

# Molecular identification of potential pathogens in water and air of a hospital therapy pool

Largus T. Angenent<sup>†‡</sup>, Scott T. Kelley<sup>§</sup>, Allison St. Amand<sup>¶</sup>, Norman R. Pace<sup>¶</sup>, and Mark T. Hernandez<sup>||</sup>

<sup>†</sup>Department of Chemical Engineering and Environmental Engineering Science Program, Washington University, St. Louis, MO 63130; <sup>§</sup>Department of Biology, San Diego State University, San Diego, CA 92182; and Departments of <sup>¶</sup>Molecular, Cellular, and Developmental Biology and <sup>||</sup>Civil, Environmental, and Architectural Engineering, University of Colorado, Boulder, CO 80309

Contributed by Norman R. Pace, February 14, 2005

Indoor warm-water therapy pool workers in a Midwestern regional hospital were diagnosed with non-tuberculosis pulmonary hypersensitive pneumonitis and *Mycobacterium avium* infections. In response, we conducted a multiseason survey of microorganisms present in this therapy pool water, in biofilms associated with the pool containment walls, and in air immediately above the pool. The survey used culture, microscopy, and culture-independent molecular phylogenetic analyses. Although outfitted with a state-of-the-art UV-peroxide disinfection system, the numbers of bacteria in the therapy pool water were relatively high compared with the potable water used to fill the pool. Regardless of the source, direct microscopic counts of microbes were routinely  $\approx 1,000$  times greater than conventional plate counts. Analysis of clone libraries of small subunit rRNA genes from environmental DNA provided phylogenetic diversity estimates of the microorganisms collected in and above the pool. A survey of  $>1,300$  rRNA genes yielded a total of 628 unique sequences, the most common of which was nearly identical to that of *M. avium* strains. The high proportion of clones with different *Mycobacterium* spp. rRNA genes suggested that such organisms comprised a significant fraction of microbes in the pool water (to  $>30\%$ ) and preferentially partition into aerosols (to  $>80\%$ ) relative to other waterborne bacteria present. The results of the study strongly validate aerosol partitioning as a mechanism for disease transfer in these environments. The results also show that culture protocols currently used by public health facilities and agencies are seriously inadequate for the detection and enumeration of potential pathogens.

16S rRNA genes | bioaerosols | hypersensitive pneumonitis | molecular survey | *Mycobacterium*

Public warm-water therapy pools and hot tubs are implicated as environments with high exposure to common waterborne and airborne pathogens, such as *Legionella* and *Mycobacterium* spp. (1–5). Monitoring and regulation of the microbiology relevant to public health in warm-water recreational pools and other environments has relied primarily on culture-based analyses that specifically target classical indicator organisms. Although culture can be successful for assessment of some microbes, a large body of gene sequence-based studies shows that standard enrichment techniques significantly underestimate the actual quantity and diversity of microorganisms in a wide variety of environments (6). Further, some waterborne pathogens (e.g., *Vibrio cholera* and *Legionella pneumophila*) are documented to remain viable for extended periods of time, but are unrecoverable by otherwise successful culture protocols (7, 8). Microbes associated with warm-water recreational and therapy pool environments have been studied only by culture-based methods and, because of the limitations of the culture approach, remain poorly characterized.

Microorganisms and other microbiological materials can become airborne and then are termed “bioaerosols.” The enrichment and partitioning of waterborne microorganisms into bioaerosols may be an important factor in disease transmission associated with warm-water pool use. Adverse health effects can be caused by inhalation of viable airborne pathogens and also by inhalation of inactive

microorganisms and/or their component parts (9, 10). For example, toxic or inflammatory pneumonia, bronchitis, and asthma are of growing public health concern and do not depend on viability of microbes to produce effect. Although there is abundant evidence to suggest that airborne microorganisms or their components can cause serious respiratory illness in indoor environments (11–13), there have been few systematic investigations of bioaerosols generated by warm-water pools and their appurtenances.

During March 2000, an indoor warm-water therapy pool at a Midwestern regional hospital was closed in response to employee complaints of extended respiratory problems. In the months preceding the closure, as many as nine employees that worked in the pool area reported the following symptoms: shortness of breath, wheezing, coughing, and night sweats. At least two of the employees were diagnosed by clinical culture with pulmonary infection by *Mycobacterium avium*. In response to this situation, we conducted and report here the results of a multiseason survey to characterize the identity, distribution, and abundance of bacteria in different niches of the therapy pool environment, including the aerosol immediately above the pool. We analyzed the microbial components of samples with standard culture techniques and also a culture-independent method in which small subunit (SSU) rRNA genes are amplified from samples and sequenced for identification. A main attribute of the rRNA sequence approach for characterization of microbial contents is that all organisms, not only specific targets, are detected and identified, and identification does not depend on viability. Molecular techniques have been used previously for water and bioaerosol research, but only for detection of specific microorganisms [e.g., *L. pneumophila* and *Mycobacterium tuberculosis* (14, 15)]

This study of a therapeutic swimming pool is a general molecular analysis of public pool bioaerosols and of a public health application in general. We report that conventional disinfection systems may not inhibit the growth of pathogenic microorganisms at elevated operating temperatures common to therapy pools and hot tubs, and that potential pathogens can preferentially partition into aerosols. We also specifically identify in this pool a host of potential pathogens that may be widespread in such settings and should be considered of public health concern.

## Materials and Methods

**Hospital Therapy Pool Environment.** The hospital therapy pool sampled used an independent ventilation system, separated from the ventilation system of the regional hospital located in the same building complex. The pool air relative humidity was between 25% and 40%, and the temperature was between 25°C and 30°C,

Abbreviations: SSU, small subunit; SAC, swirling aerosol collector; MP, maximum parsimony; NJ, neighbor-joining.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AY897621–AY898248).

<sup>†</sup>To whom correspondence should be addressed at: Environmental Engineering Science Program, Washington University, Campus Box 1180, One Brookings Drive, St. Louis, MO 63130. E-mail: angenent@seas.wustl.edu.

© 2005 by The National Academy of Sciences of the USA

whereas the pool water was held at 33°C. The volume of the pool water was  $\approx 208 \text{ m}^3$ , and the nominal volume of air immediately above the pool was  $1,100 \text{ m}^3$ . The air in the pool building was exchanged 48 times daily with outdoor air by the ventilation system (16). Pool water was cycled four times daily through six parallel, pressurized sand filters (BakerHydro Filtrations, Augusta, GA) and 12 parallel UV units with a UV dose of  $40 \mu\text{W}\cdot\text{s}^{-1}\cdot\text{cm}^{-2}$  and a water residence time of 3 s (Advanced UV Systems, Glendale, CA). Immediately after UV irradiation, hydrogen peroxide was added with a residuum of  $\approx 100 \text{ mg}\cdot\text{liter}^{-1}$ .

**Sample Collection.** Samples for the molecular surveys were collected on February 15, August 16, and December 31, 2001. The pool was occupied at the times of sampling. On Feb. 15 and Aug. 16, inside pool air was sampled with a closed-face,  $\gamma$ -radiated filter cassette that contained a  $0.45\text{-}\mu\text{m}$  pore size GN-6 Metricel (hydrophilic mixed cellulose esters) membrane (Pall Gelman Laboratories, Ann Arbor, MI) and two swirling aerosol collectors (SACs) (Biosampler, SKC, Eighty Four, PA). At the sampling location, 20 ml of sterile 0.01 M PBS containing 0.01% Tween 80 (Sigma) was added to the SACs. The flow rate of all samplers was maintained at a constant level of  $12.5 \text{ liters}\cdot\text{min}^{-1}$  for a sampling period of 1 h (February sampling event:  $0.75 \text{ m}^3$  sample volume) or 2 h 40 min (August sampling event:  $2 \text{ m}^3$  sample volume). The air above the pool was sampled by mounting the inlets of the filter and the SAC at  $\approx 20 \text{ cm}$  above the water surface, taking care to avoid direct splash from entering the samplers. The liquid from one SAC was filtered through an open-face,  $\gamma$ -radiated filter cassette, containing a  $0.20\text{-}\mu\text{m}$  pore size Supor (hydrophilic polyethersulfone) membrane (Pall Gelman), which then was used for the molecular survey. The liquid from the other SAC was transferred to sterile polystyrene 50-ml vials (Corning) for culture and direct count. Approximately 100 ml of pool water sample was filtered through an open-face,  $\gamma$ -radiated filter cassette with a  $0.20\text{-}\mu\text{m}$  pore size Supor membrane.

On August 16, in addition to pool air, the air immediately outside was sampled upwind of the pool building at a distance of 25 m, in a wooded surrounding, by using an otherwise identical sampling protocol except that air samplers used outside were covered with aluminum foil to protect them from sunlight. Biofilm associated with the side of the pool was collected by scraping with an aquarium glass scrubber (Lee's Aquarium and Pet Products, San Marcos, CA) modified to aspirate material through a closed-face,  $\gamma$ -radiated filter cassette with a  $0.20\text{-}\mu\text{m}$  pore size Supor membrane. Approximately 250 ml of pool water was pulled through the filter while scraping the side of the pool just below the water surface.

Beginning in the spring of 2000, grab samples of pool water were collected monthly for direct microscopic counts. On Dec. 31, 2001, 1 ml of sand (and interstitial water) from the pressurized filters was collected by a random grab from a wet filter pan and immediately fixed with 4% paraformaldehyde or frozen for subsequent nucleic acid extraction.

**Bacterial Colony Count and Microscopy Count.** Bacterial colony counts were determined by spreading undiluted SAC liquid onto nutrient-rich tryptic soy agar ( $0.1 \text{ g}\cdot\text{liter}^{-1}$  cycloheximide to inhibit

fungal growth), incubating at 37°C for 2–3 weeks, and counting colonies. For direct microscopic quantification, we followed the method by Hobbie *et al.* (17); pool or SAC liquid was filtered through a  $0.45\text{-}\mu\text{m}$  pore size polycarbonate membrane filter (Osmotics, Minnetonka, MN), which was then stained with  $0.1 \mu\text{g}/\text{ml}$  DAPI (Sigma), washed with sterile PBS, mounted in antifadent (CitiFluor, Leicester, England), and cells were counted with an epifluorescence microscope.

**DNA Extraction and PCR Amplification.** Filters for nucleic acid extraction were stored at  $-80^\circ\text{C}$  until used. Material on filters was eluted with 2 ml 100 mM KCl, 10 mM Tris-HCl (pH 7.4), 10 mM  $\text{NH}_4\text{Cl}$ , and 1% CA-630 (Igepal) (Sigma), and DNA was extracted with a bead-beating protocol (18). SSU rRNA genes were amplified by PCR from the extracted DNA samples according to Frank *et al.* (18). DNA extracts from the February 2001 sampling event were amplified with the universal primer pairs 515F (5'-GTGCCAGC-MGCCGCGGTAA) and 907R (5'-CCGTCAATTCCTT-TRAGTTT), and 515F and 1391R (5'-GACGGCGGTG-WGTRCA). DNA extracts from the August 2001 sampling were amplified with the universal primer pair 515F and 1391R, and the primer pair targeting bacterial genes 27F (5'-AGAGTTTGATC-CTGGCTCAG) and 1391R. DNA extracts from the December 2001 sampling were amplified with the primer pair targeting bacterial genes: 27F and 907R, and the primer pair targeting mycobacterial genes 515F and 1027R (5'-GCACACAGGCCA-CAAGGG), and 515F and 1037R (5'-CATGCACCACCTGCA-CACAG). A typical 50- $\mu\text{l}$  PCR included: 17.5  $\mu\text{l}$  of  $\text{H}_2\text{O}$ ; 5  $\mu\text{l}$  of  $10\times$  PCR buffer; 5  $\mu\text{l}$  of 50 mM  $\text{MgCl}_2$ ; 4  $\mu\text{l}$  of dNTP mix (2.5 mM of each dNTP); 4  $\mu\text{l}$  of 10 mg/ml BSA; 10  $\mu\text{l}$  of 10 M betaine; 200 ng of forward primer; 200 ng of reverse primer; 0.5  $\mu\text{l}$  of AmpliTaq Gold Polymerase (Applied Biosystems); and 2  $\mu\text{l}$  of DNA sample. PCR was conducted with a Mastercycler gradient Machine (Eppendorf, Westbury, NY) by running 20 cycles with a gradient from  $65^\circ\text{C}$  to  $45^\circ\text{C}$  ( $92^\circ\text{C}$  for 30 s;  $65^\circ\text{C}$  to  $1^\circ\text{C}/\text{cycle}$  for 30 s;  $72^\circ\text{C}$  for 90 s) after a hot start at  $94^\circ\text{C}$  for 12 min, and an additional 20 cycles at a  $45^\circ\text{C}$  annealing temperature ( $92^\circ\text{C}$  for 30 s;  $45^\circ\text{C}$  for 90 s;  $72^\circ\text{C}$  for 90 s) before a final extension at  $72^\circ\text{C}$  for 20 min.

**Cloning and Sequence Analyses.** Cloning, restriction fragment length polymorphism (RFLP), and sequencing were performed according to methods described by Frank *et al.* (18) by using pGEM vectors (pGEM-T vector System I, Promega). pGEM vector primers equidistant from the DNA insert, pGEM forward (5'-GAATACT-CAAGCTATGC) and pGEM reverse (5'-AGTGAATTGTAAT-ACGACT), were used to amplify the plasmid inserts before RFLP analysis. Sequence analyses were conducted with a NEN Global IR<sup>2</sup> DNA sequencer (Li-Cor, Lincoln, NE) and a MegaBase 1000 capillary sequencer (Amersham Pharmacia, Piscataway, NJ). In all instances, both strands were sequenced.

**Phylogenetic Analyses.** Approximate microbial species identifications were made by comparing sequences with those in the databases by using BLASTN (19). All sequences identified as mycobacterial were also aligned by using the computer application ARB, which provides rRNA secondary structure information

**Table 1. Heterotrophic plate counts and direct microscopic counts of bacteria in therapy pool water, air above the therapy pool, and air immediately outside therapy pool building during sampling events in 2001**

Sampling event	Environment	Sampler	Bacterial colony count, $\text{m}^{-3}$	Bacterial microscopy count, $\text{m}^{-3}$	Ratio colony and microscopy count, %
Feb. 15, 2001	Pool air	Impinger	$1.7 \times 10^4$	$8.1 \times 10^6$	0.21
	Pool water	Grab		$3.8 \times 10^{11}$	
Aug. 16, 2001	Pool air	Impinger	$6.1 \times 10^2$	$7.3 \times 10^5$	0.08
	Pool water	Grab		$8.5 \times 10^{12}$	
	Outside air	Impinger	$9.0 \times 10^1$	$5.9 \times 10^5$	0.02





**Table 3. Summary and comparison of rRNA sequence identities associated with potentially pathogenic microbes from public databases and the relative abundance of selected sequences to all sequences compiled in clone libraries**

Nearest relative for potential pathogens*	% identity†	Clones‡	Clone	Sample	Sampling date
<i>M. avium</i> complex	99	108/325	MB1008c6	Pool air	Feb. 2001
<i>M. avium</i> complex	99	5/357	MC0401c1	Pool air	Aug. 2001
<i>M. avium</i> complex	99	7/183	MC0424c9	Pool sides <sup>§</sup>	Aug. 2001
<i>M. asiaticum</i>	98	2/325	MB0510c4	Pool air	Feb. 2001
<i>M. asiaticum</i>	99	10/357	MC0411c5	Pool air	Aug. 2001
<i>M. asiaticum</i>	99	55/183	MC0508c9	Pool sides <sup>§</sup>	Aug. 2001
<i>M. fortuitum</i>	99	34/325	MB051512	Pool air	Feb. 2001
<i>M. goodii</i>	99	1/325	MB100811	Pool air	Feb. 2001
<i>M. diemhoferi</i>	99	7/357	MC0610c3	Pool air	Aug. 2001
<i>Staphylococcus hominis</i>	99	2/325	MB050710	Pool air	Feb. 2001
<i>Staphylococcus hominis</i>	99	8/357	MC061013	Pool air	Aug. 2001
<i>Streptococcus pneumoniae</i>	99	6/357	MC0604c7	Pool air	Aug. 2001
<i>P. melaninogenica</i>	99	1/325	MB050715	Pool air	Feb. 2001
<i>P. melaninogenica</i>	98	1/51	MB1001c1	Pool water	Feb. 2001
<i>A. otitis</i>	98	2/357	MC040916	Pool air	Aug. 2001
<i>Gemella morbillorum</i>	99	4/357	MC061116	Pool air	Aug. 2001
<i>Veillonella atypica</i>	99	2/51	MB1003c3	Pool water	Feb. 2001
<i>Coriobacterium glomerans</i>	99	2/51	MB100111	Pool water	Feb. 2001

\*Based on ARB alignment and identification.

†Based on BLASTN identification.

‡Number of clones per total number of clones in sample.

§Pool water and sides of the pool.

involved in the deposition of biofilms in the pool water system. The enrichment of mycobacterial sequences in pool water with side biofilm (Fig. 1F) may indicate that these organisms, too, preferentially associate with and perhaps deposit biofilms. We did not encounter mycobacterial sequences in the one sand filter sample analyzed with universal probes. However, we did detect mycobacterial sequences with mycobacteria-specific primers (Table 2).

**Potential Pathogens in Pool Environments.** A total of 77 different mycobacterial rRNA genes, including some closely related to those of known pathogens, were detected among indoor air sequences. No mycobacteria were detected in the outside air sample. Because SSU rRNA sequences from named *Mycobacterium* spp. are relatively similar to one another, we conducted phylogenetic analyses to resolve the relationships between the pool mycobacteria and previously studied organisms. The alignment used for the analyses included 42 pool sequences and 42 reference sequences representing both “slow growing” and “fast growing” mycobacteria, the two large phylogenetic groups of those organisms (25, 26). One phylogenetic tree that includes the new sequences is shown in Fig. 2. Maximum likelihood, MP, and NJ analyses produced similar trees, with similar bootstrap support at resolved branches. Table 3 lists some particular pathogens to which pool sequences are closely related. The level of relatedness of the pool sequences to those of the selected pathogens, 98–99% identity, corresponds roughly to strain-level differences between the pool organisms and known pathogens, but does not necessarily indicate that the pool organisms detected by the sequences are specifically pathogenic. Nonetheless, the general distribution and relatively high abundance of mycobacterial sequences in pool water and air, particularly those associated with the *M. avium* complex, constitute potential health hazards.

Several potentially pathogenic bacteria besides mycobacteria were detected in the inside pool air and pool water (Table 3). *Prevotella melaninogenica*, for instance, is implicated in a number of afflictions such as vaginosis, periodontal disease, and sinusitis; *Staphylococcus* spp. and *Streptococcus* spp. detected in the pool air and water are associated with diverse human infections. Association of environmental sequences with named pathogens can be misleading as to health risk in some cases, however. *Alloiococcus otitis*,

for instance, detected in one of the samples analyzed, is associated clinically with ear infections but also is implicated as a commensal inhabitant of the outer ear (18). We did not detect *Legionella* spp., which is often cited as common waterborne potential pathogens, in any of the pool samples.

Fungal rRNA sequences, possibly from spores, were abundant in August in both outdoor and indoor aerosol samples, although not in February. Fungal sources evidently were distributed inside the pool by means of the facility’s ventilation system, because no fungal spores were detected in water or biofilm samples (Table 2). No known pathogenic fungi were detected. Predominant fungal rRNA gene sequences outside and inside were from the phyla *Basidiomycota* and *Ascomycota*. The most prevalent fungal sequences in the pool air sample corresponded to those of *Ustilago hordei* (a barley-smut fungus) and *Raciborskiomyces longisetosum* (phylum *Ascomycota*). In addition to eucaryal DNA from fungi, rRNA genes from a nematode (99% identity to the rRNA genes of *Diplolaimelloides meylli*) and a plant (99% identity to *Atropa belladonna* and *Pisum sativum*) were found in the pool air.

## Discussion

In our daily lives we are surrounded by a complex microbiota that is mainly of unknown character. Indeed, before the advent of molecular methods for identification of organisms without culture, it was not possible to know the nature of those unknown microbes, because most evade standard culture identification. The comparison (Table 1) of colony counts with direct microscopic counts of microbes in pool water and air shows clearly the inadequacy of culture for detection of unknown microbes in the pool setting. Identification of environmental organisms by gene sequences does not require culture, and the results are unambiguous in the context of the gene sequences used in the phylogenetic analysis. The rRNA genes are particularly useful for casting a wide phylogenetic net because all organisms contain rRNA genes. Although the rRNA gene is only a rough indicator of phenotype, the rRNA sequences reveal the basic nature of the environmental organisms relative to known organisms. The census of rRNA sequences also is an assessment of the relative abundances of the different organisms represented by the sequences.

