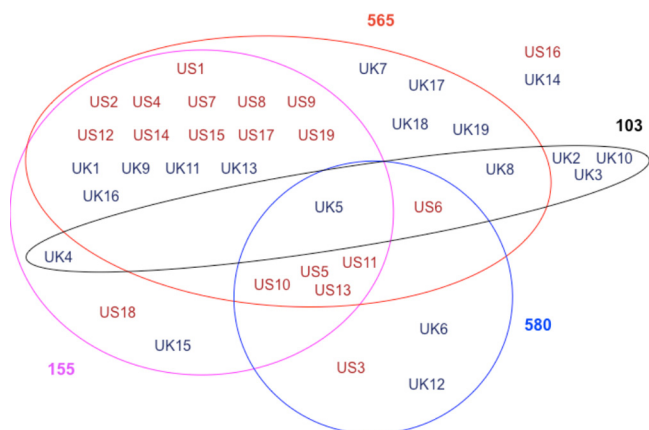


FIG. 2. Venn diagram grouping patients according to the presence of one or more of the four most commonly occurring T-RF band lengths in the sample set. Circles correspond to T-RF band lengths; patients within one or more circles shared these particular T-RF band lengths. T-RF band lengths corresponded to the following organisms: 103 bases, *Prevotella* spp.; 155 bases, *P. aeruginosa*; 565 bases, *Pseudomonas/Bordetella/Achromobacter* spp.; 580 bases, not assigned.

patients, and a cluster diagram was constructed using unweighted pair-group averages (UPGMA) (Fig. 3). Values above 0.95 indicate pairs of samples that were more similar than expected by chance, and thus indicated significant clustering. Using this 0.95 threshold, Fig. 3 is comprised of three clusters; cluster 1 of four U.S. patients (US5, US10, US11, and US13), cluster 2 of five United Kingdom patients (UK7, UK8, UK17, UK18, and UK19), and a substantial mixed cluster (3) of 10 U.S. patients and six United Kingdom patients (US1, US2, US4, US7, US8, US12, US14, US15, US17, US19, UK1, UK5, UK9, UK11, UK13, and UK16).

The degree of similarity of the U.S. and United Kingdom samples was further assessed using the Raup and Crick probability-based similarity index calculated on species presences and absences. This statistically resolved whether the samples were significantly similar ( $S_{RC} \geq 0.95$ ), significantly dissimilar ( $0.05 \geq S_{RC}$ ), or neither ( $0.05 < S_{RC} < 0.95$ ). Table 3 shows the numbers and proportions of significant and nonsignificant  $S_{RC}$  values for a series of pairwise comparisons of  $S_{RC}$  values for United Kingdom with United Kingdom patients, U.S. with

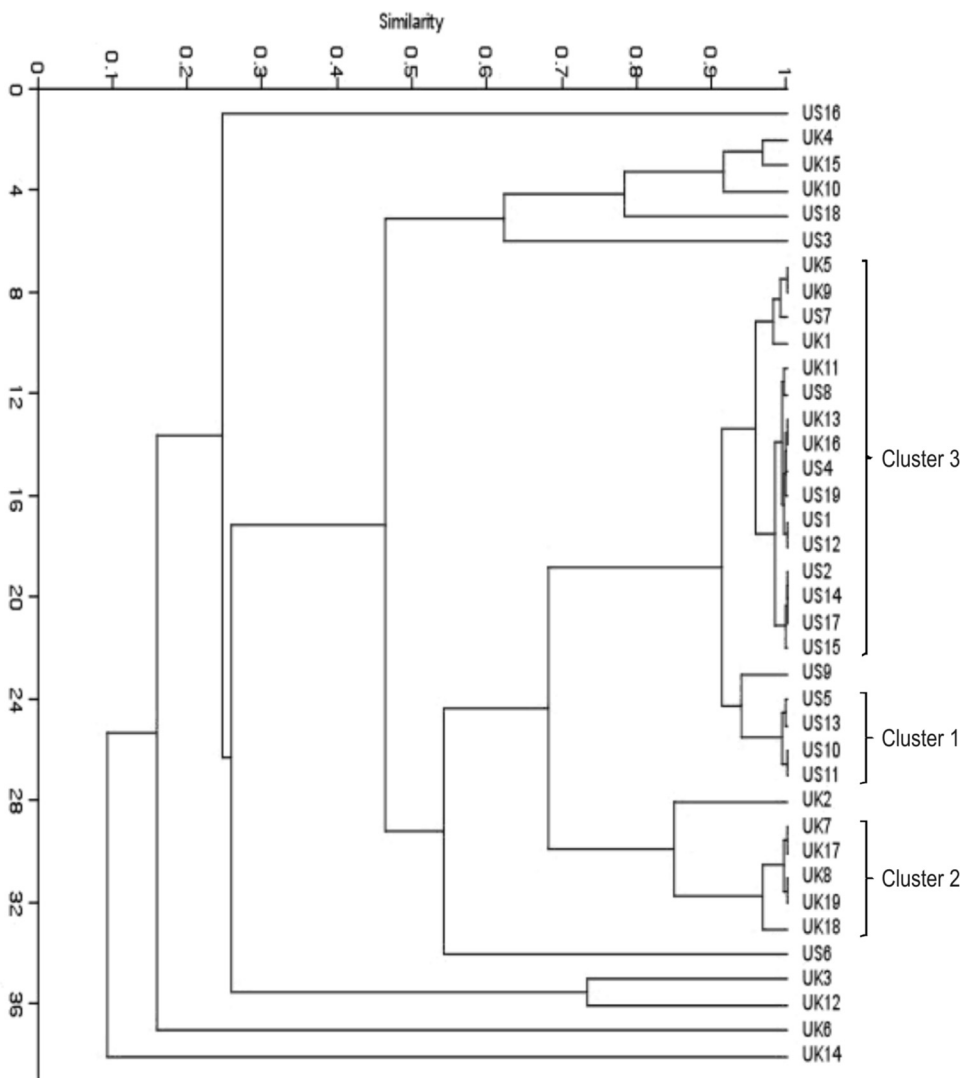


FIG. 3. Cluster diagram of the Raup and Crick similarity coefficient assessment of United Kingdom and U.S. patient T-RFLP data.

TABLE 3. Numbers and percentages of significant and nonsignificant Raup and Crick similarity indices<sup>a</sup>

Sample patients	No. (%) with $S_{RC}$ value:			Total no. (%) of $S_{RC}$ value comparisons
	$\geq 0.95$	$\leq 0.05$	Not significant	
U.K.-U.K.	25 (14.6)	7 (4.1)	139 (81.3)	171 (100)
U.S.-U.S.	70 (40.9)	2 (1.2)	99 (57.9)	171 (100)
U.K.-U.S.	163 (23.2)	28 (4)	512 (72.8)	703 (100)

<sup>a</sup> Indices are shown for pairwise comparisons for United Kingdom with United Kingdom patients (U.K.-U.K.), U.S. with U.S. patients (U.S.-U.S.), and United Kingdom with U.S. patients (U.K.-U.S.) from T-RFLP profiling. Values equal to or above 0.95 indicate pairs of samples that are more similar than expected by chance, and values equal to or less than 0.05 indicate pairs of samples that are less similar than expected by chance.

U.S. patients, and the matched United Kingdom and U.S. patients. Here, 23% of United Kingdom-U.S. pairs showed a significant level of similarity. United Kingdom-U.S. pairs were more similar than United Kingdom-United Kingdom pairs (15%) but less similar than U.S.-U.S. pairs (41%). A significant level of dissimilarity was detected for 4% of United Kingdom-U.S. pairs. For United Kingdom patients, 4% were more dissimilar from each other than expected by chance, and only 1%

of U.S. patients were significantly dissimilar from each other. Further, this meant that many United Kingdom samples were more similar to U.S. samples than to each other.

**Community structure.** In addition to species presence, the similarity of the abundance of these species in samples was investigated using T-RFLP band volume data. Morisita similarity coefficients ( $I_M$ ) were calculated in a pairwise manner for the relative volumes of bands of all 38 T-RFLP profiles, and a cluster dendrogram was plotted using UPGMA (Fig. 4).

Two distinct clusters of exclusively United Kingdom patients were formed with four and six patients, respectively: cluster 1 (UK7, UK8, UK17, and UK19) and cluster 2 (UK4, UK5, UK11, UK13, UK15, and UK16). A third cluster of 11 U.S. and one United Kingdom patient also was observed (US1, US4, US7, US8, US11, US12, US13, US14, US15, US17, US19, and UK9). The average  $I_M$  value for United Kingdom-United Kingdom patient comparisons was 0.27 (SD, 0.39;  $n = 171$ ), for U.S.-U.S. comparisons was 0.59 (SD, 0.38;  $n = 171$ ), and for United Kingdom-U.S. comparisons was 0.35 (SD, 0.38;  $n = 361$ ). These three average Morisita index values all were significantly different at [ $P < 0.0001$ ,  $F_{(df\ 2, 700)} = 33.21$ , and post hoc tests]. Again the U.S. samples showed a stronger association with each other than the United Kingdom samples showed

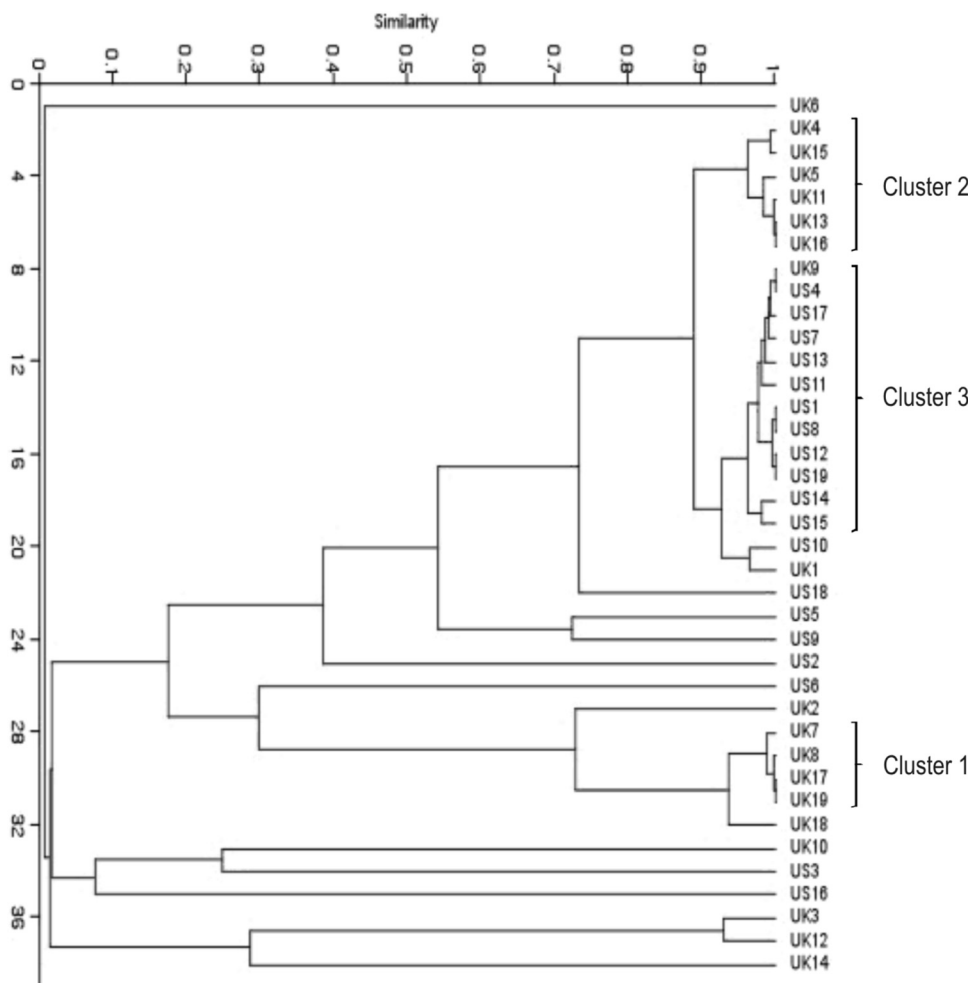


FIG. 4. Cluster diagram for the Morisita similarity coefficient assessment of United Kingdom and U.S. patient T-RFLP data.



with each other, and the United Kingdom samples were significantly more associated with the U.S. samples than with one another.

In summary, the comparison of the data presented in Fig. 3 and 4 showed that some degree of overlap was identified in terms of cluster membership in relation to the two different strategies of data analysis used. Figure 3 identified three clusters, one cluster of 4 patients (US5, US10, US11, and US13), one of 5 patients (UK7, UK8, UK17, UK18, and UK19), and another of 16 patients (US1, US2, US4, US7, US8, US12, US14, US15, US17, US19, UK1, UK5, UK9, UK11, UK13, and UK16). Figure 4 identified another three clusters, one of 4 patients (UK7, UK8, UK17, and UK19), one of 6 patients (UK4, UK5, UK11, UK13, UK15, and UK16), and another of 12 patients (US1, US4, US7, US8, US11, US12, US13, US14, US15, US17, US19, and UK9).

**Comparison of T-RFLP and clone sequence data.** For 90% of species identified by clone sequence analysis, a T-RF band length for the corresponding cut site was detected. For *Brevundimonas diminuta* and *Parvimonas micra*, no corresponding T-RF band length could be detected, and for *Oribacterium sinus*, the first CfoI cut site lay outside the T-RFLP detection range. For 26% of T-RF bands a species with a corresponding cut site was detected by clone sequence analysis. The detection of species by clone sequencing that corresponded to T-RF band lengths in individual patient's profiles ranged from 13% (United Kingdom patient 1) to 67% (United Kingdom patients 10 and 12).

## DISCUSSION

The aim of this study was to compare the bacterial communities present in sputum collected from adult CF patients attending CF centers in the United States and United Kingdom. Patients in the two countries were paired prior to microbiological assessment according to clinical parameters. Two culture-independent, molecular strategies were used to analyze sample pairs, T-RFLP profiling and 16S rRNA clone sequence analysis. Both relied on the PCR amplification of the 16S rRNA gene from DNA extracted directly from bacteria in the sputum samples collected. These approaches have been used previously to characterize the communities of bacteria in the adult CF lung (3, 15, 28–32, 35, 36). To the best of our knowledge, this is the first application of such techniques to assess the impact of geographical location on bacterial community composition in CF respiratory samples.

To normalize for variation other than geographical grouping, patients were matched into pairs to align major clinical parameters considered important by the treating clinicians (age, sex, lung obstruction, BMI, and CFTR genotype). An attempt to match for other parameters, such as distance to clinic and treatment, also was made. However, due to differences in the size of the catchment area and some aspects of the treatment regimes, lower priority was given to these parameters. While we recognize that no system of matching patients is ideal, this approach allowed 19 pairings to be defined based on these criteria.

The T-RFLP profiling of the generated 16S rRNA PCR products was used to assess the bacterial community. A total of 73 distinct T-RF band lengths were generated for the 38 pa-

tients studied, which each were regarded as an individual bacterial species. The paired U.S. and United Kingdom patients showed similar numbers of bacterial species within their samples, suggesting that one or more of the clinical parameters used to assign patients to their pair were relevant. No marked difference was identified in the mean number of species per sample for the United Kingdom or U.S. patient group. While no comparable studies could be identified in the literature, the mean number of species obtained here was lower than that in earlier studies of stable adult CF patients (29, 30, 36), with this difference accounted for by the selection of a higher detection threshold. Statistical assessment showed a modest significant correlation in species number in relation to the patient pairs. As such, the number of species per sample did not provide evidence for a geographical difference in samples from the United Kingdom or U.S.

No difference was identified in the number of singletons from patients in the United Kingdom or U.S. Overall, approximately one third of all species detected occurred as singletons. Taken together, these findings imply that the lung community has a strong component that is patient specific regardless of patient geographical location. Other evidence reinforcing the highly different nature of bacterial communities between patients was found in particular for two patients, one from the United Kingdom and one from the United States, whose sputa contained no species that were common with those found in the other 36 patient samples studied here.

An examination of the species most frequently detected in the sample set (detected in four or more patients and at least once in each geographical cohort) suggested that certain species not traditionally associated with CF respiratory infections are common inhabitants of the CF lung. This is particularly noteworthy given the high degree of microbial diversity reported in this and earlier studies (29, 30, 36). The presence of species not recognized as key CF pathogens in the airways of a substantial number of patients may have important implications for treatment, and this is an area that warrants further investigations (37).

Indeed, a wide range of different bacterial species other than the key CF pathogens has been identified in the past decade (15, 28, 29, 36). Clone sequence analysis from 10 patients, with five patients representing each geographical group, revealed 30 different species from 22 different genera. Species traditionally associated with the CF lung, e.g., *Pseudomonas aeruginosa*, *Staphylococcus aureus*, and *Stenotrophomonas maltophilia* (1), as well as a number of species previously detected in CF samples primarily by culture-independent means, were identified here (15, 28, 29, 36), including one species that had not previously been reported in the context of CF, *Phenyllobacterium koreense*. Eleven (37%) of the species, identified within the genera *Pseudomonas*, *Streptococcus*, *Actinomyces*, *Staphylococcus*, *Parvimonas*, *Neisseria*, *Prevotella*, and *Veillonella*, were common to both the United Kingdom and U.S. patient group. Interestingly, species from within the latter two genera require anaerobic conditions for growth. Anaerobes previously have been detected in CF sputum (3, 15, 29, 37), nonetheless it was striking that even in geographically distinct locations, similar lung physiological conditions were reflected in the species present. However, the majority of species were uniquely found

in one geographical location, suggesting a strong local dominance in their biogeography.

To determine whether there is an uneven geographical distribution of the four most commonly detected T-RF bands, patients were grouped according to whether each band was present in their sputum sample. These groupings are illustrated in the form of a Venn diagram (Fig. 2). Here, two large clusters formed that contained both U.S. and United Kingdom patients to approximately 80% of all patients studied. This suggests that certain species are core to the CF lung by adulthood. Here, it was found that species common by this definition were present within at least one patient from both sample pools. As such, this provides evidence that these species are not endemic to either geographical region.

Similarity indices were employed to identify the degree of similarity between United Kingdom and U.S. samples. The first of these used here was the Raup and Crick similarity index. This analysis generated a set of three clusters representing similar bacterial communities. None of these clusters was found to relate to patient pairs. In contrast, these three clusters comprised one containing only U.S. patients, another containing only United Kingdom patients, and the last containing 10 patients from both the United States and United Kingdom. As such, this provides evidence both for and against the transatlantic distribution patterns of members of the communities detected.

Raup and Crick probability-based similarity indices ( $S_{RC}$ ) were analyzed to determine the degree of significance of the similarities identified between bacterial community memberships (38). Here, a large number of sample pairs with nonsignificant values ( $0.05 > S_{RC} < 0.95$ ) would indicate stochastic dispersal, while a large number of significant  $S_{RC}$  values would indicate a structuring influence, such as niche selection. The paired patients from the United States and United Kingdom showed little tendency to cluster on the basis of community composition, with evidence of this in only ca. 25% of patient pairs. In patient groups from either of the two geographic regions, where no attempt was made to pair on clinical grounds, U.S. patients had communities that clustered to a greater degree (ca. 40%) than U.S.-United Kingdom pairs, whereas United Kingdom patients showed a lower degree of significant clustering (ca. 15%). This suggests that the pairing had only a limited impact on the community findings of this aspect of the study. The reason that the United Kingdom samples as well as the United Kingdom-U.S. matches differed so much in terms of this analysis is not clear, although possible explanations may rest in the differential treatments, exposure to different microbes, or factors that were not assessed. It also is possible that the lower diversity in U.S. samples, as confirmed within the detection limit of this study, allows for the tighter grouping of these samples.

Even though samples may contain similar numbers or types of species, the abundance of those species in the samples could vary substantially, presenting completely different types of ecosystems. Morisita similarity coefficients were used to determine whether relative bacterial species abundance within samples differed between patient groups in the United Kingdom and U.S. As for the Raup and Crick analysis based on species presence or absence, three clusters were identified. Two distinct clusters of exclusively United Kingdom patients (com-

prised of 4 and 6 patients, respectively) were formed with a third cluster of 11 U.S. and 1 United Kingdom patient. Further, when average Morisita index values for United Kingdom, U.S., and paired United Kingdom-U.S. patient groups were examined, U.S. samples again showed a greater level of similarity. Thus, relative species abundance was more affected by biogeography than was community membership.

Comparative studies of bacterial presence in the lungs of patients attending different CF centers previously have only focused on one or more species regarded as pathogens. Johansen et al. (21) found that *P. aeruginosa* and *B. cepacia* complex were more frequently detected in patients attending a CF center in Toronto than in those attending an equivalent center in Copenhagen. This focus on single species made it difficult to carry out a full comparative analysis with the findings here. In our study, bacterial species found to be common to both geographical patient groups included both known CF pathogens, such as *P. aeruginosa*, and those species not traditionally associated with CF lung infections, such as members of the genus *Prevotella*. However, the analysis of larger patient groups will be required to better characterize the geographic distribution of these species that were commonly detected here but as yet are not recognized as CF pathogens.

Certain bacterial species, typically regarded as being of environmental origin, can act as opportunistic pathogens. A number of such species were detected in the samples analyzed here; however, whether these are derived from the wider environment, the immediate treatment environment, or directly from other CF patients is not known. While some such species were detected in both United Kingdom and U.S. patient groups, others showed an uneven geographic distribution. This distribution may be related to differences in the treatment regimes between the two centers or environmental and biogeographical factors. Again, a wider study of this topic is warranted to determine the likely origin of the differences under the condition of patients attending different CF centers (23).

In conclusion, this study has considered the bacterial species present in the lungs of CF patients in two centers in the United States and United Kingdom. Species numbers per sample were similar in the two sites, but differences in terms of species identities were detected, as could be confirmed by the detection limit here. Despite this, the four most frequently detected T-RF bands were present in 80% of all patients, suggesting that certain species are common to the CF lung on both continents. Overall, the U.S. samples showed a stronger association with each other than the United Kingdom samples, which may reflect the lower diversity detected in U.S. samples. Geographical differences were identified in species composition and more strongly in species prevalence in samples taken from patients attending the two CF centers. Overall, it is clear that the bacteriology of the CF lung is complex. Given the importance of lung function to CF patient health, it is by extension important to be able to understand this complexity as the first step in advancing therapy for these patients.

#### ACKNOWLEDGMENTS

This work was supported by the Anna Trust.

Work conducted at UNC-Chapel Hill was supported by a grant from the National Institutes of Health (HL092964).

## REFERENCES

1. Balfour-Lynn, I. A., and J. S. Elborn. 2007. Respiratory disease: infection, p. 137–158. In M. Hodson, D. Geddes, and A. Bush (ed.), *Cystic Fibrosis*, 3rd ed. Hodder Arnold, London, United Kingdom.
2. Bilton, D. 2008. Cystic fibrosis. *Medicine* **36**:273–278.
3. Bittar, F., H. Richet, J. C. Dubus, M. Reynaud-Gaubert, N. Stremmer, J. Sarles, D. Raoult, and J. M. Rolain. 2008. Molecular detection of multiple emerging pathogens in sputa from cystic fibrosis patients. *PLoS One* **3**:e2908.
4. Cho, J.-C., and J. M. Tiedje. 2000. Biogeography and degree of endemicity of fluorescent *Pseudomonas* strains in soil. *Appl. Environ. Microbiol.* **66**: 5448–5456.
5. Corey, M., F. J. McLaughlin, M. Williams, and H. Levison. 1988. A comparison of survival, growth, and pulmonary function in patients with cystic fibrosis in Boston and Toronto. *J. Clin. Epidemiol.* **41**:583–591.
6. Cystic Fibrosis Trust. 2009. Antibiotic treatment for Cystic Fibrosis—Report of the UK Cystic Fibrosis Trust Antibiotic Working Group, 3rd ed., Cystic Fibrosis Trust, Kent, United Kingdom.
7. Davies, J. C., E. W. Alton, and A. Bush. 2007. Cystic fibrosis. *BMJ* **335**:1255–1259.
8. de Wit, R., and T. Bouvier. 2006. “Everything is everywhere, but, the environment selects”; what did Baas Becking and Beijerinck really say? *Environ. Microbiol.* **8**:755–758.
9. Dodge, J. A., P. A. Lewis, M. Stanton, and J. Wilsher. 2007. Cystic Fibrosis mortality and survival in the UK: 1947–2003. *Eur. Respir. J.* **29**:522–526.
10. Flume, P. A., P. J. Mogayzel, Jr., K. A. Robinson, C. H. Goss, R. L. Rosenblatt, R. J. Kuhn, B. C. Marshall, and the Clinical Practice Guidelines for Pulmonary Therapies Committee. 2009. Cystic Fibrosis pulmonary guidelines—treatment of pulmonary exacerbations. *Am. J. Respir. Crit. Care Med.* **180**:802–808.
11. Frederiksen, B., S. Lannig, C. Koch, and N. Hoiby. 1996. Improved survival in the Danish center-treated cystic fibrosis patients: results of aggressive treatment. *Pediatr. Pulmonol.* **21**:153–158.
12. Fierer, N., and R. B. Jackson. 2006. The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci. U. S. A.* **103**:626–631.
13. Gordon, D. M., and F. FitzGibbon. 1999. The distribution of enteric bacteria from Australian mammals: host and geographical effects. *Microbiology* **145**: 2663–2671.
14. Goss, C. H., and M. Rosenfeld. 2004. Update on cystic fibrosis epidemiology. *Curr. Opin. Pulm. Med.* **10**:510–514.
15. Harris, J. K., M. A. De Groot, S. D. Sagel, E. T. Zemanick, R. Kapsner, C. Penvari, H. Kaess, R. R. Deterding, F. J. Accurso, and N. R. Pace. 2007. Molecular identification of bacteria in bronchoalveolar lavage fluid from children with cystic fibrosis. *Proc. Natl. Acad. Sci. U. S. A.* **104**:20529–20533.
16. Health Protection Agency. 2008. British standard operating procedure 57, national standard method—investigation of bronchoalveolar lavage sputum and associated specimens. BSOP 5712. Health Protection Agency, London, United Kingdom.
17. Hodson, M. E., N. J. Simmonds, W. J. Warwick, E. Tullis, C. Castellani, B. Assael, J. A. Dodge, M. Corey, and International Study of Aging in Cystic Fibrosis. 2008. An international/multicentre report on patients with cystic fibrosis (CF) over the age of 40 years. *J. Cyst. Fibros.* **7**:537–542.
18. Hoiby, H. N. 1982. Microbiology of lung infections in cystic fibrosis patients. *Acta Paediatr. Scand. Suppl.* **301**:33–54.
19. Horner-Devine, M. C., M. Lage, J. B. Huges, and B. J. M. Bohannon. 2004. A taxa-area relationship for bacteria. *Nature* **432**:750–753.
20. Hutchison, M. L., and J. R. W. Govan. 1999. Pathogenicity of microbes associated with cystic fibrosis. *Microbes Infect.* **1**:1005–1014.
21. Johansen, H. K., T. A. Kovesi, C. Koch, M. Corey, N. Hoiby, and H. Levison. 1998. *Pseudomonas aeruginosa* and *Burkholderia cepacia* infection in cystic fibrosis patients treated in Toronto and Copenhagen. *Pediatr. Pulmonol.* **26**:89–96.
22. Karpati, F., and J. Jonasson. 1996. Polymerase chain reaction for the detection of *Pseudomonas aeruginosa*, *Stenotrophomonas maltophilia* and *Burkholderia cepacia* in sputum patients with cystic fibrosis. *Mol. Cell Probes* **10**:397–403.
23. Littlewood, J. M. 2005. European cystic fibrosis society consensus on standards—a roadmap to “best care.” *J. Cyst. Fibros.* **4**:1–5.
24. Liu, W. T., T. L. Marsh, H. Cheng, and L. J. Forney. 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphism of genes encoding 16S rRNA. *Appl. Environ. Microbiol.* **63**:4516–4522.
25. Lyczak, J., C. Cannon, and G. Pier. 2002. Lung infections associated with cystic fibrosis. *Clin. Microbiol. Rev.* **15**:194–222.
26. Martiny, J. B., B. J. M. Bohannon, J. H. Brown, R. K. Colwell, J. A. Fuhrman, J. L. Green, M. C. Horner-Devine, M. Kane, J. Adams-Krumins, C. R. Kuske, P. J. Morin, S. Naem, L. Ovreas, A. L. Reysenbach, V. H. Smith, and J. T. Staley. 2006. Microbial biogeography: putting microorganisms on the map. *Nat. Rev. Microbiol.* **4**:102–112.
27. Roemling, U., A. Kader, D. D. Sriramulu, R. Simm, and G. Kronvall. 2005. Worldwide distribution of *Pseudomonas aeruginosa* clone C strains in the aquatic environment and cystic fibrosis patients. *Environ. Microbiol.* **7**:1029–1038.
28. Rogers, G. B., C. A. Hart, J. R. Mason, M. Hughs, M. J. Walshaw, and K. D. Bruce. 2003. Bacterial diversity in cases of lung infection in cystic fibrosis patients: 16S ribosomal DNA (rDNA) length heterogeneity PCR and 16S rDNA terminal restriction fragment length polymorphism profiling. *J. Clin. Microbiol.* **41**:3548–3558.
29. Rogers, G. B., M. P. Carroll, D. J. Serisier, P. M. Hockey, G. Jones, and K. D. Bruce. 2004. Characterization of bacterial community diversity in cystic fibrosis lung infections by use of 16S ribosomal DNA terminal restriction fragment length polymorphism profiling. *J. Clin. Microbiol.* **42**:5176–5183.
30. Rogers, G. B., M. P. Carroll, D. J. Serisier, P. M. Hockey, G. Jones, V. Kehagia, G. J. Connett, and K. D. Bruce. 2006. Use of 16S rRNA gene profiling by terminal restriction fragment length polymorphism analysis to compare bacterial communities in sputum and mouthwash samples from patients with cystic fibrosis. *J. Clin. Microbiol.* **44**:2061–2064.
31. Rogers, G. B., T. W. Daniels, A. Tuck, M. P. Carroll, G. J. Connett, G. J. David, and K. D. Bruce. 2009. Studying bacteria in respiratory specimens by using conventional molecular microbiological approaches. *BMC Pulm. Med.* **15**:9–14.
32. Rogers, G. B., S. Skelton, D. J. Serisier, C. J. van der Gast, and K. D. Bruce. 2010. Determining cystic-fibrosis-affected lung microbiology: comparison of spontaneous and serially induced sputum samples by use of terminal restriction fragment length polymorphism profiling. *J. Clin. Microbiol.* **48**:78–86.
33. Salunkhe, P., C. H. Smart, J. A. Morgan, S. Panagea, M. J. Walshaw, C. A. Hart, R. Geffers, B. Tuemmler, and C. Winstanley. 2005. A cystic fibrosis epidemic strain of *Pseudomonas aeruginosa* displays enhanced virulence and antimicrobial resistance. *J. Bacteriol.* **187**:4908–4920.
34. Shale, D. J. 1996. Cystic fibrosis. Introduction. *Thorax* **45**:880.
35. Sibley, C. D., H. Rabin, and M. G. Surette. 2006. Cystic fibrosis: a polymicrobial infectious disease. *Future Microbiol.* **1**:53–61.
36. Sibley, C. D., M. D. Parkins, H. R. Rabin, K. Duan, J. C. Norgaard, and M. G. Surette. 2008. A polymicrobial perspective of pulmonary infections exposes an enigmatic pathogen in cystic fibrosis patients. *Proc. Natl. Acad. Sci. U. S. A.* **105**:15070–15075.
37. Tunney, M. M., T. R. Field, T. F. Moriarty, S. Patrick, G. Doering, M. S. Muhlebach, M. C. Wolfgang, R. Boucher, D. F. Gilpin, A. McDowell, and J. S. Elborn. 2008. Detection of anaerobic bacteria in high numbers in sputum from patients with cystic fibrosis. *Am. J. Respir. Crit. Care Med.* **177**:995–1001.
38. Van der Gast, C. J., D. Ager, and A. K. Lilley. 2008. Temporal scaling of bacterial taxa is influenced by both stochastic and deterministic ecological factors. *Environ. Microbiol.* **10**:1411–1418.
39. Yankaskas, J. R., B. C. Marshall, B. Sufian, R. H. Simon, and D. Rodman. 2004. Cystic fibrosis adult care—consensus conference report. *Chest* **125**:1S–39S.