Extensive Genomic Plasticity in *Pseudomonas aeruginosa* Revealed by Identification and Distribution Studies of Novel Genes among Clinical Isolates

Kai Shen,¹ Sameera Sayeed,^{1†} Patricia Antalis,¹ John Gladitz,¹ Azad Ahmed,¹ Bethany Dice,¹ Benjamin Janto,¹ Richard Dopico,¹ Randy Keefe,¹ Jay Hayes,¹ Sandra Johnson,¹ Sujun Yu,¹ Nathan Ehrlich,^{1†} Jennifer Jocz,^{1‡} Laura Kropp,^{1§} Ray Wong,^{1‡} Robert M. Wadowsky,³ Malcolm Slifkin,^{4¶} Robert A. Preston,¹ Geza Erdos,^{1†} J. Christopher Post,^{1,2} Garth D. Ehrlich,^{1,2*} and Fen Z. Hu^{1,2}

Center for Genomic Sciences, Allegheny-Singer Research Institute, Pittsburgh, Pennsylvania 15212¹; Department of Microbiology and Immunology, Drexel University College of Medicine, Allegheny Campus, Pittsburgh, Pennsylvania 15212²; Children's Hospital of Pittsburgh, Pittsburgh, Pennsylvania 15213³; and Clinical Microbiology Laboratory, Department of Pathology, Allegheny General Hospital, Pittsburgh, Pennsylvania 15212⁴

Received 4 April 2006/Returned for modification 25 May 2006/Accepted 23 June 2006

The distributed genome hypothesis (DGH) states that each strain within a bacterial species receives a unique distribution of genes from a population-based supragenome that is many times larger than the genome of any given strain. The observations that natural infecting populations are often polyclonal and that most chronic bacterial pathogens have highly developed mechanisms for horizontal gene transfer suggested the DGH and provided the means and the mechanisms to explain how chronic infections persist in the face of a mammalian host's adaptive defense mechanisms. Having previously established the validity of the DGH for obligate pathogens, we wished to evaluate its applicability to an opportunistic bacterial pathogen. This was accomplished by construction and analysis of a highly redundant pooled genomic library containing approximately 216,000 functional clones that was constructed from 12 low-passage clinical isolates of *Pseudomonas aeruginosa*, 6 otorrheic isolates and 6 from other body sites. Sequence analysis of 3,214 randomly picked clones (mean insert size, ~1.4 kb) from this library demonstrated that 348 (10.8%) of the clones were unique with respect to all genomic sequences of the *P. aeruginosa* prototype strain, PAO1. Hypothetical translations of the open reading frames within these unique sequences demonstrated protein homologies to a number of bacterial virulence factors and other proteins not previously identified in *P. aeruginosa*. PCR and reverse transcription-PCR-based assays were performed to analyze the distribution and expression patterns of a 70-open reading frame subset of these sequences among 11 of the clinical strains. These sequences were unevenly distributed among the clinical isolates, with nearly half (34/70) of the novel sequences being present in only one or two of the individual strains. Expression profiling revealed that a vast majority of these sequences are expressed, strongly suggesting they encode functional proteins.

Pseudomonas aeruginosa is a broadly disseminated environmental gram-negative bacillus with a large genome (>6.2 Mb) that is highly adaptive and able to flourish in a wide range of free-living, commensal, and pathogenic environments. It has invasive and toxigenic characteristics that can cause disease across a broad range of eukaryotic phyla, including insects, vertebrates, and higher plants. In humans, it is usually an opportunistic pathogen that can colonize locally or be disseminated systemically in patients who have sustained trauma or are immunodeficient. *P. aeruginosa* is the major pathogen as-

sociated with spontaneously perforating otorrhea, it is the scourge of burn victims, and it can cause life-threatening infections in diabetics and patients receiving chemotherapy or those with acute ulcerative keratitis. In addition, it is associated with otitis externa (swimmer's ear) and hot tub folliculitis. In cystic fibrosis (CF) patients, *P. aeruginosa* infects the lung and is the major chronic pathogen associated with morbidity and mortality. The ability of this bacterium to persist and multiply in oligotrophic environments, such as hospital-based humidifiers, contributes to its importance as a nosocomial pathogen (18, 32).

Many *P. aeruginosa* clinical isolates are multiply resistant to antibiotics at concentrations achievable in vivo through systemic administration, and single or combination antimicrobial therapy fails in 5% to 70% of instances, depending on the series. The basic mechanisms of antibiotic resistance, including alteration of the drug target, prevention of drug access to the target, and drug inactivation, mostly arise via the acquisition of exogenous genetic material through horizontal gene transfer (HGT) mechanisms (29, 58). These studies suggested that other genes associated with virulence may also be commonly

* Corresponding author. Mailing address: Center for Genomic Sciences, Allegheny-Singer Research Institute, Allegheny General Hospital, 320 East North Ave., 11th Floor South Tower, Pittsburgh, PA 15212. Phone: (412) 359-4228. Fax: (412) 359-6995. E-mail: gehrich@wpahs.org.

† Present address: University of Pittsburgh, Pittsburgh, PA 15213.

‡ Present address: Carnegie-Mellon University, Pittsburgh, PA 15213.

§ Present address: Pennsylvania State University, State College, PA 16802.

¶ Deceased.

exchanged among the pseudomonads and even among phylogenetically diverse genera via similar mechanisms. Studies of multiple clinical isolates of both the gram-negative *Haemophilus influenzae* and the gram-positive *Streptococcus pneumoniae* demonstrated that both of these naturally transformable obligate human pathogens have a population-based supragenome that is multiple times the size of a given bacterium's genome. Moreover, it has been shown that the infecting populations of these species are polyclonal in nature and that recombination takes place among the strains during infection, leading to the development of novel strains (17, 33, 44, 45, 47–49; J. Hogg et al., unpublished data). We wished to determine if the distributed genome hypothesis (DGH) (15) and its associated virulence corollary were also applicable to bacterial species that had adapted to a broader array of ecological niches.

The natural population structure of *P. aeruginosa* infections is not fully understood with regard to genotypic clonality, but there is consensus that this organism can form heterogeneous three-dimensional biofilms and that its life cycle is complex, composed of multiple discrete stages including planktonic, reversible attachment, irreversible attachment, biofilm growth and maturation, and active dispersion wherein planktonic showers vent from mature biofilm structures (40, 61). It is also clear that *P. aeruginosa* biofilms are composed of multiple types of differentiated cells (6, 23, 40) even when grown in vitro from a clonal population. Thus, these cells have the intrinsic capacities for development and differentiation, phenomena long believed to be unique to eukaryotes (38). Clearly, their ability to undergo myriad phenotypic transformations provides the population as a whole with profound resistances to both host defenses and antibiotics under clinical conditions (9, 11, 14, 15, 43). A further layer of complexity may result if natural infecting populations are polyclonal. Such a scenario would provide for continuous reassortment of genic characters among strains during the infectious process that would result in the production of new strains, some of which would have a reproductive advantage in the host environment.

It has long been known that the various *P. aeruginosa* strains display different phenotypic characters, but the mechanisms underlying these differences could not definitively be ascribed to expression or genomic heterogeneity. Serotyping of *P. aeruginosa* has been used for nearly a century to differentiate among strains (3), and multiple schemata have been developed to aid in this process. Liu et al. have identified 20 *P. aeruginosa* serotypes based on the structure of the lipopolysaccharide (30, 31), but this method lacks sufficient discriminating power, identifies differences at only a very limited set of loci, and is not applicable to rough-type colonies. Thus, more comprehensive methods of strain discrimination using genomic fingerprinting approaches have been developed including, restriction fragment length polymorphism typing, pulsed-field gel electrophoresis, and PCR-based methods. Each of these techniques is capable of discriminating among large numbers of strains (52); however, none of these methods is capable of identifying the individual coding elements that compose a genome, and so it is not possible to distinguish between allelic and genic differences using these techniques.

Comparative whole-genome DNA sequence analysis provides the most comprehensive means to identify differences among bacterial isolates. The complete genomic sequences of

two *P. aeruginosa* strains have been determined: PAO1, the prototype lab strain (54), and UCBPP-PA14, a pathogenic strain (53). Comparative evaluations of these strains have revealed that each strain contains unique genes with respect to the other. Whole-genome comparisons, while ideal, are currently impractical under most circumstances, especially for organisms with larger genomes, but the random sequencing of clones from multiple strains compared against a fully sequenced prototype strain can provide an estimate of the genomic plasticity within a species (44, 45). Recently, Spencer et al. partially sequenced (average coverage of $\sim 0.5\times$) the genomes of three *P. aeruginosa* strains, two CF isolates and one aquatic isolate, wherein they showed that at least 10% of genes were novel with respect to PAO1.

The current study was designed to estimate the degree of genomic plasticity (i.e., genic differences, as opposed to allelic differences) among human pathogenic strains of *P. aeruginosa*. This was accomplished through the random sequencing of clones from a highly redundant, pooled genomic library prepared from 11 low-passage clinical isolates obtained from patients with otorrhea ($n = 6$) or other chronic infections ($n = 5$) and 1 clinical strain from the American Type Culture Collection (ATCC).

MATERIALS AND METHODS

Acquisition of clinical strains and bacterial growth conditions. Six *P. aeruginosa* strains, designated Pitt A, Pitt B, Pitt C, Pitt D, Pitt E, and Pitt F, were obtained as first-plate pure cultures from the Children's Hospital of Pittsburgh, where they had been cultured from the discharges of pediatric patients suffering from otorrhea. An additional five isolates, designated M18851, M18858, W27912, W27931, and W28869, were cultured from patients with various chronic infections at Allegheny General Hospital in Pittsburgh. A 12th strain, 27853, was obtained from the ATCC. All strains were streaked for single colonies upon receipt. Using a single colony for inoculation, each strain was grown to early stationary phase in a 200-ml Luria-Bertani (LB) broth (Difco Laboratories, Detroit, MI) culture that was then used to make 300 aliquots in 15% glycerol for storage at -80°C . All subsequent experiments for all strains were performed using one of the original freezes in an effort to limit the number of passages in vitro. The strains were grown at 37°C prior to genomic DNA extraction. Due to a procedural error, the Pitt B strain was lost after the library construction.

TOP10 One Shot *Escherichia coli* competent cells (Invitrogen Corp., California) were grown in LB broth or on LB agar. Ampicillin (50 $\mu\text{g/ml}$) or kanamycin (30 $\mu\text{g/ml}$) was supplemented in the culture medium for clone selection.

Isolation of bacterial genomic DNA and total RNA. Genomic DNA was extracted essentially as described elsewhere (4). Briefly, cells were collected by centrifugation from 100-ml overnight cultures and resuspended in TE buffer (10 mM Tris-HCl, pH 8.0; 1 mM EDTA, pH 8.0). Bacteria were lysed by the addition of sodium dodecyl sulfate (SDS; Invitrogen) to a final concentration of 0.5%. The lysates were then digested with RNase A (50 $\mu\text{g/ml}$; Gentra Systems, Inc., Minneapolis, MN) at 37°C for 1 h and then with proteinase K (100 $\mu\text{g/ml}$; Invitrogen) at 37°C for 1 h. Cetyltrimethylammonium bromide (Sigma) was added to a final concentration of 1%, and the mixture was incubated at 65°C for 20 min. Following a chloroform-isoamyl alcohol (24:1) extraction, the DNA was precipitated from the aqueous phase with 0.6 volume of isopropanol. The DNA was then pelleted by centrifugation and washed with 70% ethanol. After air drying, the pellets were resuspended in TE buffer at 65°C for 1 h. The purified DNA was tested for quantity and quality by UV absorbance at 260/280 nm (DU650 spectrophotometer; Beckman) and agarose gel electrophoresis, respectively.

Total bacterial RNAs were extracted using the hot phenol method (10). Briefly, stationary-phase cultures (optical density at 600 nm of 4) were treated with a cell stop solution (1 volume of buffer-saturated phenol, pH 6.6, and 19 volumes of 100% ethanol), and the bacteria were pelleted at $5,500 \times g$ for 5 min at 4°C . The pellet was resuspended in cell resuspension solution (10 mM KCl, 5 mM MgCl_2 , 10 mM Tris, pH 7.5) followed by the addition of fresh hot phenol solution (five parts hot phenol buffer [400 mM NaCl, 40 mM EDTA, 1% β -mercaptoethanol, 1% SDS, 20 mM Tris pH 7.5] and one part buffer-saturated phenol

[pH 6.6; Ambion, Inc., Austin, TX)], heated at 95°C for 2 min, and centrifuged at $20,000 \times g$ for 15 min at room temperature. The aqueous layer was recovered and extracted, once with acid phenol-chloroform (5:1) and then twice with chloroform, followed by centrifugation at $20,000 \times g$ for 15 min at room temperature. The RNA was precipitated with isopropanol, placed on dry ice for 10 min, and centrifuged at $16,000 \times g$ for 15 min at 4°C. The RNA pellet was washed twice with 70% ethanol and resuspended in $1 \times$ TE buffer at 37°C for 30 min followed by treatment with TURBO DNase (Ambion). RNA quality was checked using an Agilent 2100 Bioanalyzer with the RNA 6000 Nano assay kit (Agilent Technologies, Palo Alto, CA).

Construction of the pooled genomic library. A pooled genomic library was constructed using equimolar quantities of DNA from each of 12 clinical isolates of *P. aeruginosa* as described previously (16, 46). Briefly, the genomic DNA from each isolate was hydrodynamically fragmented to an average size of ~ 1.5 kb using the HydroShear apparatus (GeneMachines, San Carlos, CA) according to the manufacturer's instructions. Aliquots (10 μ g each) of the sheared DNA preparations were pooled, end repaired, ligated into the vector pCR4Blunt-TOPO, and transformed into *E. coli* TOP10 competent cells according to the manufacturer's protocol (Invitrogen Corp., Carlsbad, CA). The transformants were plated onto 22-cm² LB agar plates (Genetix, Christchurch, United Kingdom) at densities between 2,000 and 3,000 colonies per tray and incubated at 37°C for 16 h. A total of 241,152 bacterial colonies were picked and arrayed using the CGS multifunctional three-axis Q-bot (Genetix Limited, United Kingdom). Arrayed clones were chosen randomly for further analysis.

DNA sequencing. Plasmid DNA templates for sequencing were prepared using the QIAprep Miniprep kits (QIAGEN Inc., Valencia, CA) on a Beckman FX robot (Beckman Instruments), digested with EcoRI (Invitrogen), and analyzed on ethidium bromide-stained 1% agarose gels in Tris-acetate-EDTA buffer. Only constructs containing inserts of >0.5 kb were used as sequencing templates. Dideoxy sequencing was performed as described for both LiCor IR² Gene ReadIR instruments (LiCor, Inc., Lincoln, Nebr.) and Beckman CEQ 2000 XL automated capillary electrophoresis sequencing instruments (Beckman Coulter, Inc., Fullerton, CA) (44). Confirmatory sequencing was performed using an ABI 3730xl DNA analyzer (Applied Biosystems Inc., Foster City, CA) in which the sequencing reactions were prepared in 3- μ l volumes, using the Parallab 350 nanoliter genomic work station (Brooks Automation, Inc., Chelmsford, MA). For Parallab-based sequencing, the reaction mixtures consisted of the following: (i) 1.4 μ l of plasmid template (approximately 100 ng DNA); (ii) 0.5 μ l of 10-pmol/ μ l primer; and (iii) 1.1 μ l from the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems). Reaction aliquots of 500 nl were then thermal cycled and purified within the Nano-Pipetter of the Parallab. Cycling conditions were as follows: 35 cycles with a 0-s denaturation step at 96°C, a 0-s annealing step at 50°C, and a 60°C extension step for 45 s. The sequencing reactions were then run on the ABI 3730xl and analyzed using the ABI Analysis software v5.1.

DNA sequence analysis. Sequencing results were analyzed and contigs were determined using Sequencher (version 4.1.4; Gene Codes Corporation, Ann Arbor, MI). DNA sequence similarity searches utilizing the BLASTn and BLASTx algorithms (2) were performed using the CGS high-speed BLAST cluster (J. Gladitz, G. Erdos, and S. Yu, unpublished data) that is automatically updated on a weekly basis from the National Center for Biotechnology Information website (<http://www.ncbi.nlm.nih.gov/>). This system, including a custom-designed software package, automatically performs vector trimming, sequence quality checks, and BLAST homology searches, permitting fast and accurate analyses of thousands of clones daily.

PCR-based gene distribution studies. Primer pairs were designed using Oligo 6.65 (Molecular Biology Insights, Cascade, CO) and Vector NTI 9.0 (Invitrogen) for the performance of distribution studies among the parent strains for the 70 unique genes. Primer sequences are available on the CGS website (Supplementary Table 1 at www.centerforgenomicsciences.org). Genomic DNAs isolated from each of the clinical strains were used as amplification templates to determine which clinical strains possessed each of the non-PAO1 sequences. Positive controls for each genomic template DNA were also performed using primers specific for the *P. aeruginosa* polyphosphate kinase (*ppk*) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) genes (see Supplementary Table 1 at the CGS website). Amplifications were performed in a volume of 12.5 μ l containing 0.3 units of *Taq* DNA polymerase, 50 ng of template DNA, 5 pmol of each primer, 1.5 mM MgCl₂, and 0.4 mM deoxynucleoside triphosphates and using the Eppendorf MasterTaq kit (Brinkman Instruments, Inc., Westbury, NY). PCR was performed using a suite of Perkin-Elmer 9600 thermal cyclers programmed as follows: 1 cycle at 95°C for 2 min; 35 cycles at 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s; 1 cycle for 7 min at 72°C; and then a 4°C hold. The PCR products were electrophoresed through 1.75% agarose gels, stained with ethidium bro-

me, photographed with a Kodak Image Station 440 UV light scanner (NEN Life Science Products, Inc., Boston, MA), and analyzed with Kodak 1D Image Analysis software (Eastman Kodak Company, Rochester, NY).

RT-PCR detection of transcripts encoded by novel ORFs. Reverse transcriptase PCR (RT-PCR) was used to determine if the novel sequences corresponded to functional gene units. An RNA premix was prepared by mixing random hexamers (Invitrogen) with 4 μ g of total RNA, followed by heat denaturation, quenching on ice, and aliquoting into two tubes. Two master mixes were prepared, one containing all of the RT components, including Moloney murine leukemia virus RT (Invitrogen) (+RT) and the other one lacking the RT enzyme (-RT). Each pair of RNA specimens received an aliquot of the +RT mixture in one tube and the -RT mixture in the other. PCR amplifications were carried out using 2.5 μ l of the first-strand cDNA from each RT reaction mixture. *P. aeruginosa* GAPDH mRNA was amplified as a positive control for each strain. Negative controls contained identical reaction mixtures but no template nucleic acid.

Southern blot analysis. *P. aeruginosa* genomic DNAs were isolated from liquid cultures as described above, digested with EcoRI, and electrophoresed into 1% agarose gels. Each gel also contained one lane with a pool of the unique plasmid clones for which we were probing, which served as a positive control. The DNAs were transferred to positively charged nylon membranes by capillary blotting using 0.4 M NaOH (Amersham Pharmacia Biotech, Buckinghamshire, United Kingdom) according to the method of Southern (51). Probes for the unique genes under study were produced by PCR-based amplification of the corresponding plasmid inserts, followed by purification of the amplicons by using the QIAquick PCR purification kit (QIAGEN). Radioactive labeling of the probes was performed using the random primer DNA labeling system (Invitrogen) according to the manufacturer's instructions. Probes were purified via gel exclusion chromatography (G-50 Sephadex columns; Roche Diagnostics, Indianapolis, IN). Specific activity was measured in a Bioscan QC4000XER counter (Bioscan, Washington, D.C.). A 30-min prehybridization of the filters was carried out at 42°C (5 $\times$ SSC [1 $\times$ SSC is 0.15 M NaCl plus 0.015 M sodium citrate], 5 $\times$ Denhardt solution, 50% [wt/vol] formamide, 1% [wt/vol] SDS) with heat-denatured sheared salmon sperm DNA (Sigma) added immediately before incubation. Hybridization of the novel gene probes to the blotted DNAs was accomplished by adding $\sim 2 \times 10^7$ dpm of heat-denatured probe to each prehybridization reaction mixture and incubating at 42°C overnight. Following hybridization, the membranes were washed in 2 $\times$ SSC-0.1% (wt/vol) SDS for 5 min at room temperature three times, then in 0.2 $\times$ SSC-0.1% (wt/vol) SDS at room temperature for 15 min, then in 0.2 $\times$ SSC-0.1% (wt/vol) SDS at 42°C for 15 min, and finally in 0.1 $\times$ SSC-0.1% SDS at 68°C for 15 min. The membranes were then autoradiographed using Kodak XAR film.

Nucleotide sequence accession numbers. The 70 novel nucleotide sequences reported in this paper have been deposited in GenBank with the following accession numbers: CL422968 to CL422970, CL422972 to CL422973, CL422976 to CL422983, CL440257 to CL440259, CL440261 to CL440265, CL440267 to CL440270, CL440272 to CL440280, and DU708885 to DU708920.

RESULTS AND DISCUSSION

Characterization of *P. aeruginosa* pooled genomic library. A pooled genomic DNA library comprised of 241,152 clones was constructed and arrayed using equimolar amounts of hydrodynamically sheared DNA from 12 *P. aeruginosa* clinical isolates. Sampling and analysis of 25,975 clones determined that $\sim 96\%$ contained plasmids. Sequencing of 3,453 randomly chosen clones demonstrated that $\sim 93\%$ contained inserts of >100 bp, suggesting a functional library of $\sim 216,000$ clones. The average insert size was 1.4 kb, resulting in a library of $\sim 3.0 \times 10^8$ bp of DNA, which translates to $\sim 4\times$ coverage of each of the 12 genomes, assuming a genome size of 6.2 Mb. The average readable length for each sequencing reaction was ~ 650 bases. For clone sequences that were longer than 1.3 kb, primer walking was used to sequence the interior of the fragment.

Identification of novel (non-PAO1) sequences from clinical isolates. The nucleotide sequences from the 3,214 clones containing inserts of >100 bp were analyzed by BLAST against the reference PAO1 genome (54) to identify novel sequences contained within the genomes of the 12 clinical isolates. A total of

89.2% (2,866/3,214) of the clones revealed a minimum of 350 bp of contiguous homology (at least 80% nucleotide identity) to PAO1 and were classified as PAO1-like; however, most of these sequences showed allelic differences, including small indels and point mutations. The remaining 348 (10.8%) sequences were classified as novel and did not contain any blocks of nucleotide homology (>100 bases) when compared with all known sequences in the GenBank database at the time of the initial search. BLASTx searches using conceptual translations of the novel sequence ORFs identified varying degrees of similarity to proteins from throughout the prokaryotic domain (see Supplementary Table 2 at the CGS website). Some of the novel sequences displayed very low levels of similarity to any known proteins and thus may encode novel peptides with new functions not previously recognized, whereas other sequences showed relatively higher similarity to known virulence factors or other structural or metabolic proteins of known function. Seventy high-quality sequences were randomly selected for genetic distribution studies among the component strains that make up the pooled library to gauge the degree of interstrain genomic plasticity. Among these unique sequences, we identified a range of hypothetical genes by ORF analysis that are predicted to encode phage proteins or virulence-related determinants, as well as proteins associated with cellular metabolism. Below we describe a subset of these putative novel gene products.

Identification of novel genes. We undertook to examine in more detail a minority of these randomly chosen unique pseudomonal sequences to try to determine if the distributed genes fell within groups of genes known to be associated with horizontal gene transfer and originated within a limited set of related species. Our analyses of the various hypothetical genes indicated that the genomic plasticity observed was not dependent on a single group of genes, but instead included genes from numerous functional classes. The only group of genes previously associated with high rates of HGT that were over-represented among our study set were phage-associated genes. Interestingly, many of the novel genes we identified were present in only a minority of the strains examined, suggesting that there is a continual accrual of foreign genes into the pseudomonal supragenome and that many of them may not be fixed in the population. Moreover, the sources of the various unique genes, based on homology searches, suggest that a very broad array of species shares genes across many functional groups.

(i) Phage and phage-like sequences. Hypothetical phage genes were the most prevalent class of novel clones, with 20/70 (29%) of the unique sequences demonstrating some level of protein similarity to known or putative phage genes (Table 1; see also Supplementary Table 2 at the CGS website). Bacteriophages play a critically important role in HGT among the pseudomonads, as they are incompetent for natural transformation (38, 42). Over 60 temperate bacteriophages have been isolated from *P. aeruginosa* (1, 57), all of which can integrate their genomes as well as transduce genes from their previous hosts into their new host's genome via transposition. Clone L001_PA_0515_A13 contains four ORFs with between 84 and 94% amino acid identity to the *Pseudomonas* phage D3; however, the nucleotide homology was too low in all cases to identify a match. This phage contains an *O*-acetylase gene

(encoded by ORF28) that adds an acetyl group to the lipopolysaccharide O-antigen which results in a serotype switch from O5 to O16 (27) in the presence of D3-encoded α -polymerase inhibitor and a β -polymerase gene (34). The ORFs represented in our clone are ORF21, ORF22, ORF25, and ORF26, suggesting that ORFs 23 and 24 have been deleted. We also identified six ORFs (24, 53, 59, 60, 62, and 63) from our random clone sequencing that showed limited protein similarity to the phage F116. Clone L001_PA_0166_M20 appears to be a chimeric gene that resulted from recombination between two known integrases, (ORF100) from bacteriophage pf1 and an integrase present in PA0728. Clone L001_PA_0022_J20 contains two ORFs with 86% and 88% amino acid identity to proteins of phage B3, and clone L001_PA_0021_F02 contains an ORF whose hypothetical translation would produce a protein with 95% amino acid identity to the CP7 gene of the plasmid pKLC102, which can integrate into the genome of *P. aeruginosa* C strains (24) at the tRNA^{Lys} gene. Clone L001_PA_0018_I13 contains an ORF whose highest similarity (70% amino acid identity) is to the NLP/P60 amidohydrolase gene in *Pseudomonas syringae*. This protein acts on γ -glutamyl-containing substrates such as peptidoglycans (41); however, most of the identified protein homologs for this gene are phage tail assembly proteins, such as amidase protein K of phage lambda. Moreover, the adjacent ORF in this clone shows similarity to another phage tail protein, suggesting that both ORFs are probably phage encoded.

(ii) Novel *soxR*-like gene. Clone L001_PA_0002_J02 contains an ORF that encodes a hypothetical protein with homology (59% amino acid identity, 75% similarity) to the putative SoxR protein of *P. putida* KT2440 (see Supplementary Table 2 at the CGS website). This gene and its transcripts were identified in 3 out of 11 of the clinical strains. In PAO1 there is another *soxR* gene (PA2273) that controls a six-gene regulon in response to oxidative stress (36); this gene is also required for virulence in a mouse model of intrapulmonary infection and is highly induced in *P. aeruginosa* strain PAK in the mouse burn model (20). The *soxR*-like clone identified here has only a short region (nucleotides 92 to 168) of limited nucleotide homology (75%) to the A2273 gene, indicating that it is most likely a unique gene within the same gene family.

Orthologs of this gene in other bacterial species, including *E. coli*, also encode transcriptional regulators associated with superoxide responsiveness (22) which in turn activate another transcriptional regulator, *soxS*, that activates a regulon that includes dozens of oxidative and nitric oxide stress response genes (12, 39). *P. aeruginosa* does not have a *soxS* homolog, suggesting that other genes, perhaps *soxR* homologs, may function in additional oxidative stress pathways in this species (25).

(iii) *hlyD*-like gene. Clone L001_PA_0020_A06 contains an ORF encoding a hypothetical protein that shares similarity to the putative HlyD protein (50% amino acid identity, 64% similarity) of the marine oligotroph *Sphingopyxis alaskensis* RB2256. The *hlyD* gene product in *E. coli* is known to be part of an oligomeric protein that also contains the products of the *hlyB* and *-C* genes and that is essential for the secretion of α -hemolysin, a cytotoxin encoded by the *hlyA* gene (37). There are numerous *hlyB*-, *-C*-, and *-D*-like genes throughout the bacterial domain that are associated with various membrane transporter mechanisms for the secretion of extracellular pro-

TABLE 1—Continued

Clone or contig name	Amplimer size (bp)	Data set	<i>P. aeruginosa</i> strain										No. of strains with non-PAO1 sequences	ORF similarity ^a	
			Pitt A	Pitt C	Pitt D	Pitt E	Pitt F	27853	M18851	M18858	W27912	W27931			W28869
L001_PA_0223_C22	92	PCR	+	-	+	-	-	-	+	+	+	+	-	7	Putative pathogenesis-related protein
L001_PA_0286_K13	80	RT-PCR	+	-	+	-	-	-	-	-	-	-	-	4	Wzz
		PCR	-	-	+	-	-	-	-	-	-	-	-	1	
L001_PA_0322_I08	58	RT-PCR	-	+	+	+	+	+	+	+	+	+	-	10	Conserved hypothetical protein
L001_PA_0325_C06	179	PCR	-	+	+	+	-	-	-	-	+	+	-	4	
		RT-PCR	-	+	+	+	-	-	-	-	+	+	-	5	Putative transposase
L001_PA_0410_L05	88	RT-PCR	-	-	+	+	+	+	+	-	+	-	-	4	
		PCR	+	-	-	+	+	+	-	-	-	-	-	6	Hypothetical protein
L001_PA_0419_M18	534	RT-PCR	+	-	+	+	-	-	+	-	-	-	-	2	
		PCR	+	+	-	+	+	+	+	+	-	-	-	6	Hypothetical protein mesoDRAFT_3262
L001_PA_0450_P07	283	RT-PCR	+	+	-	-	+	-	+	+	-	-	-	5	
		PCR	-	-	-	+	-	-	-	-	-	-	-	1	Hypothetical protein F116p54
L001_PA_0452_M03-0582_F16	87	RT-PCR	-	-	-	+	-	-	-	+	-	-	-	1	
		PCR	-	-	-	+	-	-	-	-	-	-	-	2	ABC-type amino acid transport/signal transduction systems, periplasmic component/domain
L001_PA_0454_K21	66	RT-PCR	-	-	-	-	+	-	-	-	-	-	-	1	
		PCR	-	-	-	-	-	-	-	-	-	+	-	1	α-Amylase, catalytic subdomain
L001_PA_0459_F21	211	RT-PCR	-	-	-	+	-	+	-	-	+	+	-	1	
		PCR	-	-	-	+	-	-	-	-	-	-	-	3	Bacteriophage Mu Gp30-like protein
L001_PA_0465_E21	94	RT-PCR	-	-	-	+	-	-	-	-	-	-	-	1	
		PCR	-	-	-	+	-	-	-	-	-	-	-	2	Hypothetical protein t2948
L001_PA_0501_J07	87	RT-PCR	-	-	+	+	+	+	-	+	-	-	-	1	
		PCR	-	-	+	+	+	-	-	+	-	-	-	2	Endonuclease VII
L001_PA_0515_A13	69	RT-PCR	-	-	+	+	-	-	-	-	-	-	-	1	
		PCR	-	-	+	+	-	-	-	-	-	-	-	2	Pseudomonas phage D3 Orf26
L001_PA_0516_I04	53	RT-PCR	-	+	-	-	-	-	-	-	-	-	-	1	
		PCR	-	+	-	-	-	-	-	-	-	-	-	2	Hypothetical protein BGp138
L001_PA_0541_D04	65	RT-PCR	-	-	+	+	+	+	-	-	-	-	-	2	
		PCR	-	-	+	+	+	-	-	-	-	-	-	2	Conserved hypothetical protein
L001_PA_0550_D02	63	RT-PCR	-	-	-	-	-	-	-	-	-	+	-	1	
		PCR	-	-	-	-	-	-	-	-	-	-	-	2	ABC-type multidrug transport system, ATPase and permease components
L001_PA_0553_G08	80	RT-PCR	-	-	-	-	-	-	-	-	-	+	-	1	
		PCR	-	-	-	+	-	-	-	-	-	-	-	1	Hypothetical protein F116p60
L001_PA_0592_N09	55	RT-PCR	-	-	-	+	-	-	-	-	-	-	-	1	
		PCR	-	-	-	+	-	-	-	-	-	-	-	1	Hypothetical protein pspt03407
L001_PA_0607_O20	202	RT-PCR	-	-	-	+	-	-	-	-	-	+	-	1	
		PCR	-	-	-	-	-	-	-	-	-	-	+	1	Na ⁺ /H ⁺ antiporter NhaA
L001_PA_0613_O07	57	RT-PCR	-	-	-	-	-	-	-	-	-	+	-	1	
		PCR	-	-	-	-	-	-	-	-	-	+	+	1	alkBFGHKL regulator

^a See Supplementary Table 2 at the CGS website for detailed information.

TABLE 4. Frequencies of *P. aeruginosa* strains showing unique sequences

No. of strains with trait	No. or % of sequences or transcripts found in the indicated no. of strains ^a			
	No. of unique DNA sequences	% Unique DNA sequences	No. of unique RNA transcripts	% Unique RNA sequences
11	1	1.4	0	0
10	3	4.3	0	0
9	1	1.4	2	2.9
8	1	1.4	0	0
7	7	10.0	0	0
6	6	8.6	3	4.3
5	4	5.7	2	2.9
4	7	10.0	7	10.0
3	6	8.6	11	15.7
2	15	21.4	13	18.6
1	19	27.1	32	45.7

^a The percentages are based on the number of occurrences out of a maximum of 70.

(*ppk*) and GAPDH genes were used as positive controls and were successfully amplified from all 11 clinical strains and PAO1. All distribution and expression assays were performed in quadruplicate, and a positive call required that at least 3/4 of the independent results were positive. Our PCR-based distribution analyses are theoretically limited by the fact that divergent alleles of the novel genes that vary significantly at the primer-binding sites could produce a false-negative result, which would lead to an underestimation of their penetration into the pseudomonal gene pool. Thus, we performed Southern hybridizations for six of the clones as a test of the PCR results. Overall, these two methods were highly corroborative, but there were a few minor differences (Table 3). Nevertheless, it remains possible that genes encoding surface-exposed antigens, which come in contact with the host immune system, could display heterogeneity beyond what could be detected even by Southern hybridization analyses.

The 70 novel DNA sequences were found in 33% (258/770) of all possible occurrences, with at least one strain harboring each of the novel sequences, indicating that none of these sequences were contaminants (Table 1). PAO1 DNA did not support amplification with any of the 70 primer pairs devel-

oped for the unique sequences. The Pitt E strain contains the most novel sequences (31/70) based on this sampling, whereas the Pitt C strain has the least (11/70). On average each of the 11 clinical strains contained less than one-third (23/70) of the unique genes. Nearly half of the sequences (34/70, or 48.5%) were found in only one or two of the strains, and only one sequence (L001_PA_0031_E04) was identified in all 11 clinical isolates (Table 4).

RNA transcripts were detected in concert with 64% (165/258) of the genic occurrences, and all 70 of the novel sequences were expressed in at least one of the clinical strains, demonstrating that in all likelihood each of these novel sequences represents a functional gene. The mean number of genes expressed across all strains was 15/70 and, similar to the DNA sequence distribution studies, the Pitt E strain demonstrated expression of the highest number (29/70) of RNA transcripts. In contrast the ATCC strain 27853 expressed the fewest (2/70) of these novel genes (Table 2). A majority (45/70, or 64.3%) of the genes were expressed in only one or two of the strains when grown under in vitro planktonic conditions. Two sequences (L001_PA_0026_A07 and 0031_E04) were expressed in 9 strains, but no single transcript was observed in 10 or more strains (Table 4).

Pairwise comparisons of the 11 clinical strains using the PCR results were performed for all 70 novel sequences under study (Table 5). The greatest differences in genic content were observed between the strain pairs Pitt E and M18858 and Pitt E and W28869, both of which differed at 42 loci. The smallest degree of difference was observed between strains W27912 and W27931, which differed at only six loci. The Pitt E strain, which has the greatest number of the novel sequences, displays the greatest number of differences (315) when compared against all other strains. Conversely, the Pitt A strain has the fewest genic differences (191) when compared against all other strains (Table 6). A similar comparison using the RT-PCR results demonstrated that Pitt E and W28869 have the largest number (39) of gene expression differences. Interestingly, strains M18851 and M18858 had identical expression patterns for the 70 novel sequences. Overall, Pitt E and W28869 are the two strains most different from other strains, which is consistent with the fact that these two strains have the highest number (eight each) of unique DNA sequences.

TABLE 5. Pairwise tabulations of unique gene content and expression among 11 *P. aeruginosa* clinical isolates

Strain	No. of genic and expression differences between strain pairs based on 70 unique (non-PAO1) sequences ^a										
	Pitt A	Pitt C	Pitt D	Pitt E	Pitt F	27853	M18851	M18858	W27912	W27931	W28869
Pitt A		14	18	28	18	12	17	22	19	21	22
Pitt C	14		18	30	24	18	25	28	29	25	26
Pitt D	18	22		32	34	22	33	32	25	21	28
Pitt E	24	32	30		30	18	33	42	31	29	42
Pitt F	14	14	28	32		20	17	14	25	29	36
27853	13	3	21	29	13		23	30	19	21	30
M18851	13	13	25	35	11	12		11	24	26	29
M18858	13	13	25	35	11	12	0		25	23	32
W27912	24	26	16	30	30	25	25	25		6	31
W27931	26	28	18	32	32	27	27	27	2		31
W28869	21	21	25	39	25	20	22	22	29	31	

^a Numbers in bold refer to the number of different genes between each pair of strains, and numbers that are not bold refer to the number of unique genes that are differentially expressed between members of a strain pair. In all cases, the number of differences is over a denominator of 70.

TABLE 6. Summary of pairwise comparisons for each strain based on unique gene content and transcripts among 11 *P. aeruginosa* clinical isolates

Strain	No. of differences in pairwise comparisons of ^a :	
	Novel gene sequences	Novel RNA transcripts
Pitt A	191	180
Pitt C	237	186
Pitt D	263	228
Pitt E	315	318
Pitt F	247	210
27853	213	175
M18851	238	183
M18858	259	183
W27912	234	232
W27931	232	250
W28869	307	255

^a Results shown are for summed pairwise comparisons of the indicated strain with the other 10 clinical strains studied.

The advent of bacterial genomic sequencing has provided a wealth of information regarding the extent of intraspecies genetic diversity. The documentation of very extensive genomic plasticity among professional bacterial pathogens reflects the fact that these organisms possess sophisticated mechanisms for horizontal gene transfer and that they live in a highly selective environment driven by the host's adaptive immune response (44, 45). In the current study we wished to determine whether an opportunistic pathogen such as *P. aeruginosa* displayed the same high level of interstrain genomic plasticity; thus, we embarked on a study to compare a group of unrelated clinical isolates to the reference strain, PAO1. The DGH states that chronic pathogenic processes are polyclonal and that there is a population-based supragenome that is substantially larger than the genome size of any of the component strains. The DGH also states that no two strains would be identical in terms of their genic content and that through HGT processes new strains with unique genic complements are continually being created by recombining the unique genic characters of the parent strains. To test this hypothesis, we constructed a highly redundant pooled genomic library prepared from the DNA of 12 *P. aeruginosa* clinical strains and surveyed approximately 1.5% of the functional library. These studies identified ~11% of the clones as being novel with respect to the PAO1 genome, i.e., a BLASTn search revealed no identifiable homologs. A BLASTx search of 117 of these novel sequences employing hypothetical translations of the ORFs within these clones (see Supplementary Table 2 at the CGS website) revealed that 4 had their greatest similarity to PAO1 genes, 17 were most similar to known or putative phage genes, 29 showed their greatest similarity to genes from other *P. aeruginosa* strains, and another 28 had their greatest similarity to various other species within the pseudomonas family, including *Ralstonia*, *Burkholderia*, *Azotobacter*, and *Rhodopseudomonas* species among others. This last finding suggests that there are not significant barriers to HGT among the pseudomonads generally. Even more surprising is the finding that 39 of the 117 novel genes had their greatest similarity to genes from a very wide range of organisms, including archaea and eukaryotes. This of course does not mean that these genes had their origins in these species identified,

only that they likely did not have their origin within the pseudomonads.

Spencer et al. surveyed three sequenced *P. aeruginosa* genomes (0.5-fold coverage) and found that 11.6% of their high-quality sequence was unique with respect to the PAO1 genome (53). Taken together with our findings, it would appear that globally *P. aeruginosa* strains vary by ~11% on average. The genic distribution of the 70 selected sequences examined in the current study demonstrate the nonuniform distribution of genes within the population-based supragenome. These findings support the virulence corollary to the DGH that says chronic pathogens use gene distribution, recombination, and natural selection during infection as supravirulence factors that contribute to persistence in the face of host defense mechanisms. As further support of this hypothesis, we identified a number of novel sequences that, based upon protein similarities, would appear to encode highly variant forms of known virulence factors in *P. aeruginosa* and other pathogenic bacterial species. Nineteen of the 70 novel sequences were identified exclusively in only 1 of the 11 strains; therefore, these genes may be useful as strain-specific markers in epidemiological investigations and for studying gene flow in polyclonal animal model studies.

We previously presented evidence that extensive genomic plasticity exists among both gram-negative and gram-positive obligate respiratory pathogens (44, 45), and in the current study we have extended these findings to include opportunistic pathogens. Comparing the genomic plasticity characteristics among clinical strains of *P. aeruginosa*, *S. pneumoniae*, and *H. influenzae*, we found 48.5% of the novel *P. aeruginosa* sequences present in only one or two strains; this was in contrast to 37.9% for *S. pneumoniae* and only 3.8% for *H. influenzae*. Thus, *P. aeruginosa*, despite the fact that it is not naturally competent, displays the greatest genomic plasticity and the highest number of unfixed genes. This suggests that other HGT mechanisms must also play important roles in diversity generation. *P. aeruginosa* is a robust biofilm former and is associated with numerous medical biofilms, including those found with otorrhea, chronic obstructive pulmonary disease, and CF. *P. aeruginosa* biofilms contain high concentrations of DNA within their extracellular matrices (59), and this observation combined with the fact that the levels of HGT in biofilms have been demonstrated to be several orders of magnitude higher than in planktonic cells of the same strain make it likely that much of the strain evolution that occurs in vivo takes place during the biofilm phase of the bacterial life cycle. It is interesting to speculate that *P. aeruginosa*, which can form pili or pseudopili (13, 21), may be able to directly take up DNA from within the biofilm environment. Regardless of the mechanism of diversity generation, it is clear that *P. aeruginosa* is genomically highly plastic. The data presented here will become a base for future in vivo polyclonal animal studies for testing the DGH.

ACKNOWLEDGMENTS

This work was supported by Allegheny General Hospital, Allegheny-Singer Research Institute, and National Institute on Deafness and Other Communication Disorders grants DC 04173 (G.D.E.) and DC 02148 (G.D.E.).

REFERENCES

- Akhverdian, V. Z., E. A. Khrenova, V. G. Bogush, T. V. Gerasimova, and N. B. Kirsanov. 1984. Wide distribution of transposable phages in natural *Pseudomonas aeruginosa* populations. *Genetika* **20**:1612–1619.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403–410.
- Aoki, K. 1926. Agglutinatorische einteilung von Pyocyaneusbazillen welche bei verschiedenen menschenerkrankungen nach gewiesen wurden. *Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg. Abt. 1 Orig.* **98**:186–195.
- Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. 1990. Current protocols in molecular biology. Greene Publishing Associates and Wiley-Interscience, New York, N.Y.
- Bick, J. A., J. J. Dennis, G. J. Zylstra, J. Nowack, and T. Leustek. 2000. Identification of a new class of 5'-adenylylsulfate (APS) reductases from sulfate-assimilating bacteria. *J. Bacteriol.* **182**:135–142.
- Boles, B. R., M. Thoendel, and P. K. Singh. 2004. Self-generated diversity produces "insurance effects" in biofilm communities. *Proc. Natl. Acad. Sci. USA* **101**:16630–16635.
- Canosa, I., J. M. Sanchez-Romero, L. Yuste, and F. Rojo. 2000. A positive feedback mechanism controls expression of AlkS, the transcriptional regulator of the *Pseudomonas oleovorans* alkane degradation pathway. *Mol. Microbiol.* **35**:791–799.
- Carroll, K. S., H. Gao, H. Chen, C. D. Stout, J. A. Leary, and C. R. Bertozzi. 2005. A conserved mechanism for sulfonucleotide reduction. *PLoS Biol.* **3**:1418–1435.
- Characklis, W. G., R. Bakke, and M. H. Turakhia. 1988. Theoretical and experimental analysis of a *Pseudomonas aeruginosa* biofilm. *Water Sci. Technol.* **20**:45–51.
- Chuang, S. E., D. L. Daniels, and F. R. Blattner. 1993. Global regulation of gene expression in *Escherichia coli*. *J. Bacteriol.* **175**:2026–2036.
- Costerton, J. W., Z. Lewandowski, D. E. Caldwell, D. R. Korber, and H. M. Lappin-Scott. 1995. Microbial biofilms. *Annu. Rev. Microbiol.* **49**:711–745.
- Ding, H., and B. Dimple. 2000. Direct nitric oxide signal transduction via nitrosylation of iron-sulfur centers in the SoxR transcription activator. *Proc. Natl. Acad. Sci. USA* **97**:5146–5150.
- Durand, E., A. Bernadac, G. Ball, A. Lazdunski, J. N. Sturgis, and A. Filloux. 2003. Type II protein secretion in *Pseudomonas aeruginosa*: the pseudopilus is a multifibrillar and adhesive structure. *J. Bacteriol.* **185**:2749–2758.
- Ehrlich, G. D., R. Veeh, X. Wang, J. W. Costerton, J. D. Hayes, F. Z. Hu, B. J. Daigle, M. D. Ehrlich, and J. C. Post. 2002. Mucosal biofilm formation on middle-ear mucosa in the chinchilla model of otitis media. *JAMA* **287**:1710–1715.
- Ehrlich, G. D., F. Z. Hu, K. Shen, P. Stoodley, and J. C. Post. 2005. Bacterial plurality as a general mechanism driving persistence in chronic infections. *Clin. Orthop. Relat. Res.* **437**:20–24.
- Erdos, G., S. Sayeed, P. Antalis, F. Z. Hu, J. Hayes, J. Goodwin, R. Dopico, J. C. Post, and G. D. Ehrlich. 2003. Development and characterization of a pooled *Haemophilus influenzae* genomic library for the evaluation of gene expression changes associated with mucosal biofilm formation in otitis media. *Int. J. Pediatr. Otorhinolaryngol.* **67**:749–755.
- Farjo, R. S., B. Foxman, M. J. Patel, L. Zhang, M. M. Pettigrew, S. I. McCoy, C. F. Marrs, and J. R. Gilsdorf. 2004. Diversity and sharing of *Haemophilus influenzae* strains colonizing healthy children attending day-care centers. *Pediatr. Infect. Dis. J.* **23**:41–46.
- Fiorillo, L., M. Zucker, D. Sawyer, and A. N. Lin. 2001. The pseudomonas hot-foot syndrome. *N. Engl. J. Med.* **345**:335–338.
- Guzzo, J., F. Duong, C. Wandersman, M. Murgier, and A. Lazdunski. 1991. The secretion genes of *Pseudomonas aeruginosa* alkaline protease are functionally related to those of *Erwinia chrysanthemi* proteases and *Escherichia coli* alpha-haemolysin. *Mol. Microbiol.* **5**:447–453.
- Ha, U., and S. Jin. 1999. Expression of the *soxR* gene of *Pseudomonas aeruginosa* is inducible during infection of burn wounds in mice and is required to cause efficient bacteremia. *Infect. Immun.* **67**:5324–5331.
- Keizer, D. W., C. M. Slupsky, M. Kalisiak, A. P. Campbell, M. P. Crump, P. A. Sastry, B. Hazes, R. T. Irvin, and B. D. Sykes. 2001. Structure of a pilin monomer from *Pseudomonas aeruginosa*: implications for the assembly of pili. *J. Biol. Chem.* **276**:24186–24193.
- Kiley, P. J., and H. Beinert. 2003. The role of Fe-S proteins in sensing and regulation in bacteria. *Curr. Opin. Microbiol.* **6**:181–185.
- Kirisits, M. J., L. Prost, M. Starkey, and M. R. Parsek. 2005. Characterization of colony morphology variants isolated from *Pseudomonas aeruginosa* biofilms. *Appl. Environ. Microbiol.* **71**:4809–4821.
- Klockgether, J., O. Reva, K. Larbig, and B. Tummeler. 2004. Sequence analysis of the mobile genome island pKLC102 of *Pseudomonas aeruginosa* C. *J. Bacteriol.* **186**:518–534.
- Kobayashi, K., and S. Tagawa. 2004. Activation of SoxR-dependent transcription in *Pseudomonas aeruginosa*. *J. Biochem.* **136**:607–615.
- Kopriva, S., T. Buchert, G. Fritz, M. Suter, R. Benda, V. Schunemann, A. Koprivova, P. Schurmann, A. X. Trautwein, P. M. Kroneck, and C. Brunold. 2002. The presence of an iron-sulfur cluster in adenosine 5'-phosphosulfate reductase separates organisms utilizing adenosine 5'-phosphosulfate and phosphoadenosine 5'-phosphosulfate for sulfate assimilation. *J. Biol. Chem.* **277**:21786–21791.
- Kropinski, A. M. 2000. Sequence of the genome of the temperate, serotype-converting, *Pseudomonas aeruginosa* bacteriophage D3. *J. Bacteriol.* **182**:6066–6074.
- Kuroda, T., N. Fujita, J. Utsugi, M. Kuroda, T. Mizushima, and T. Tsuchiya. 2004. A major Li⁺ extrusion system NhaB of *Pseudomonas aeruginosa*: comparison with the major Na⁺ extrusion system NhaP. *Microbiol. Immunol.* **48**:243–250.
- Landman, D., J. M. Quale, D. Mayorga, A. Adedeji, K. Vangala, J. Ravishankar, C. Flores, and S. Brooks. 2002. Citywide clonal outbreak of multi-resistant *Acinetobacter baumannii* and *Pseudomonas aeruginosa* in Brooklyn, N.Y.: the preantibiotic era has returned. *Arch. Intern. Med.* **162**:1515–1520.
- Liu, P. V., H. Matsumoto, H. Kusama, and T. Bergan. 1983. Survey of heat-stable, major somatic antigens of *pseudomonas aeruginosa*. *Int. J. Syst. Bacteriol.* **33**:256–264.
- Liu, P. V., and S. Wang. 1990. Three new major somatic antigens of *Pseudomonas aeruginosa*. *J. Clin. Microbiol.* **28**:922–925.
- Lyczak, J. B., C. L. Cannon, and G. B. Pier. 2000. Establishment of *Pseudomonas aeruginosa* infection: lessons from a versatile opportunist. *Microbes Infect.* **2**:1051–1060.
- Murphy, T. F., S. Sethi, K. L. Klingman, A. B. Brueggemann, and G. V. Doern. 1999. Simultaneous respiratory tract colonization by multiple strains of nontypeable *Haemophilus influenzae* in chronic obstructive pulmonary disease: implications for antibiotic therapy. *J. Infect. Dis.* **180**:404–409.
- Newton, G. J., C. Daniels, L. L. Burrows, A. M. Kropinski, A. J. Clarke, and J. S. Lam. 2001. Three-component-mediated serotype conversion in *Pseudomonas aeruginosa* by bacteriophage D3. *Mol. Microbiol.* **39**:1237–1247.
- Padan, E., M. Venturi, Y. Gerchman, and N. Dover. 2001. Na⁺/H⁺ antiporters. *Biochim. Biophys. Acta* **1505**:144–157.
- Palma, M., J. Zurita, J. A. Ferreras, S. Worgall, D. H. Larone, L. Shi, F. Campagne, and L. E. Quadri. 2005. *Pseudomonas aeruginosa* SoxR does not conform to the archetypal paradigm for SoxR-dependent regulation of the bacterial oxidative stress adaptive response. *Infect. Immun.* **73**:2958–2966.
- Pimenta, A. L., J. Young, I. B. Holland, and M. A. Blight. 1999. Antibody analysis of the localisation, expression and stability of HlyD, the MFP component of the *E. coli* haemolysin translocator. *Mol. Gen. Genet.* **261**:122–132.
- Pirnay, J. P., D. De Vos, C. Cochez, F. Bilocq, A. Vanderkelen, M. Zizi, B. Ghysels, and P. Cornelis. 2002. *Pseudomonas aeruginosa* displays an epidemic population structure. *Environ. Microbiol.* **4**:898–911.
- Pomposiello, P. J., and B. Dimple. 2001. Redox-operated genetic switches: the SoxR and OxyR transcription factors. *Trends Biotechnol.* **19**:109–114.
- Purevdorj-Gage, B., J. W. Costerton, and P. Stoodley. 2005. Phenotypic differentiation and seeding dispersal in non-mucoid and mucoid *Pseudomonas aeruginosa* biofilms. *Microbiology* **151**:1569–1576.
- Rigden, D. J., M. J. Jedrzejewski, and M. Y. Galperin. 2003. Amidase domains from bacterial and phage autolysins define a family of gamma-D,L-glutamate-specific amidohydrolases. *Trends Biochem. Sci.* **28**:230–234.
- Ripp, S., O. A. Ogunseitan, and R. V. Miller. 1994. Transduction of a freshwater microbial community by a new *Pseudomonas aeruginosa* generalized transducing phage, UT1. *Mol. Ecol.* **3**:121–126.
- Sauer, K., A. K. Camper, G. D. Ehrlich, J. W. Costerton, and D. G. Davies. 2002. *Pseudomonas aeruginosa* displays multiple phenotypes during development as a biofilm. *J. Bacteriol.* **184**:1140–1154.
- Shen, K., P. Antalis, J. Gladitz, S. Sayeed, A. Ahmed, S. Yu, J. Hayes, S. Johnson, B. Dice, R. Dopico, R. Keefe, B. Janto, W. Chong, J. Goodwin, R. M. Wadowsky, G. Erdos, J. C. Post, G. D. Ehrlich, and F. Z. Hu. 2005. Identification, distribution, and expression of novel genes in 10 clinical isolates of nontypeable *Haemophilus influenzae*. *Infect. Immun.* **73**:3479–3491.
- Shen, K., J. Gladitz, P. Antalis, B. Dice, B. Janto, R. Keefe, J. Hayes, A. Ahmed, R. Dopico, N. Ehrlich, J. Jocz, L. Kropp, S. Yu, L. Nistico, D. P. Greenberg, K. Barbadora, R. A. Preston, J. C. Post, G. D. Ehrlich, and F. Z. Hu. 2006. Characterization, distribution and expression of novel genes among eight clinical isolates of *Streptococcus pneumoniae*. *Infect. Immun.* **74**:321–330.
- Shen, K., X. Wang, J. C. Post, and G. D. Ehrlich. 2003. Molecular and translational research approaches for the study of bacterial pathogenesis in otitis media. p. 91–119. *In* R. M. Rosenfeld and C. D. Bluestone (ed.), Evidence-based otitis media, 2nd ed. BC Decker Inc., Hamilton, Ontario, Canada.
- Smith-Vaughan, H. C., A. J. Leach, T. M. Shelby-James, K. Kemp, D. J. Kemp, and J. D. Mathews. 1996. Carriage of multiple ribotypes of non-encapsulated *Haemophilus influenzae* in aboriginal infants with otitis media. *Epidemiol. Infect.* **116**:177–183.
- Smith-Vaughan, H. C., K. S. Sriprakash, J. D. Mathews, and D. J. Kemp. 1995. Long PCR-ribotyping of nontypeable *Haemophilus influenzae*. *J. Clin. Microbiol.* **33**:1192–1195.
- Smith-Vaughan, H. C., K. S. Sriprakash, J. D. Mathews, and D. J. Kemp. 1997. Nonencapsulated *Haemophilus influenzae* in aboriginal infants with otitis media: prolonged carriage of P2 porin variants and evidence for horizontal P2 gene transfer. *Infect. Immun.* **65**:1468–1474.

50. Smits, T. H., M. Rothlisberger, B. Witholt, and J. B. van Beilen. 1999. Molecular screening for alkane hydroxylase genes in gram-negative and gram-positive strains. *Environ. Microbiol.* **1**:307–317.
51. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**:503–517.
52. Speert, D. P. 2002. Molecular epidemiology of *Pseudomonas aeruginosa*. *Front. Biosci.* **7**:e354–e361.
53. Spencer, D. H., A. Kas, E. E. Smith, C. K. Raymond, E. H. Sims, M. Hastings, J. L. Burns, R. Kaul, and M. V. Olson. 2003. Whole-genome sequence variation among multiple isolates of *Pseudomonas aeruginosa*. *J. Bacteriol.* **185**:1316–1325.
54. Stover, C. K., X. Q. Pham, A. L. Erwin, S. D. Mizoguchi, P. Warrenner, M. J. Hickey, F. S. Brinkman, W. O. Hufnagle, D. J. Kowalik, M. Lagrou, R. L. Garber, L. Goltry, E. Tolentino, S. Westbrook-Wadman, Y. Yuan, L. L. Brody, S. N. Coulter, K. R. Folger, A. Kas, K. Larbig, R. Lim, K. Smith, D. Spencer, G. K. Wong, Z. Wu, I. T. Paulsen, J. Reizer, M. H. Saier, R. E. Hancock, S. Lory, and M. V. Olson. 2000. Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature* **406**:959–964.
55. van Beilen, J. B., S. Panke, S. Lucchini, A. G. Franchini, M. Rothlisberger, and B. Witholt. 2001. Analysis of *Pseudomonas putida* alkane-degradation gene clusters and flanking insertion sequences: evolution and regulation of the alk genes. *Microbiology* **147**:1621–1630.
56. van Beilen, J. B., M. M. Marin, T. H. Smits, M. Rothlisberger, A. G. Franchini, B. Witholt, and F. Rojo. 2004. Characterization of two alkane hydroxylase genes from the marine hydrocarbonoclastic bacterium *Alcanivorax borkumensis*. *Environ. Microbiol.* **6**:264–273.
57. Wang, P. W., L. Chu, and D. S. Guttman. 2004. Complete sequence and evolutionary genomic analysis of the *Pseudomonas aeruginosa* transposable bacteriophage D3112. *J. Bacteriol.* **186**:400–410.
58. Waterer, G. W., and R. G. Wunderink. 2001. Increasing threat of gram-negative bacteria. *Crit. Care Med.* **29**:N75–N81.
59. Whitchurch, C. B., T. Tolker-Nielsen, P. C. Ragas, and J. S. Mattick. 2002. Extracellular DNA required for bacterial biofilm formation. *Science* **295**:1487.
60. Whyte, L. G., T. H. Smits, D. Labbe, B. Witholt, C. W. Greer, and J. B. van Beilen. 2002. Gene cloning and characterization of multiple alkane hydroxylase systems in *Rhodococcus* strains Q15 and NRRL B-16531. *Appl. Environ. Microbiol.* **68**:5933–5942.
61. Wilson, S., M. A. Hamilton, G. C. Hamilton, M. R. Schumann, and P. Stoodley. 2004. Statistical quantification of the detachment rates and size distribution of cell clumps from wild-type (PAO1) and cell signaling mutant (JP1) *Pseudomonas aeruginosa* biofilms. *Appl. Environ. Microbiol.* **70**:5847–5852.

Editor: J. T. Barbieri