Genome-wide recombination drives diversification of epidemic strains of *Acinetobacter baumannii*

Evan S. Snitkin[a], Adrian M. Zelazny[b], Clemente I. Montero[b], Frida Stock[b], Lilia Mijares[a,b], NISC Comparative Sequence Program[c,1], Patrick R. Murray[b,2,3], and Julie A. Segre[a,2]

[a]Epithelial Biology Section, Genetics and Molecular Biology Branch, National Human Genome Research Institute, Bethesda, MD 20892; [b]Department of Laboratory Medicine, National Institutes of Health Clinical Center, National Institutes of Health, Bethesda, MD 20892; and [c]National Institutes of Health Intramural Sequencing Center, National Institutes of Health, Rockville, MD 20852

Edited by Eric S. Lander, Broad Institute of MIT and Harvard, Cambridge, MA, and approved July 15, 2011 (received for review March 18, 2011)

Acinetobacter baumannii **is an emerging human pathogen and a significant cause of nosocomial infections among hospital patients worldwide. The enormous increase in multidrug resistance among hospital isolates and the recent emergence of pan-drug–resistant strains underscores the urgency to understand how *A. baumannii* evolves in hospital environments. To this end, we undertook a genomic study of a polyclonal outbreak of multidrug-resistant *A. baumannii* at the research-based National Institutes of Health Clinical Center. Comparing the complete genome sequences of the three dominant outbreak strain types enabled us to conclude that, despite all belonging to the same epidemic lineage, the three strains diverged before their arrival at the National Institutes of Health. The simultaneous presence of three divergent strains from this lineage supports its increasing prevalence in international hospitals and suggests an ongoing adaptation to the hospital environment. Further genomic comparisons uncovered that much of the diversification that occurred since the divergence of the three outbreak strains was mediated by homologous recombination across 20% of their genomes. Inspection of recombinant regions revealed that several regions were associated with either the loss or swapping out of genes encoding proteins that are exposed to the cell surface or that synthesize cell-surface molecules. Extending our analysis to a larger set of international clinical isolates revealed a previously unappreciated ability of *A. baumannii* to vary surface molecules through horizontal gene transfer, with subsequent intraspecies dissemination by homologous recombination. These findings have immediate implications in surveillance, prevention, and treatment of *A. baumannii* infections.**

bacterial evolution | microbial genomics | clinical microbiology | hospital epidemiology | hospital infection

In the last three decades, the rates of hospital-acquired infections with *Acinetobacter baumannii* have increased sharply (1). Of even greater cause for concern are rapidly increasing rates of multidrug resistance among *Acinetobacter* isolates, with a recent study of the US hospital system noting a jump in the rate of multidrug resistance from 7% to 30% between 1993 and 2004 (2). The capacity of *A. baumannii* to evolve resistance to antibiotics has recently culminated in the emergence of pan-drug–resistant strains (3), which threatens to render *A. baumannii* infections virtually untreatable. In addition to drug resistance, two other factors have largely contributed to the emergence of *A. baumannii* as a significant hospital pathogen (4, 5). First, *A. baumannii* is resistant to desiccation (6), facilitating spread via hospital personnel, infrastructure, and medical devices (7). Second, in addition to causing prolonged infections, *A. baumannii* can stably reside within a human host in the form of extended asymptomatic patient colonization (8), enabling its spread both within and between hospitals (9).

Because of its capacity to survive and spread within the hospital environment, *A. baumannii* has a propensity for causing hospital outbreaks, wherein there is a sudden spike of infections for a prolonged period (7). To track the origin and progression of a given outbreak, molecular typing methods capable of distinguishing between divergent strains of *A. baumannii* are applied, typically retrospectively (10). Earlier studies found that most outbreaks were caused by one or two dominant strains spreading within and between institutions (7). More recently, there have been several reports of outbreaks with either the simultaneous or sequential presence of multiple dominant strains (11, 12). It is unclear whether these polyclonal outbreaks are a consequence of a rapid divergence of *A. baumannii* during an outbreak resulting in multiple distinct clones or of the increasing prevalence of *A. baumannii*, increasing the likelihood of multiple clones simultaneously coming to the same institution.

Molecular typing studies analyzing outbreaks across different hospitals have demonstrated that a small number of epidemic strains are responsible for a disproportionate number of outbreaks in geographically scattered hospitals (13–15). Among the most widespread are three particular lineages, commonly referred to as European clonal (EC) lineages I, II, and III (16, 17). To gain insight into the success of these lineages, groups have sequenced the genomes of select representatives (18–21). Among the most notable features of these *A. baumannii* genomes was an observed propensity for horizontal gene transfer, with the most striking examples being large genomic islands enriched in transposable elements and antibiotic resistance determinants (21).

Here we apply whole-genome sequencing to clinical isolates of *A. baumannii* to gain insight into the origin of a polyclonal outbreak and thus develop an understanding of how epidemic lineages of *A. baumannii* evolve in the hospital environment. Analyzing the genome sequences of representatives of the three major multidrug-resistant (MDR) strain types allowed us to dis-

The members of the NISC Comparative Sequencing Program are Jim Mullikin, Robert Blakesley, Alice Young, Grace Chu, Colleen Ramsahoye, Sean Lovett, Joel Han, Richelle Legaspi, Christina Sison, Michael Gregory, Casandra Montemayor, Marie Gestole, April Hargrove, Taccara Johnson, Jerlil Myrick, Nancy Riebow, Brian Schmidt, Betsy Novotny, Jyoti Gupta, Betty Benjamin, Shelise Brooks, Holly Coleman, Shi-ling Ho, Karen Schandler, Lauren Smith, Mal Stantripop, Quino Maduro, Gerry Bouffard, Mila Dekhtyar, Xiaobin Guan, Cathy Masiello, Baishali Maskeri, Jenny McDowell, Morgan Park, and Pamela Thomas.

Author contributions: E.S.S., A.M.Z., P.R.M., and J.A.S. designed research; E.S.S., A.M.Z., C.I.M., F.S., L.M., and N.C.S.P. performed research; E.S.S., A.M.Z., and J.A.S. contributed new reagents/analytic tools; E.S.S. and P.R.M. analyzed data; and E.S.S., P.R.M., and J.A.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under accession nos. AFSZ00000000, AFTA00000000, AFTB00000000, AFTC00000000. The sequences can also be found under BioProject (Genome Project) (accession nos. 63335, 63337, 63339, 63341).

[1]A complete list of the investigators of the NISC Comparative Sequencing Program is available at http://nisc.nih.gov/open_page.cgi?path=/staff.html.

[2]To whom correspondence may be addressed. E-mail: Patrick_Murray@bd.com or jsegre@nhgri.nih.gov.

[3]Present address: BD Diagnostics, Sparks, MD 21152.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1104404108/-/DCSupplemental.

cern that two events likely contributed to the polyclonal outbreak: a single patient bringing two strains and the other strain evolving from a non-MDR strain present in the hospital more than a year before the outbreak. Additional comparisons of the genomes of the outbreak strains revealed evidence of extensive homologous recombination since their divergence, which had facilitated variation in gene sets contributing to the collection of *A. baumannii* cell-surface antigens. In total, our findings demonstrate the capacity of *A. baumannii* to rapidly alter its genome to evade both natural host defenses and antibiotic intervention.

Results

Molecular Typing Demonstrates Genetic Diversity Among Outbreak Isolates. Between May and December of 2007, *A. baumannii* was isolated from 45 patients at the research-based National Institutes of Health Clinical Center (NIHCC). Of these 45 patients, 29 harbored MDR *A. baumannii* (MDRAB), with 18 infected and 11 subclinically colonized. Pulsed-field gel electrophoresis (*Materials and Methods*) typing of the 29 MDRAB patients identified three distinct pulsotypes during the outbreak (*SI Appendix*, Fig. S1). Type A dominated early, type B dominated late, and type C was observed midoutbreak (Fig. 1). Multilocus sequence typing (MLST) (*Materials and Methods*) across seven regions revealed that A, B, and C were nearly identical to one another (*SI Appendix*, Table S1) and that all three were members of the EC II lineage.

The co-occurrence of three distinct strains from the same clonal lineage at the NIHCC may have two explanations. First, a single founder strain may have undergone rapid genomic changes that altered the pulsed-field banding pattern. Alternatively, two or three founder strains could have simultaneously come to the NIHCC. Although the second scenario seems less likely, the EC II lineage has emerged as the most prevalent epidemic lineage of *A. baumannii* (22), increasing the likelihood that three distinct members of this lineage simultaneously came to the NIHCC.

Whole-Genome Sequencing Suggests That Multiple Events Contributed to the Polyclonal Outbreak. To explore their history and evolution, we generated more than 15-fold unidirectional and paired-end genome sequences with the Roche 454 XLR platform for representatives of each of the three observed MDRAB pulsotypes: A, B, and C (*Materials and Methods*). Summary statistics pertaining to sequencing, assembly, and annotation of the draft genomes are provided in *SI Appendix*, Table S2. To determine whether the polyclonal outbreak was caused by one or multiple founders, we first determined the sequence variation among the genomes of A, B, and C by identifying all high-confidence SNPs (*SI Appendix*, Table S3) by using a series of stringent filtering criteria (*Materials and Methods*). To provide context to the sequence comparison of the outbreak genomes, we included strain ACICU, a member of the EC II lineage isolated from an intensive care unit outbreak in Rome in 2005, whose genome had been previously sequenced (20).

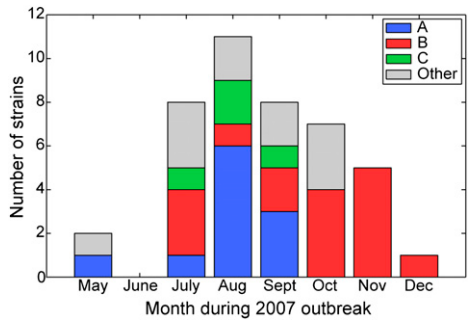

Fig. 1. Occurrence of different *A. baumannii* strain types during the NIHCC polyclonal outbreak. Isolates from patients were classified with pulsed-field gel electrophoresis. Three MDR strain types were observed (A, blue; B, red; and C, green) as well as non-MDR strains (other, gray).

We next explored whether there is any evidence of homologous recombination in these *A. baumannii* clinical isolates. For phylogenetic analyses, SNPs acquired horizontally through homologous recombination are not independently informative of a lineage's history and therefore should be excluded. Although homologous recombination has been shown to be prevalent in other *Acinetobacter* species (23), its role in the evolution of *A. baumannii* clinical isolates has not been previously explored. To identify recombinant regions, we searched for evidence of the localized sequence variation associated with recombination events, where foreign DNA has replaced a homologous region of the host chromosome. Fig. 2 displays the spatial distribution of all SNPs among the A, B, and C genomes, relative to ACICU, and revealed several contiguous genomic regions of unusually high sequence divergence. Polymorphic sites were not uniformly distributed but occurred in discrete clusters, with 98% (13,235 of 13,493) of polymorphic sites occurring in 20% of the genome. Thus, most of the observed sequence differences among these EC II genomes are a consequence of recombinant DNA that has been acquired since their divergence.

Reconstructing the phylogeny of EC II strains based on SNPs in nonrecombinant regions does not support the outbreak strains having diverged from a single ancestor while at the NIHCC (Fig. 3). The observation that A is no more closely related to B and C than it is to ACICU lends weight to A's arrival at the NIHCC being independent of B and C. Furthermore, although the clustering of B and C on the tree supports their divergence from a common ancestor, the divergence of B and C occurring before their arrival at the NIHCC is supported by the fact that B and C were both first isolated from the same patient at the same time. This patient had spent time in several other hospitals, which is consistent with a scenario in which B and C diverged from an ancestral strain, either in the patient himself or in one of the hospitals in which he previously resided.

Recombination Events Facilitate Swapping of Genes That Encode Cell-Surface Proteins and Synthesize Cell-Surface Molecules. We next returned to the striking observation of recombinant DNA spanning large regions of the EC II genomes. To gain insight into the selective forces driving recombination in the EC II lineage, we focused on individual recombinant regions associated with the gain or loss of genes. We hypothesized that these variations in gene content were likely to have had major functional consequences and were therefore good candidates to have driven the fixation of the recombinant regions with which they are associated. For the purposes of this analysis, we defined individual recombinant regions as encompassing SNP clusters with a uniform allele distribution across the EC II genomes, although multiple overlapping recombination events may underlie the observed patterns. Fig. 2 highlights four of the recombinant regions associated with variation in gene content (specific variable genes are shown in *SI Appendix*, Table S4).

Region I shows evidence of multiple recombination events, with the presence of distinct alleles in all four genomes providing evidence that recombination has occurred in at least three of the strains since their divergence. This recombinant region encompasses the cluster of genes synthesizing the O-antigen component of lipopolysaccharide (LPS), and we observed that all four EC II strains have unique gene content at their O-antigen clusters (Fig. 4 and *SI Appendix*, Fig. S3). Thus, as has been observed in other species (24), recombination is facilitating the previously documented O-antigen switching in *A. baumannii* (19). Region II, which also shows evidence of recurrent recombination, is associated with the loss of a cluster of genes previously annotated as being involved in iron acquisition (25). Region III is a 45-kb

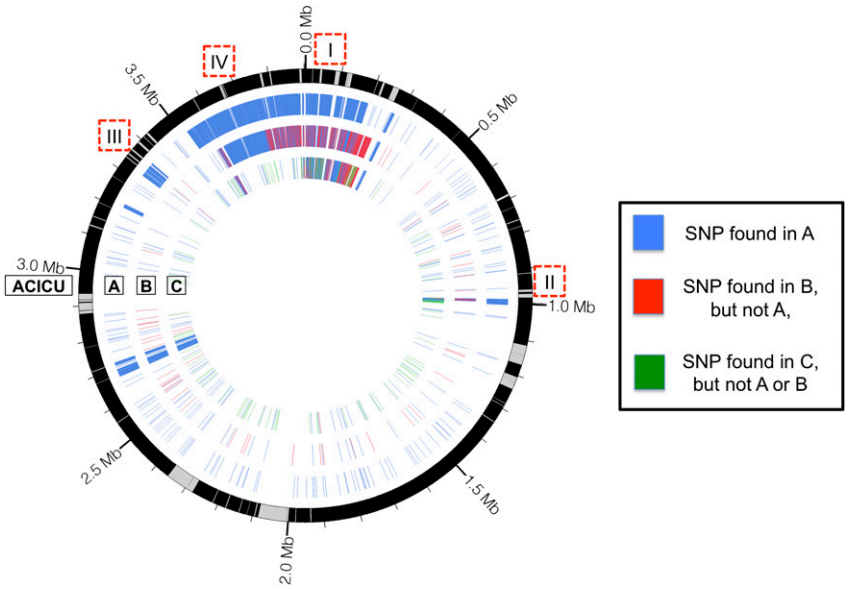

Fig. 2. Distribution of nucleotide differences among three outbreak strains with the genome of ACICU as reference. SNPs in the genome alignments of A, B, and C relative to ACICU were determined by using Mauve and displayed with Circos (39). Positions differing from ACICU (outer black circle) are indicated by colored marks in the three inner circles for A, B, and C. Different colors represent shared alleles among different sets of genomes. Blue marks represent a position in which A differs from ACICU, with a blue mark also occurring in the circles for B and/or C if either of them also share this variant. Red marks represent a position in which B differs from both A and ACICU, with a red mark also occurring in the circle for C if it shares this variant. Finally, green marks represent a position in which C differs from ACICU, A, and B. In the outer circle, gray regions represent sequence present in ACICU but not in all three of the other strains, which were therefore not considered for the SNP analysis. Such variable sequences include both truly variable regions and repetitive sequences that may be collapsed in the assemblies of A, B, and/or C.

section that is recombinant only in genome A and is associated with the swapping out of several genes, including LPS glycosyltransferases, for an alternate set of genes. Finally, region IV covers more than 100 kb, with A and B sharing one set of alleles across the region and C and ACICU sharing another set. Although region IV lacks explicit variation in gene content, there are two alternate blocks of highly divergent genes, including several putative outer membrane proteins and efflux pumps. Thus, in all four cases, recombination appears to have facilitated either the deletion (region II) or swapping out (regions I, II, and IV) of genes encoding proteins either exposed to the extracellular environment (e.g., membrane proteins) or involved in the synthesis of extracellular molecules (e.g., LPS). Building phylogenies based on these recombinant regions provided evidence that the immediate sources of recombinant material were other *A. baumannii* (*SI Appendix*, Fig. S4), which is consistent with previous work demonstrating a rapid decline in recombination efficiency with decreasing sequence similarity (23).

Alternative Genes Sets in EC II Genomes Show Dispersed Presence Across Distant *A. baumannii* Lineages. The observation of extensive recombination in the EC II lineage prompted us to search for evidence of recombination in other *A. baumannii* lineages. We first examined clustering of SNPs among three previously sequenced genomes in the EC I lineage (19, 21) and found several large SNP clusters (e.g., surrounding the O-antigen gene cluster) in addition to several smaller clusters (*SI Appendix*, Fig. S2). Strikingly, several of these clusters encompassed the sets of variable genes identified in the EC II genomes, with variation in gene content also occurring among the EC I genomes. Extending our analysis to more distantly related *Acinetobacter* genomes, we found that variation of these regions occurs across a large phylogenetic distance (Fig. 5 and *SI Appendix*, Fig. S3). Fig. 5 also shows that the number of alternative gene sets at different loci vary across the population, with the O-antigen cluster having many variants and the MDR pumps locus possibly having only two.

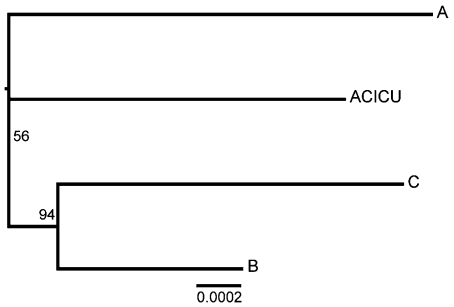

Fig. 3. Phylogeny of EC II strains based on SNPs in nonrecombinant regions. A maximum likelihood tree including the EC II strains, and rooted with EC I strain AB0057 (not shown for clarity), was constructed by using only polymorphisms that did not occur in the clusters observed in Fig. 2. Numbers on internal nodes are based on 100 bootstraps.

Monitoring O-Antigen Gene Clusters Leads to the Discovery of a Putative Non-MDR Ancestor of One of the Outbreak Strains. The observed diversity in recombinant regions across distantly related *A. baumannii* led us to wonder whether these regions were as unstable over the relatively short timescale of the outbreak. To this end, we examined the variability of O-antigen gene clusters in strains isolated from the NIHCC in the time leading up to and during the 2007 outbreak, by designing PCR assays specific for the O-antigen clusters of A, B, and C. We found that, among the strains isolated during the outbreak, the O-antigen cluster tracked with strain pulsotype, indicating that the O-antigen cluster was stable during the outbreak (*SI Appendix*, Table S5).

However, the non-MDR strain HC64, which was isolated more than a year before the outbreak began, tested positive for the O-antigen cluster of strain A despite being classified as a distinct pulsotype (*SI Appendix*, Fig. S1). The presence of the A O-antigen cluster in this non-MDR strain could be explained by (*i*) one of the two strains acquiring the O-antigen cluster from the other, (*ii*) both HC64 and A diverging from a common ancestor at the NIHCC, or (*iii*) HC64 being distantly related to A and coincidentally having the same O-antigen cluster. To distinguish these scenarios, we sequenced the complete genome of HC64. Comparison of the SNPs between HC64 and A, relative to ACICU, suggested a close relationship (*SI Appendix*, Fig. S5). This relationship was confirmed by incorporating HC64 into the phylogeny of EC II strains (*SI Appendix*, Fig. S6) and observing that A and HC64 clearly group together. Thus, the ancestor of one of the three major MDR outbreak strains was most likely present in the hospital more than a year before the outbreak.

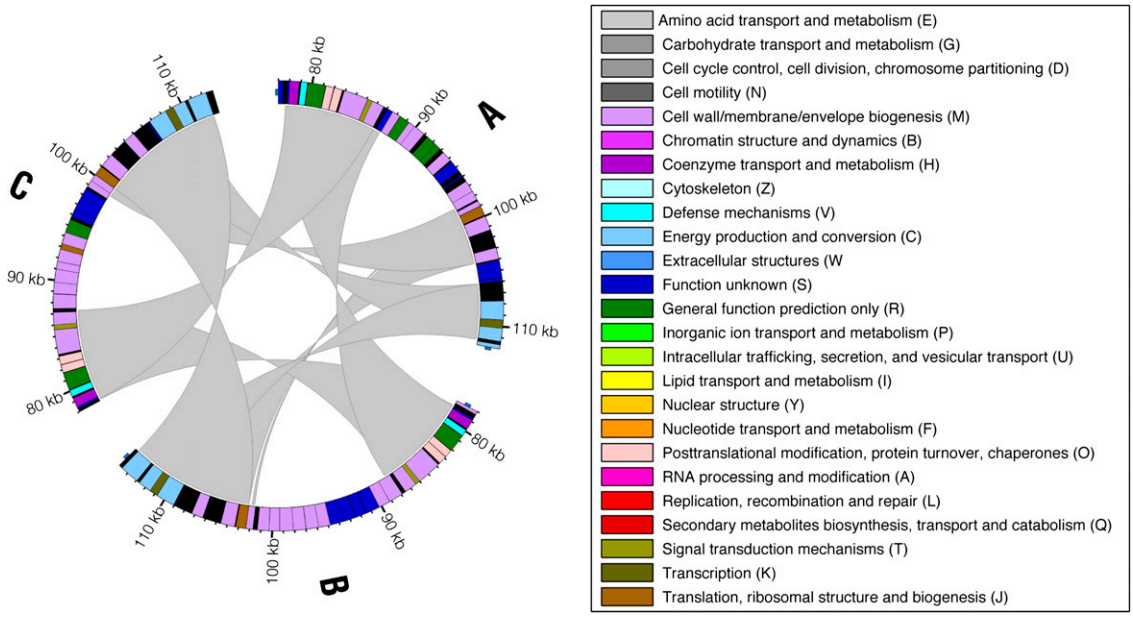

Fig. 4. Variable gene content in the O-antigen biosynthetic clusters of the three dominant outbreak strains. The ~30-kb regions surrounding the O-antigen biosynthetic gene clusters of strains A, B, and C are each represented by semicircles as marked. The colored bands on each semicircle indicate the Clusters of Orthologous Groups of Proteins (COG) functions of genes at the given position, with the COG/color relationship indicated in the key on the right. The gray ribbons represent sequences that can be aligned between a pair of genomes. The sequences flanking the O-antigen biosynthetic cluster align in all three genomes, but the biosynthetic clusters themselves contain unique genes.

Although A and HC64 are genetically very similar, they differ in the set of drugs to which they are resistant (*SI Appendix*, Table S6). Specifically, A is resistant to gentamicin, trimethoprim-sulfamethoxazole, and imipenem, whereas HC64 is susceptible to all three. Searching for genetic differences to account for the variable resistance, we found that A uniquely possessed a gentamicin 3′-*N*-acetyltransferase and a sulfonamide-resistant dihydropteroate synthase, which have been shown to provide resistance to gentamicin and sulfamethoxazole, respectively (5). Both of these genes reside in the same region of genome A, interspersed with mobile genetic elements.

Discussion

We applied whole-genome sequencing to elucidate the basis for a polyclonal outbreak at a single institution. We found the comparison of complete genome sequences to provide essential information beyond that of standard molecular typing techniques that allowed us to conclude that, despite being closely related, the three major MDRAB outbreak strains did not diverge during the course of the outbreak. Furthermore, whole-genome analysis enabled us to determine that one of the MDR outbreak strains originated from a non-MDR ancestor at the NIHCC, whereas the other two strains likely came to the NIHCC via a single patient. As whole-genome sequencing technologies continue to improve with respect to both price and speed, these approaches should become the gold standard for rigorous epidemiological analysis. In addition to this study, several recent studies have shown that epidemiological questions ranging from single-hospital outbreaks (26) to worldwide epidemics (27), and from retrospective analysis (28) to real-time monitoring (29), can be effectively addressed with a whole-genome approach.

Our results have additional relevance specific *to A. baumannii* epidemiology. First, our observation of three divergent members of the EC II lineage causing simultaneous outbreaks is further indication of the rising prevalence and continued diversification of this epidemic lineage. A second finding, important to future studies, was the observation of widespread recombination, which must be taken into account to effectively interpret both MLST and genomic data in the context of epidemiology. Furthermore, the observed positional bias of recombination events should be considered when selecting MLST regions (*SI Appendix*, Table S1). This bias is exemplified by comparing the outbreak strains using the two different sets of *A. baumannii* MLST regions, where only one of the two sets contains regions in recurrently recombinant regions (*SI Appendix*, *Supplementary Methods* and Table S1). Finally, our observation that recombination is facilitating gene content variation of specific antigenic loci, with a potential relevance to intrahospital transmission, suggests that typing of these loci could prove useful in outbreak control.

In addition to its relevance to epidemiology, recombination in *A. baumannii* may be important to virulence in the form of immune evasion. Among several large gene gain/loss events associated with recombinant regions, the common thread is the variability of genes that contributed to the repertoire of cell-surface molecules. Recombination-mediated variation in surface molecules has been previously observed in organisms including *Neisseria meningitidis* (30), *Streptococcus pneumoniae* (31), and *Escherichia coli* (32, 33). As has been hypothesized for these species, variation of cell-surface molecules in *A. baumannii* may be driven by evasion of the human immune system. In addition to facilitating repeat infection, this variation may contribute to the spread of *A. baumannii* by contributing to the propensity of *A. baumannii* for long-term patient colonization by allowing strains with novel surface markers to remain invisible to the host immune system.

Although immune evasion is a plausible driver of the observed variation in cell-surface molecules, an additional selective pressure may come in the form of phage predation. The genomes of EC II strains, and other sequenced *A. baumannii*, reveal a strong prophage presence (*SI Appendix*, *Supplementary Methods*). Many of the same molecules targeted by the immune system are exploited in phage attachment to cells, making it difficult to distinguish between these two possibilities (34). However, even if the immediate selective pressure for antigenic variation is phage

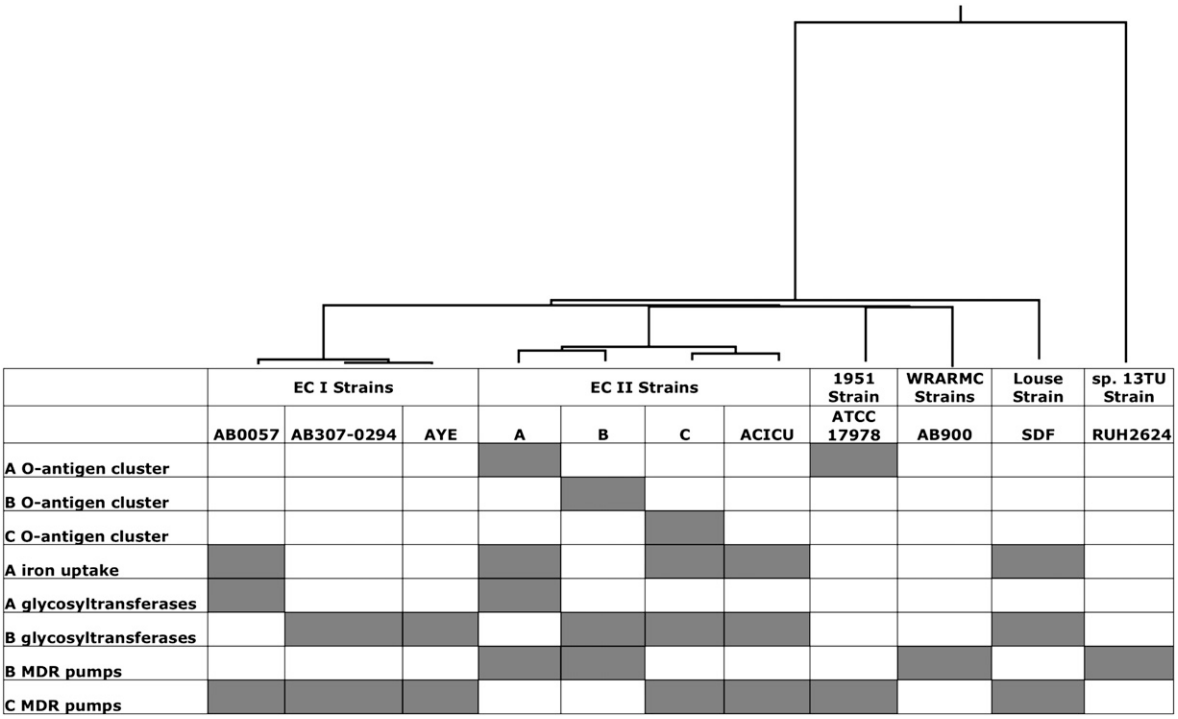

Fig. 5. Alignment of regions of variable gene content in EC II recombinant regions to other sequenced *Acinetobacter* genomes. Each column in the table represents a different *Acinetobacter* strain that has been sequenced either here (A, B, or C) or previously. Phylogeny among the strains (shown above the table) was determined based on SNPs in regions common to all genomes by using the neighbor-joining algorithm. Each row of the table corresponds to variable genes found in the recombinant regions (I, II, III, and IV in Fig. 2). A gray/white box at a given position indicates that the variable region, corresponding to the row, is present/absent in the genome, corresponding to the column. A region is considered present in a given genome if more than 80% of the region aligns to the genome, as determined by BLAST. The alignments are shown in *SI Appendix*, Fig. S3.

avoidance, the resulting variation may still promote the survival and spread of *A. baumannii* in human hosts.

Regardless of the drivers of diversification, the variability in surface markers underscores a large reservoir of genetic variation among *A. baumannii*. Many vaccines are designed to train the immune system to recognize these highly immunogenic cell-surface markers (35), but, if all common variants are not included, the vaccine will not be effective. Given the diversity of markers observed here, and the ease with which *A. baumannii* seems to assimilate DNA from both other *Acinetobacter* strains as well as members of different species (21), attempts to include all possible variants will be challenging (35).

Finally, our results put in a unique light the observation that a variety of strains are often observed during the course of a hospital outbreak. Frequently, only MDR strains are carefully monitored because they pose the greatest risk to patients. However, here we find two avenues by which non-MDR strains can potentially pose a threat to patients. First, non-MDR strains can in short order gain resistance determinants and become MDR. Second, non-MDR strains could serve as a potential reservoir of antigenic variants and contribute to the diversification of MDR strains. Thus, the dynamic genome of *A. baumannii* may dictate that the entire population of hospital strains be considered in attempts to control this emerging pathogen.

Materials and Methods

Summary of Outbreak and Strains Analyzed. From May to December 2007, the NIHCC experienced an outbreak of MDRAB, with no isolates of MDRAB previously recovered. Significant risk for colonization and infection was associated with time spent in the intensive care unit, mechanical ventilation, blood product transfusion, wound care, neutropenia, renal failure, and exposure to broad-spectrum antibiotics. Termination of the outbreak was achieved by strict infection-control procedures and targeted interventions. All isolates from this outbreak were identified by conventional biochemical tests and sequencing the 16S rRNA gene, tested against a broad panel of antimicrobial agents by broth microdilution, and frozen at –80 °C in *Brucella* broth with 20% glycerol.

Pulsed-Field Gel Electrophoresis. Isolates were grown overnight in brain–heart infusion broth at 35 °C in 5% CO_2, concentrated by centrifugation, and resuspended in 1 M NaCl, and 10 mM Tris·Cl (pH 7.6) to achieve approximately the turbidity of a 4.0 McFarland turbidity standard. Extraction of genomic DNA in 0.8% InCert agarose plugs (Cambrex) was performed by standard methods and included 2 mg of lysozyme per mL followed by proteinase K digestion (100 μg/mL) for bacterial lysis. DNA was digested with 125 U of the restriction endonuclease ApaI at 25 °C. Pulsed-field gel electrophoresis was performed in a CHEF-DR III (Bio-Rad Laboratories) apparatus with ramped pulse times from 7 to 21 s for 24 h at 6 V/cm. Size standards and an *A. baumannii* control strain were included in each run for gel-to-gel comparisons.

MLST. MLST was performed by using the protocols and PCR primers previously described (36). PCR reactions were carried out with illustra PuReTaq Ready-To-Go PCR Beads (GE Healthcare). Sequencing reactions were performed with the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit (PerkinElmer Applied Biosystems). Data analysis and sequence alignments were carried out with the Lasergene suite MegAlign software (DNASTAR). Sequence allele typing was performed with the multiple locus query tool at the publicly available *A. baumannii* MLST database (http://pubmlst.org/abaumannii/).

Genome Sequencing and Assembly. Unidirectional and paired-end libraries (2- to 3-kb fragments) were constructed by using the Roche 454 Titanium General Library Reagents Kit (Roche 454 Life Sciences). Sequencing primers were added, and each library was deposited onto a PicoTiterPlate and sequenced. Contig assembly of both unidirectional and paired-end reads was performed with gsAssembler v.2.3.

Genome Annotation. Genome annotation was performed with the National Center for Biotechnology Information (NCBI) Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP).

Whole-Genome Alignments. To facilitate whole-genome comparisons, the scaffolds and large contigs from each assembly were ordered and oriented relative to a reference and then stitched together to form a pseudochromosome. Ordering and orienting was done by using the Mauve contig mover (37), with the finished ACICU genome acting as a reference (20). Multiple alignments of pseudochromosomes and other complete genome sequences were performed with Mauve. In addition to the strains sequenced in this study, the following previously sequenced *Acinetobacter* genomes were included in certain alignments as indicated: ACICU (NCBI accession no. NC_010611), AB0057 (NCBI accession no. NC_011586), AYE (NCBI accession no. NC_010410), AB307-0294 (NCBI accession no. NC_011595), AB900 (NCBI accession no. NZ_ABXK00000000), ATCC 17978 (NCBI accession no. NC_009085), SDF (NCBI accession no. NC_010400), and sp. 13TU strain RUH2624 (NCBI accession no. NZ_ACQF00000000).

Detection and Filtering of Nucleotide Differences. An initial list of nucleotide differences was generated from the Mauve alignments by using the export SNPs function. This list of SNPs was then filtered to remove those SNPs that are likely to be caused by alignment or sequencing errors as detailed in *SI Appendix, Supplementary Methods*.

Phylogenetic Tree Construction. All trees were constructed based on polymorphic positions in the indicated set of genomes, excluding indel positions. Tree construction was performed by using either the maximum likelihood or neighbor-joining algorithms, as implemented in the Seaview package (38). Tree visualization was performed with FigTree version 1.3.1.

Delineation of Recombinant Regions. To designate recombinant region, each position in a genome alignment was first designated as SNP, non-SNP, or indel. Then regions of the alignment overrepresented in SNPs were determined by using a random walk algorithm as described in the *SI Appendix, Supplementary Methods*.

PCR to Amplify Region Specific to O-Antigen Locus of Strains A, B, and C. PCR was performed with the Takara kit from Takara Bio Inc. The primer sequences used to probe the O-antigen regions of A, B, and C and the associated PCR conditions are provided in *SI Appendix*, Table S7.

ACKNOWLEDGMENTS. We thank members of the J.A.S. laboratory and National Institutes of Health Intramural Sequencing Center (NISC) for critical reading of the manuscript; Jyoti Gupta for confirming SNPs; and Effie Nomicos for chart review. E.S.S. is supported by a Pharmacology Research Associate Training Fellowship, National Institute of General Medical Sciences. This work was supported by the National Human Genome Research Institute and NIHCC Intramural Research Programs.

1. Gaynes R, Edwards JR; National Nosocomial Infections Surveillance System (2005) Overview of nosocomial infections caused by gram-negative bacilli. *Clin Infect Dis* 41: 848–854.
2. Lockhart SR, et al. (2007) Antimicrobial resistance among Gram-negative bacilli causing infections in intensive care unit patients in the United States between 1993 and 2004. *J Clin Microbiol* 45:3352–3359.
3. Mahgoub S, Ahmed J, Glatt AE (2002) Completely resistant *Acinetobacter baumannii* strains. *Infect Control Hosp Epidemiol* 23:477–479.
4. Dijkshoorn L, Nemec A, Seifert H (2007) An increasing threat in hospitals: Multidrug-resistant *Acinetobacter baumannii*. *Nat Rev Microbiol* 5:939–951.
5. Peleg AY, Seifert H, Paterson DL (2008) *Acinetobacter baumannii*: Emergence of a successful pathogen. *Clin Microbiol Rev* 21:538–582.
6. Jawad A, Seifert H, Snelling AM, Heritage J, Hawkey PM (1998) Survival of *Acinetobacter baumannii* on dry surfaces: Comparison of outbreak and sporadic isolates. *J Clin Microbiol* 36:1938–1941.
7. Villegas MV, Hartstein AI (2003) Acinetobacter outbreaks, 1977–2000. *Infect Control Hosp Epidemiol* 24:284–295.
8. Marchaim D, et al. (2007) Surveillance cultures and duration of carriage of multidrug-resistant *Acinetobacter baumannii*. *J Clin Microbiol* 45:1551–1555.
9. Munoz-Price LS, Weinstein RA (2008) Acinetobacter infection. *N Engl J Med* 358: 1271–1281.
10. Runnegar N, et al. (2010) Molecular epidemiology of multidrug-resistant *Acinetobacter baumannii* in a single institution over a 10-year period. *J Clin Microbiol* 48: 4051–4056.
11. Marchaim D, et al. (2007) Molecular and epidemiologic study of polyclonal outbreaks of multidrug-resistant *Acinetobacter baumannii* infection in an Israeli hospital. *Infect Control Hosp Epidemiol* 28:945–950.
12. Villalón P, et al. (2011) Clonal diversity of nosocomial epidemic *Acinetobacter baumannii* strains isolated in Spain. *J Clin Microbiol* 49:875–882.
13. Hamouda A, Evans BA, Towner KJ, Amyes SGB (2010) Characterization of epidemiologically unrelated *Acinetobacter baumannii* isolates from four continents by use of multilocus sequence typing, pulsed-field gel electrophoresis, and sequence-based typing of $bla_{\text{OXA-51-like}}$ genes. *J Clin Microbiol* 48:2476–2483.
14. Turton JF, Gabriel SN, Valderrey C, Kaufmann ME, Pitt TL (2007) Use of sequence-based typing and multiplex PCR to identify clonal lineages of outbreak strains of *Acinetobacter baumannii*. *Clin Microbiol Infect* 13:807–815.
15. Diancourt L, Passet V, Nemec A, Dijkshoorn L, Brisse S (2010) The population structure of *Acinetobacter baumannii*: Expanding multiresistant clones from an ancestral susceptible genetic pool. *PLoS One* 5:e10034.
16. Dijkshoorn L, et al. (1996) Comparison of outbreak and nonoutbreak *Acinetobacter baumannii* strains by genotypic and phenotypic methods. *J Clin Microbiol* 34: 1519–1525.
17. van Dessel H, et al. (2004) Identification of a new geographically widespread multiresistant *Acinetobacter baumannii* clone from European hospitals. *Res Microbiol* 155: 105–112.
18. Adams MD, Chan ER, Molyneaux ND, Bonomo RA (2010) Genomewide analysis of divergence of antibiotic resistance determinants in closely related isolates of *Acinetobacter baumannii*. *Antimicrob Agents Chemother* 54:3569–3577.
19. Adams MD, et al. (2008) Comparative genome sequence analysis of multidrug-resistant *Acinetobacter baumannii*. *J Bacteriol* 190:8053–8064.
20. Iacono M, et al. (2008) Whole-genome pyrosequencing of an epidemic multidrug-resistant *Acinetobacter baumannii* strain belonging to the European clone II group. *Antimicrob Agents Chemother* 52:2616–2625.
21. Fournier P-E, et al. (2006) Comparative genomics of multidrug resistance in *Acinetobacter baumannii*. *PLoS Genet* 2:e7.
22. Higgins PG, Dammhayn C, Hackel M, Seifert H (2010) Global spread of carbapenem-resistant *Acinetobacter baumannii*. *J Antimicrob Chemother* 65:233–238.
23. de Vries J, Wackernagel W (2002) Integration of foreign DNA during natural transformation of *Acinetobacter* sp. by homology-facilitated illegitimate recombination. *Proc Natl Acad Sci USA* 99:2094–2099.
24. Reeves PP, Wang L (2002) Genomic organization of LPS-specific loci. *Curr Top Microbiol Immunol* 264:109–135.
25. Antunes LC, Imperi F, Towner KJ, Visca P (2011) Genome-assisted identification of putative iron-utilization genes in *Acinetobacter baumannii* and their distribution among a genotypically diverse collection of clinical isolates. *Res Microbiol* 162: 279–284.
26. Lewis T, et al. (2010) High-throughput whole-genome sequencing to dissect the epidemiology of *Acinetobacter baumannii* isolates from a hospital outbreak. *J Hosp Infect* 75:37–41.
27. Harris SR, et al. (2010) Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 327:469–474.
28. Beres SB, et al. (2010) Molecular complexity of successive bacterial epidemics deconvoluted by comparative pathogenomics. *Proc Natl Acad Sci USA* 107:4371–4376.
29. Chen PE, et al. (2010) Rapid identification of genetic modifications in *Bacillus anthracis* using whole genome draft sequences generated by 454 pyrosequencing. *PLoS ONE* 5:e12397.
30. Linz B, Schenker M, Zhu P, Achtman M (2000) Frequent interspecific genetic exchange between commensal neisseriae and *Neisseria meningitidis*. *Mol Microbiol* 36: 1049–1058.
31. Croucher NJ, et al. (2011) Rapid pneumococcal evolution in response to clinical interventions. *Science* 331:430–434.
32. Milkman R, Jaeger E, McBride RD (2003) Molecular evolution of the Escherichia coli chromosome. VI. Two regions of high effective recombination. *Genetics* 163:475–483.
33. Touchon M, et al. (2009) Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet* 5:e1000344.
34. Rodriguez-Valera F, et al. (2009) Explaining microbial population genomics through phage predation. *Nat Rev Microbiol* 7:828–836.
35. Telford JL (2008) Bacterial genome variability and its impact on vaccine design. *Cell Host Microbe* 3:408–416.
36. Bartual SG, et al. (2005) Development of a multilocus sequence typing scheme for characterization of clinical isolates of *Acinetobacter baumannii*. *J Clin Microbiol* 43: 4382–4390.
37. Rissman AI, et al. (2009) Reordering contigs of draft genomes using the Mauve aligner. *Bioinformatics* 25:2071–2073.
38. Galtier N, Gouy M, Gautier C (1996) SEAVIEW and PHYLO_WIN: Two graphic tools for sequence alignment and molecular phylogeny. *Comput Appl Biosci* 12:543–548.
39. Krzywinski M, et al. (2009) Circos: An information aesthetic for comparative genomics. *Genome Res* 19:1639–1645.