



OPEN

Characterization of a SPM-1 metallo-beta-lactamase-producing *Pseudomonas aeruginosa* by comparative genomics and phenotypic analysis

Ana Paula Barbosa do Nascimento^{1✉}, Fernando Medeiros Filho¹, Heidi Pauer², Luis Caetano Martha Antunes^{2,3}, Hério Sousa⁴, Hermes Senger⁴, Rodolpho Mattos Albano⁵, Marcelo Trindade dos Santos⁶, Ana Paula D'Alincourt Carvalho-Assef⁷ & Fabrício Alves Barbosa da Silva^{1✉}

Pseudomonas aeruginosa is one of the most common pathogens related to healthcare-associated infections. The Brazilian isolate, named CCBH4851, is a multidrug-resistant clone belonging to the sequence type 277. The antimicrobial resistance mechanisms of the CCBH4851 strain are associated with the presence of the *bla*_{SPM-1} gene, encoding a metallo-beta-lactamase, in combination with other exogenously acquired genes. Whole-genome sequencing studies focusing on emerging pathogens are essential to identify key features of their physiology that may lead to the identification of new targets for therapy. Using both Illumina and PacBio sequencing data, we obtained a single contig representing the CCBH4851 genome with annotated features that were consistent with data reported for the species. However, comparative analysis with other *Pseudomonas aeruginosa* strains revealed genomic differences regarding virulence factors and regulatory proteins. In addition, we performed phenotypic assays that revealed CCBH4851 is impaired in bacterial motilities and biofilm formation. On the other hand, CCBH4851 genome contained acquired genomic islands that carry transcriptional factors, virulence and antimicrobial resistance-related genes. Presence of single nucleotide polymorphisms in the core genome, mainly those located in resistance-associated genes, suggests that these mutations may also influence the multidrug-resistant behavior of CCBH4851. Overall, characterization of *Pseudomonas aeruginosa* CCBH4851 complete genome revealed the presence of features that strongly relates to the virulence and antibiotic resistance profile of this important infectious agent.

Pseudomonas aeruginosa is one of the most common pathogens related to healthcare-associated infections in hospitalized individuals worldwide. Multidrug-resistant (MDR) isolates, particularly those non-susceptible to carbapenems, have become a major concern of health institutions. In addition to mechanisms such as loss of porins or overexpression of efflux pumps, carbapenem resistance is produced by the acquisition of genes encoding carbapenem-hydrolyzing beta-lactamases¹. These enzymes are classified into classes A and D, the active-site serine beta-lactamases (SBLs), and class B, the zinc-dependent or metallo-beta-lactamases (MBLs). SBLs have a broad spectrum of activity against beta-lactams but are inhibited by common beta-lactamase inhibitors.

¹Programa de Computação Científica, Fundação Oswaldo Cruz, Rio de Janeiro 21040-360, Brazil. ²Centro de Desenvolvimento Tecnológico em Saúde, Fundação Oswaldo Cruz, Rio de Janeiro 21040-361, Brazil. ³Escola Nacional de Saúde Pública Sergio Arouca, Fundação Oswaldo Cruz, Rio de Janeiro 21041-210, Brazil. ⁴Departamento de Computação, Universidade Federal de São Carlos, São Carlos 13565-905, Brazil. ⁵Departamento de Bioquímica, Universidade do Estado do Rio de Janeiro, Rio de Janeiro 21941-909, Brazil. ⁶Laboratório Nacional de Computação Científica, Petrópolis 25651-075, Brazil. ⁷Laboratório de Pesquisa em Infecção Hospitalar, Fundação Oswaldo Cruz, Rio de Janeiro 21040-360, Brazil. ✉email: ana.pbn@gmail.com; fabricio.silva@fiocruz.br

Feature	Strain	
	CCBH4851	PAO1
Genome status	Complete	Complete
Source	Catheter tip	Wound
Country	Brazil	Australia
Year	2008	Unknown
Genome size (bp)	6,834,257	6,264,404
G+C content (%)	66.08	66.56
No. of total genes	6,342	5,697
No. of pseudogenes	86	19
No. of total CDSs	6,236	5,572
No. of hypothetical proteins	2,476	2,256
No. of rRNAs	13	13
No. of tRNAs	64	63
No. of other RNAs	29	30

Table 1. Genome features of *P. aeruginosa* CCBH4851 compared to *P. aeruginosa* PAO1 reference strain.

MBLs have no activity only against monobactams but are not affected by SBL inhibitors. Although MBLs can be inactivated by chelating compounds, such as ethylenediaminetetraacetic acid, there is not known inhibitor available for clinical use^{2,3}. In Brazil, a recent report shows 42.9% of *P. aeruginosa* isolates recovered from primary bloodstream infections among adult patients hospitalized in intensive care units were resistant to carbapenems⁴. Currently, the main carbapenem resistance mechanism associated with Brazilian isolates is the production of an MBL denominated São Paulo metallo-beta-lactamase (SPM-1). In 2002, the SPM-1-encoding gene (*bla_{SPM-1}*) was first identified in a clinical isolate recovered from a blood culture of a patient admitted in a hospital located in the state of São Paulo. Over the past years, several isolates carrying *bla_{SPM-1}* were recovered from multiple *P. aeruginosa* hospital infection outbreaks widespread in the Brazilian territory. Multilocus sequence typing included most of these isolates within the sequence type 277 (ST-277) which contains *P. aeruginosa* strains from different countries. However, *bla_{SPM-1}* presence was only detected in isolates originating from Brazil and it is unknown why SPM-1 is restricted to a specific region whereas other MBLs tend to spread worldwide. Nevertheless, the risk of SPM-1 worldwide dissemination should not be neglected⁵. In 2008, an isolate was recovered from the catheter tip of a patient admitted in a hospital located in the state of Goiás. The strain, named CCBH4851, was subjected to antimicrobial susceptibility assays. Among the agents tested, the strain was resistant to aztreonam, amikacin, gentamicin, ceftazidime, cefepime, ciprofloxacin, imipenem, meropenem, and piperacillin-tazobactam, being susceptible only to polymyxin B⁶. Multidrug resistance scenarios like those presented by CCBH4851 highlight the urgent need for the discover of new therapeutic options. Thereby, recent whole-genome sequencing studies have focused on emerging MDR *P. aeruginosa* strains in order to identify key aspects of their physiology that may lead to the identification of new targets for drug development. This study was designed to thoroughly characterize the complete closed genome of *P. aeruginosa* CCBH4851 through the detection of genomic features, genome comparison with other *P. aeruginosa* strains, and phenotypic analysis. We believe this study will serve as a useful reference for future research using CCBH4851 as a model organism.

Results

Genome features of *P. aeruginosa* CCBH4851. A complete genome sequence was obtained by the combined assembly of Illumina short and PacBio long reads. The genome consisted of a single circular chromosome with 6,834,257 bp and a G+C content of 66.07%. The number of coding sequences (CDSs) was 6,236, with an average length of 973 bp, which represents 88.78% of the genome. Families of Clusters of orthologous groups (COG) were attributed to 5,247 CDSs comprising 20 functional categories distributed along the chromosome (Fig. 1). A summary of the genomic features found in the CCBH4851 complete genome is listed in Table 1.

Core, accessory and unique genomes. The gene repertoire of CCBH4851 was compared to other strains divided in three groups (Table 4): (A) composed by the commonly used reference strains in research studies, (B) composed by other ST-277 strains also carrying the *bla_{SPM-1}* gene, and (C) composed by other MDR strains belonging to different STs. A search for orthologs among these genomes identified genes that were shared by all genomes (core), genes specific to each genome (unique) or genes shared by two or more (but not all) genomes (accessory). In group A, 5,068 orthologous genes were identified as the core genome, 251 as the accessory genome and 697 genes were unique of CCBH4851 (Supplementary Table S1 online). Apart from genes included in the “Function unknown” COG class, the majority of unique genes were distributed among “Replication, recombination, and repair” and “Transcription” COG classes (Fig. 2A). In group B, the core genome comprised 5,645 genes. The accessory and unique genomes were comprised by 293 and 195 genes, respectively, also having the “Replication, recombination, and repair” and “Transcription” COG classes as the major functional categories with the exception of the “Function unknown” family (Fig. 2B). Group C, including MDR strains of different STs, presented 3,680 genes as the core genome, 1,745 as the accessory genome and 508 as unique genes of CCBH4851. The most evident functional categories remained the same (Fig. 2C). The comparative analysis

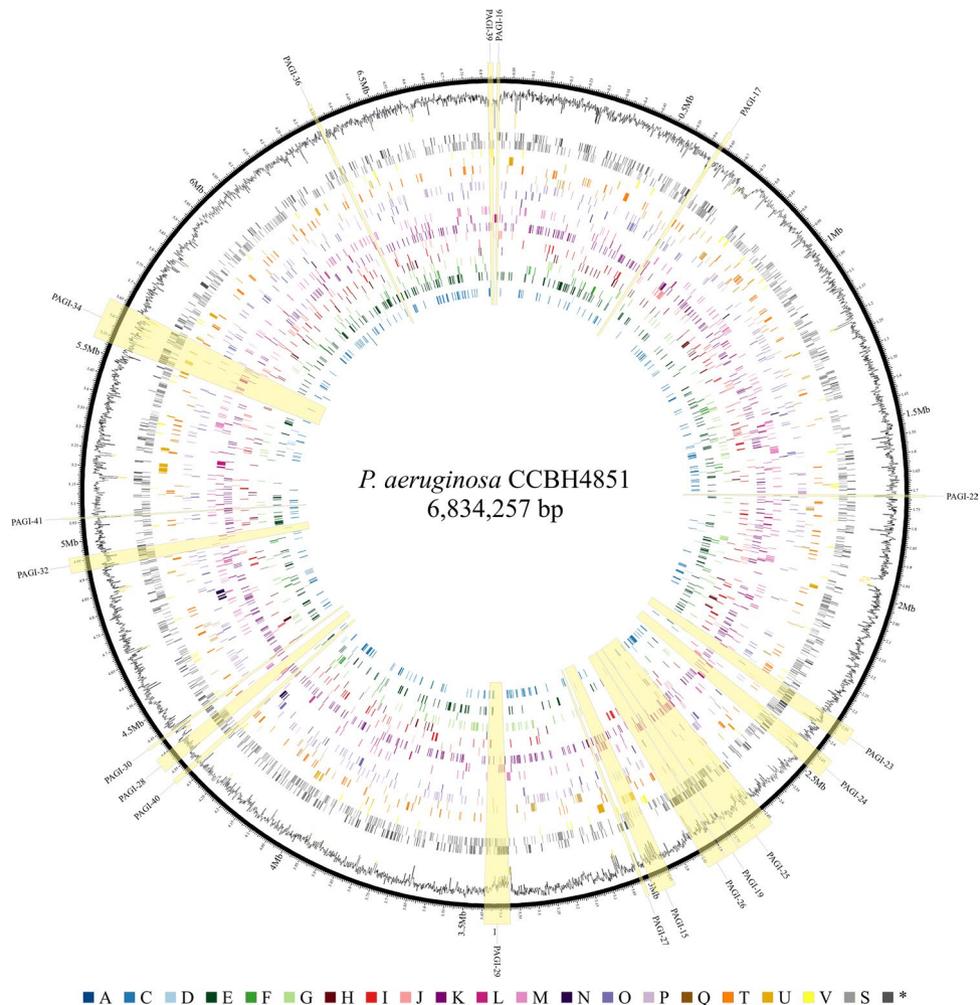


Figure 1. Circular representation of the *P. aeruginosa* CCBH4851 genome. From the inside out, rings display the distribution of genes along the chromosome based on COG classification: (A) RNA processing and modification; (C) energy production and conversion; (D) cell cycle control, cell division, chromosome partitioning; (E) amino acid transport and metabolism; (F) nucleotide transport and metabolism; (G) carbohydrate transport and metabolism; (H) coenzyme transport and metabolism; (I) lipid transport and metabolism; (J) translation, ribosomal structure, and biogenesis; (K) transcription; (L) replication, recombination, and repair; (M) cell wall, membrane, and envelope biogenesis; (N) cell motility; (O) post-translational modification, protein turnover, and chaperones; (P) inorganic ion transport and metabolism; (Q) secondary metabolites biosynthesis, transport, and catabolism; (T) signal transduction mechanisms; (U) intracellular trafficking, secretion, and vesicular transport; (V) defense mechanisms; (S) function unknown; (*) genes with no COG category assigned; following the respective colors in the legend. The outermost ring represents the G+C content. Yellow shaded regions display the localization of genomic islands.

also revealed that the CCBH4851 genome has 205 genes lacking or presenting partial homology when compared to the PAO1 genome (Supplementary Table S2 online). The predominant functional COG categories of the non-orthologous genes were “Replication, recombination, and repair”, “Transcription”, “Intracellular trafficking, secretion, and vesicular transport”, “Inorganic ion transport and metabolism”, “Energy production and conversion”, “Secondary metabolites biosynthesis, transport, and catabolism”, “Defense mechanisms”, and “Coenzyme transport and metabolism”. It is noteworthy that *oprD* (porin), *mexZ*, *liuR*, PA0207, PA0306, PA0306a, PA2100, PA2220, PA2547, PA3067, PA3094, and PA3508 (transcriptional regulators), *opdE* (membrane protein), *pilA* (type IV fimbrial protein), *algP* (alginate regulatory protein), and others were among these genes.

Genomic islands and mobile elements. The *P. aeruginosa* CCBH4851 genome contained 19 integrated genomic islands with a total of 705 CDSs (Fig. 1, yellow highlighted regions; Supplementary Table S3 online). Eighteen islands were homologous to *Pseudomonas aeruginosa* genome islands (PAGIs) previously found in other isolates belonging to ST-277^{7, 8}. An additional region found by IslandViewer was less than 5,000 bp, but had a G+C content of 55.13% and encoded 5 proteins including an integrase and an helix-turn-helix (HTH) domain-containing protein. This region was annotated as a new PAGI, numbered PAGI-41. Overall, 69.92%

Figure 2. Orthology analysis and functional classification by COG. (a) Core and accessory genomes shared between reference strains and CCBH4851, and unique genome of CCBH4851. (b) Core and accessory genomes shared between other ST-277 strains and CCBH4851, and unique genome of CCBH4851. (c) Core and accessory genomes shared between other ST strains and CCBH4851, and unique genome of CCBH4851. Legend of x-axis: (A) RNA processing and modification; (C) energy production and conversion; (D) cell cycle control, cell division, chromosome partitioning; (E) amino acid transport and metabolism; (F) nucleotide transport and metabolism; (G) carbohydrate transport and metabolism; (H) coenzyme transport and metabolism; (I) lipid transport and metabolism; (J) translation, ribosomal structure, and biogenesis; (K) transcription; (L) replication, recombination, and repair; (M) cell wall, membrane, and envelope biogenesis; (N) cell motility; (O) post-translational modification, protein turnover, and chaperones; (P) inorganic ion transport and metabolism; (Q) secondary metabolites biosynthesis, transport, and catabolism; (T) signal transduction mechanisms; (U) intracellular trafficking, secretion, and vesicular transport; (V) defense mechanisms; (S) function unknown; (*) genes with two or more COG categories assigned.

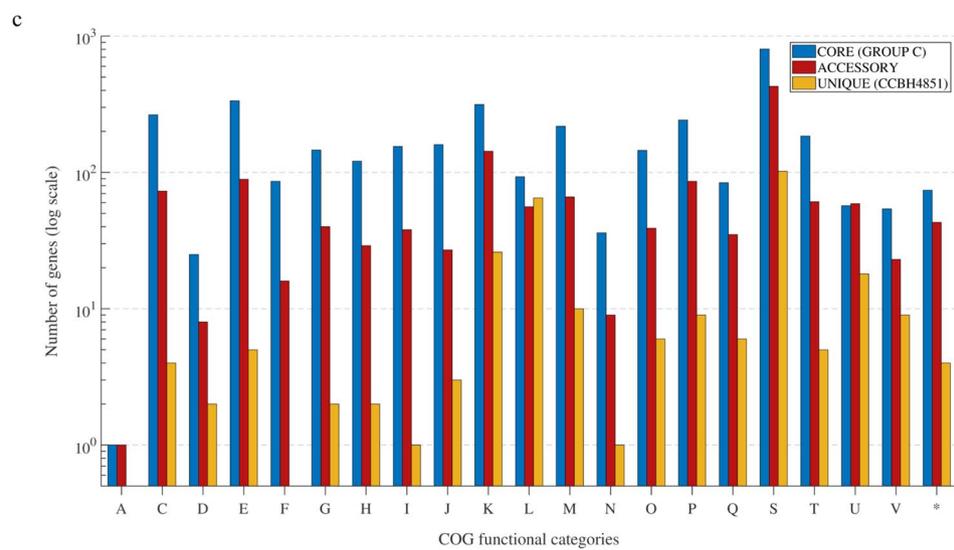
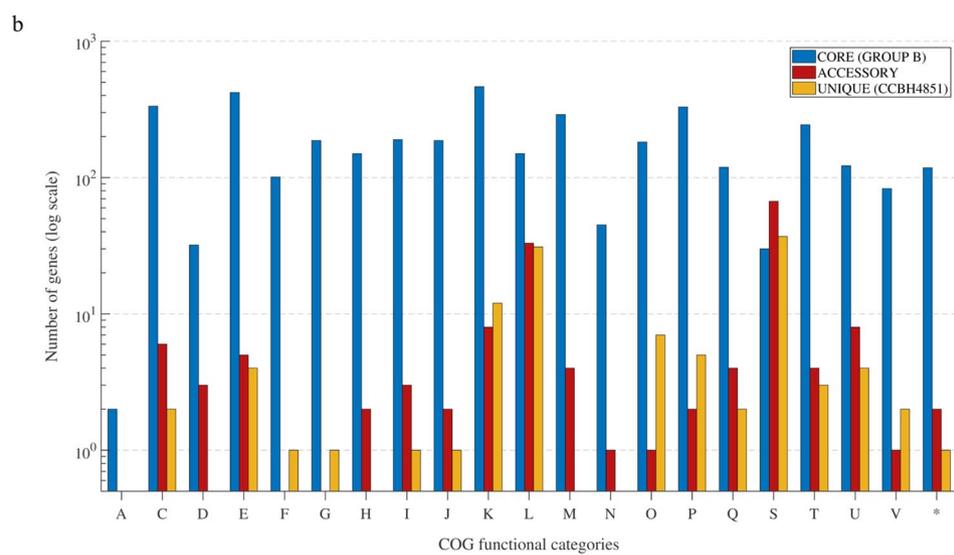
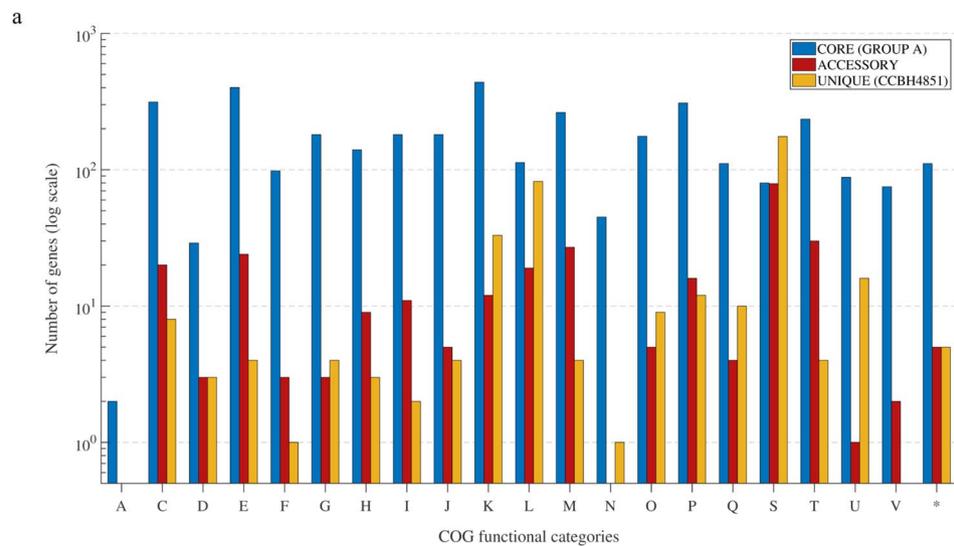
of the genes located in the PAGIs had no function assigned. The remainder were mainly classified in the following COG categories: “Replication, recombination, and repair”, “Transcription”, “Intracellular trafficking, secretion, and vesicular transport”, “Energy production and conversion”, “Secondary metabolites biosynthesis, transport, and catabolism”, “Inorganic ion transport and metabolism” and “Defense mechanisms”. A total of 13 insertion sequences’ (IS) families were found in the CCBH4851 genome with 46 predicted interspersed open reading frames (ORFs), including complete and partial sequences (Supplementary Table S4 online). A search for clustered regularly interspaced short palindromic repeats (CRISPR) revealed the presence of two CRISPR loci in CCBH4851 genome: one had only 1 spacer and direct repeats of 25 bp, and the other had 39 spacers and direct repeats of 32 bp. The first CRISPR was located at position 1,878,269 to 1,878,370 in the genome, between two hypothetical protein-encoding genes. The DNA sequence referring to this first CRISPR array was a variant of the intergenic region between PA3230 and PA3231 genes in PAO1 (homologous to AL347_00365 and AL347_00370). The second CRISPR was located at 5,576,114 to 5,578,733, partially overlapping the hypothetical protein-encoding gene AL347_11857. In addition, the second CRISPR array was located adjacent to an intact CRISPR-associated gene cluster, *cas2* (AL347_11860), *cas1* (AL347_11865), *cas4* (AL347_11870), *cas7* (AL347_11875), *cas8c* (AL347_11880), *cas5* (AL347_11885), *cas3* (AL347_11890). These *cas* genes were related to the type I-C CRISPR-Cas system and were located inside PAGI-34, as previously described in other ST-277 clones⁷.

Regulatory proteins.

Comparative genome analysis revealed strong evidences of horizontal gene transfer events in the CCBH4851 genome. Despite the predominance of genes encoding hypothetical proteins in these acquired regions, several CDSs were classified into the “Transcription” COG functional class. Regulatory protein analysis was performed to predict the presence of two-component systems, transcription factors, and other DNA-binding proteins in the CCBH4851 genome. Table 2 summarizes a comparison between CCBH4851 and PAO1 regulatory proteins. In accordance with the comparative analysis, the CCBH4851 genome lacks the histidine-kinase PA2583, the response regulator PA0034, as well as the transcriptional regulators PA3508, PA3067, and LiuR (PA2016). However, the CCBH4851 genome possessed 28 additional regulatory proteins, mostly transcriptional regulators with Xre- and LysR-type domains, distributed along PAGIs (Supplementary Table S5 online).

Antimicrobial resistance factors. Protein sequences from the CCBH4851 genome were used for searches against The Comprehensive Antibiotic Resistance Database (CARD). As previously described⁶, the CCBH4851 genome possessed additional genes conferring multidrug resistance: two copies of *blas_{SPM-1}* (AL347_32235, AL347_32285), two copies of *sul1* (AL347_30655, AL347_30700), *rmtD* (AL347_30670), *bla_{OXA-56}* (AL347_30635), *aac(6’)-I* (AL347_30630), *aadA7* (AL347_30640), *cmx* (AL347_30710), and three copies of *bcr* (AL347_32216, AL347_32266, AL347_32316). As expected, all these genes were located in PAGIs. Apart from *mexZ*, genes listed as belonging to the PAO1 resistome were also present in the CCBH4851 genome (Supplementary Table S6 online).

Single nucleotide polymorphisms. Single nucleotide polymorphisms (SNPs) detection between *P. aeruginosa* CCBH4851 and PAO1 genomes identified 25,220 variant types, including SNPs *per se*, multiple nucleotide polymorphisms (MNP), insertions, deletions and a combination of SNP/MNP⁹. Results revealed 16,690 synonymous variants, 4,972 missense variants, 7 stop gained, 5 stop lost, 5 start lost, 39 in frame deletions/insertions, 40 frameshift variants, and 3440 variant types affecting intergenic regions (Supplementary Table S7 online). Synonymous variants were left aside, as they have no presumable impact on cellular processes. Apart from intergenic regions, the remaining affected genes were classified into COG families to assess whether variant types were common to a few functional categories or were randomly distributed among all of them. Mutations occurred mainly in genes belonging to “Inorganic ion transport, and metabolism”, “Amino acid transport and metabolism”, “Cell wall, membrane, and envelope biogenesis”, “Transcription”, and “Energy production and conversion” COG categories (Fig. 3). In addition, Table 3 summarizes a list of virulence and antimicrobial resistance-associated genes with their respective amino acid substitutions. The hypothetical structural effect of new substitutions was predicted and listed in details in Supplementary Table S8 online.



Strain	Predicted regulatory proteins							
	TCS			TF				ODP
	HK	RR	PP	TR	OCS	RR	SF	
PAO1	62	73	4 ^a	170	185	49	24	36
CCBH4851	61	72	5	189	191	48	25	38

Table 2. Regulatory proteins of *P. aeruginosa* CCBH4851 compared to *P. aeruginosa* PAO1 reference strain. ^aA PP homologous sequence is present in the PAO1 DNA sequence, but it is not annotated in the deposited Genbank file. TCS two-component systems, TF transcription factors, ODP other DNA-binding proteins, HK histidine-kinases, RR response regulators, PP phosphotransferase proteins, TR transcriptional regulators, OCS one-component systems, SF sigma factors.

Phylogenomic analysis of *P. aeruginosa* CCBH4851. The relatedness of CCBH4851 genome was assessed by phylogenetic inference amongst the isolates of group A, B, and C using the whole genome to perform a SNP calling and multiple sequence alignment of the resulting core genome. *P. aeruginosa* PAO1 was used as reference genome and had 92.9% of its sequence covered by all isolates. Interesting, isolates clustered according to the sequence types in the same phyletic clade (Fig. 4). Also, phylogenomic analysis showed that CCBH4851 was more nearly related to isolates of ST-244, sharing about 2-fold less SNP variants with them than with ST-235 isolates. *P. aeruginosa* PA7, as expect, was excluded from the resulting tree since it is known to be a taxonomic outlier¹⁰.

Swarming, swimming and twitching motilities. Swarming, swimming and twitching assays showed that CCBH4851 is impaired in all three types of motility. Swarming and swarming zones of CCBH4851 were about 5 millimeters (diameter) smaller when compared to PAO1 (Figs. 5 and 6, respectively). Both differences were statistically significant (swimming: P value = 0.0048, two-tailed, unpaired t test; swarming: P value < 0.001, two-tailed, unpaired t test). Comparisons of twitching motility showed no differences concerning the diameter of migration zones (Fig. 7A). However, after removal of unattached cells and staining with crystal violet, we observed diminished attachment of CCBH4851 cells to the polystyrene surface when compared to PAO1 (Fig. 7B). Observation of the stained area showed that CCBH4851 cells produced an expanding circular adherent zone, indicating that only cells on the outer side of the twitching area were attached, whereas PAO1 cells seemed to adhere to the polystyrene surface throughout the entire area, with only a few bacteria released from the center of the colony. Absorbance measurements of solubilized crystal violet confirmed the significant difference seen between the two strains. In order to assess if the diminished attachment observed would affect the ability of CCBH4851 to form biofilms, 96-well plate biofilm formation assays were performed. Under the conditions tested, CCBH4851 formed about fourfold less biofilm than PAO1 (Fig. 7C, P value < 0.0001, two-tailed, unpaired t test).

Antimicrobial susceptibility. Antimicrobial susceptibility testing was performed using one drug of each of four major antimicrobial classes. Tests were performed using the broth microdilution method, which allowed us to determine a minimal inhibitory concentration (MIC) for each drug. CCBH4851 was resistant to all antimicrobial agents tested except for colistin. Colistin susceptibility tests presented more than one skipped well, with no pattern. According to the European Committee on Antimicrobial Susceptibility Testing (EUCAST) guidelines, tests with more than one skipped wells should not be reported¹¹. However, the majority of replicates showed no growth above the resistance breakpoint for colistin. The MICs for gentamicin, imipenem and ciprofloxacin were >1024 $\mu\text{g}/\text{mL}$, 128 $\mu\text{g}/\text{mL}$ and 16 $\mu\text{g}/\text{mL}$, respectively. In addition, an E-test strip containing ceftolozane/tazobactam (a cephalosporin in combination with a beta-lactamase inhibitor) was also employed and CCBH4851 showed no inhibition zone (Supplementary Fig. S2 online).

Discussion

Pseudomonas aeruginosa CCBH4851 is a clinical isolate originating from Brazil, which belongs to the endemic group ST-277. Previous studies demonstrated that ST-277 strains have a highly conserved DNA sequence and share several virulence and antimicrobial resistance-related features. However, to date, the presence of the *bla*_{SPM-1} gene conferring resistance to carbapenems is a singular trait of Brazilian clones^{5,7}. The purpose of this study was to reassemble and to re-annotate the CCBH4851 genome in order to perform a thorough characterization and a genomic comparison with other *P. aeruginosa* strains. The reassembly resulted in a single contig representing the complete closed genome of CCBH4851. Together with the provided annotation, data such as chromosome size, G+C content, number of CDSs, structural RNAs, and tRNAs were consistent with general features reported for *P. aeruginosa* strains¹². In addition, the CCBH4851 genome alignment with PAO1, PA14, and ST-277 strains revealed a strong synteny (Supplementary Fig. S1 online) indicating an accurate assembly and annotation.

The analyses performed in this work showed that the core genome of CCBH4851, PAO1, and PA14 strains (group A, Table 4) covered more than 80% of the CCBH4851 genome. A similar percentage was found when the core genome was defined comparing CCBH4851 with other isolates belonging to the ST-277 also carrying *bla*_{SPM-1} (group B, Table 4). However, when comparing MDR strains belonging to different STs (group C, Table 4), the core genome was smaller and an increase in the number of genes comprised in the accessory genome was

Gene name	Amino acid substitution	Product
<i>alg8</i>	L61V	Glycosyltransferase Alg8
<i>algC</i>	T113A, R120L, A124V, L274P, T289A	Phosphomannomutase
<i>algE</i>	N34G, T475A	Alginate production protein AlgE
<i>algF</i>	T11N	Alginate o-acetyltransferase AlgF
<i>algI</i>	N403S, V1360AF	Alginate o-acetylase AlgI
<i>algK</i>	M202L	Alginate biosynthesis protein AlgK
<i>algX</i>	Y49D	Alginate biosynthesis protein AlgX
<i>amgS</i>	I260V	Protein AmgS
<i>ampC</i>	R79Q ³⁴ , T105A ³⁴	Beta-lactamase
<i>ampD</i>	G148A, S175L	N-acetyl-anhydromuranmyl-L-alanine amidase
<i>ampP</i>	M87I, T172A, M177V	Transporter
<i>armZ</i>	L88P, D161G, H182Q, V243A ³³	MexZ anti-repressor ArmZ
<i>dacB</i>	A394P	D-alanyl-D-alanine carboxypeptidase
<i>exoT</i>	A83S, M107I, A135V, V216A	Exoenzyme T
<i>exoY</i>	P203L, S366G	Adenylate cyclase
<i>glpT</i>	A439T	Glycerol-3-phosphate transporter
<i>gyrA</i>	T83I ³²	DNA gyrase subunit A
<i>lasA</i>	V192A, S223R	Protease LasA
<i>lasB</i>	S241G, S436L	Elastase LasB
<i>lasR</i>	V221G	Transcriptional regulator LasR
<i>mexD</i>	E257Q, S845A	Resistance-nodulation-cell division (RND) multidrug efflux transporter MexD
<i>mexE</i>	P397Q, A407V	Resistance-nodulation-cell division (RND) multidrug efflux membrane fusion protein MexE
<i>mexI</i>	A782E	Resistance-nodulation-cell division (RND) efflux transporter
<i>mexS</i>	D249N	Oxireductase
<i>mexT</i>	F172I	Transcriptional regulator MexT
<i>mexZ</i>	L174Q	Transcriptional regulator
<i>mucD</i>	I137V	Serine protease MucD
<i>nalC</i>	G71E ¹⁴ , S209R ¹⁴	Transcriptional regulator
<i>oprD</i>	T103S, V127fs ³² , F170L ¹³	porin D
<i>oprJ</i>	DM68GV	Multidrug efflux outer membrane protein OprJ
<i>oprM</i>	A261T	Outer membrane protein OprM
PA0032	A186V, L214V	Transcriptional regulator
PA1184	A249V	Transcriptional regulator
PA2018	T543A ³³ , Q840E ³³	Multidrug efflux protein
PA2019	K329Q ³³ , L331V ³³ , W358R ³³	multidrug Efflux lipoprotein
PA3271	A212V	two-component sensor
<i>parC</i>	S87L ³² , H262Q	DNA topoisomerase IV subunit A
<i>pmrA</i>	L71R ⁶²	Two-component regulator system response regulator PmrA
<i>pmrB</i>	Y345H ⁶²	Two-component regulator system signal sensor kinase PmrB
<i>rhIA</i>	Q168P	Rhamnosyltransferase subunit A
<i>rhIC</i>	E21K	Rhamnosyltransferase
<i>rhII</i>	D83E	Acyl-homoserine-lactone synthase
<i>rpoB</i>	V51I	DNA-directed RNA polymerase subunit beta
<i>toxA</i>	T4I, L14F, V55I, I432V	Exotoxin A

Table 3. SNPs found in virulence and antimicrobial resistance-associated genes of *P. aeruginosa* CCBH4851 using *P. aeruginosa* PAO1 as reference.

observed. This result could be attributed to different factors: (i) some strains used in this analysis had over one hundred contigs and this lack of continuity could affect genome annotation; (ii) the strains used in this analysis are from different countries and had distinct resistance profiles, which could suggest a variation of mutation patterns and acquired regions, depicting interstrain differences. The larger accessory genome among these strains suggests a variability in pathogenicity and environmental flexibility.

The comparative genome analysis revealed total or partial absence of homology between several PAO1 genes and the CCBH4851 genome. Among them, there is *oprD*, which suffered a deletion of 2 nucleotides in positions 380-381, causing a frameshift resulting in the gain of a premature stop codon. OprD is an outer-membrane porin important for the diffusion of carbapenems, particularly imipenem. Disruption of OprD is a common

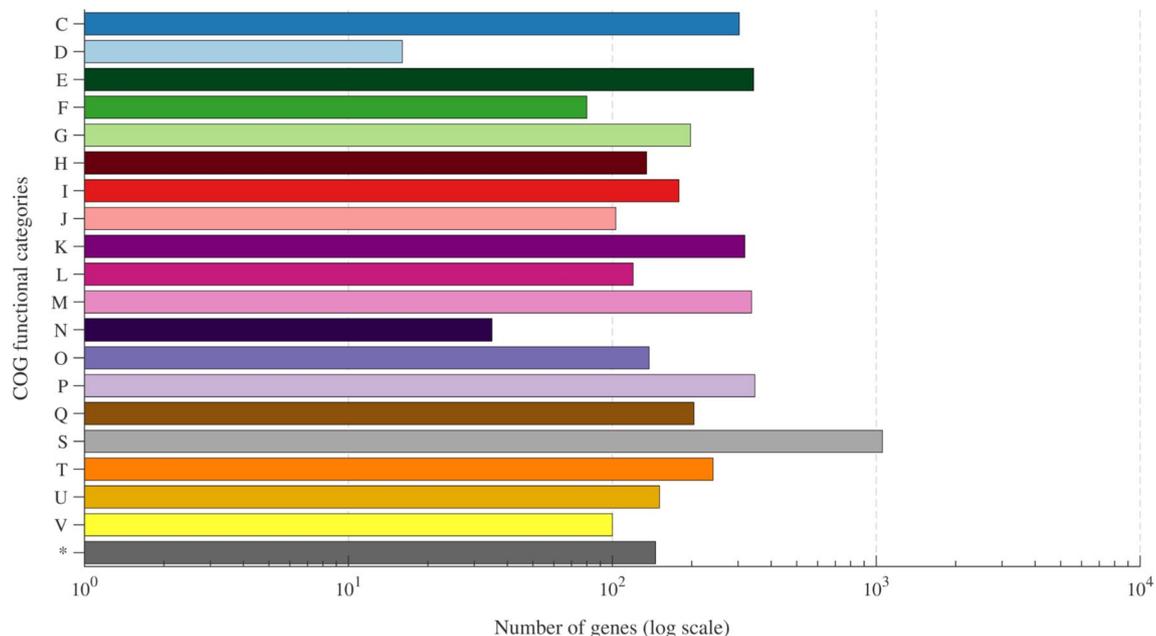


Figure 3. Functionally clustered genes affected by SNPs causing amino acid substitutions. Description of COG categories: (C) energy production and conversion; (D) cell cycle control, cell division, chromosome partitioning; (E) amino acid transport and metabolism; (F) nucleotide transport and metabolism; (G) carbohydrate transport and metabolism; (H) coenzyme transport and metabolism; (I) lipid transport and metabolism; (J) translation, ribosomal structure, and biogenesis; (K) transcription; (L) replication, recombination, and repair; (M) cell wall, membrane, and envelope biogenesis; (N) cell motility; (O) post-translational modification, protein turnover, and chaperones; (P) inorganic ion transport and metabolism; (Q) secondary metabolites biosynthesis, transport, and catabolism; (S) function unknown; (T) signal transduction mechanisms; (U) intracellular trafficking, secretion, and vesicular transport; (V) defense mechanisms; (*) genes with two or more COG categories assigned; following the respective colors in the legend.

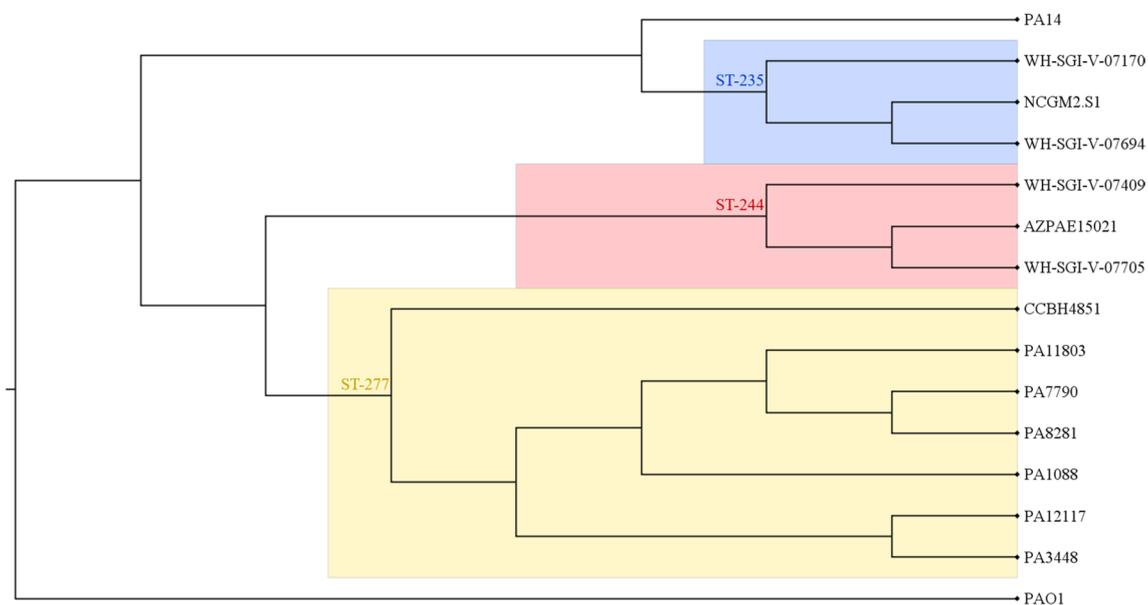


Figure 4. Relatedness of *P. aeruginosa* genomes. Core genome phylogenetic tree based on SNP calling using *P. aeruginosa* PAO1 as reference. Three distinct phyletic clades are highlighted corresponding to multilocus sequence typing: ST-277 (yellow shaded), ST-244 (red shaded), and ST-235 (blue shaded).

resistance mechanism, mainly when combined with overexpression of AmpC and efflux pump systems¹³. As result, CCBH4851 showed a MIC of 128 $\mu\text{g}/\text{mL}$ for imipenem.

Comparison with PAO1 also revealed the mutation of *mexZ* in CCBH4851 genome, suggesting the overexpression of at least one efflux pump system, MexXY, whose transcription is repressed by MexZ. A deletion of 17 nucleotides from position 439 in *mexZ* gene sequence caused a frameshift resulting in the loss of the stop codon.

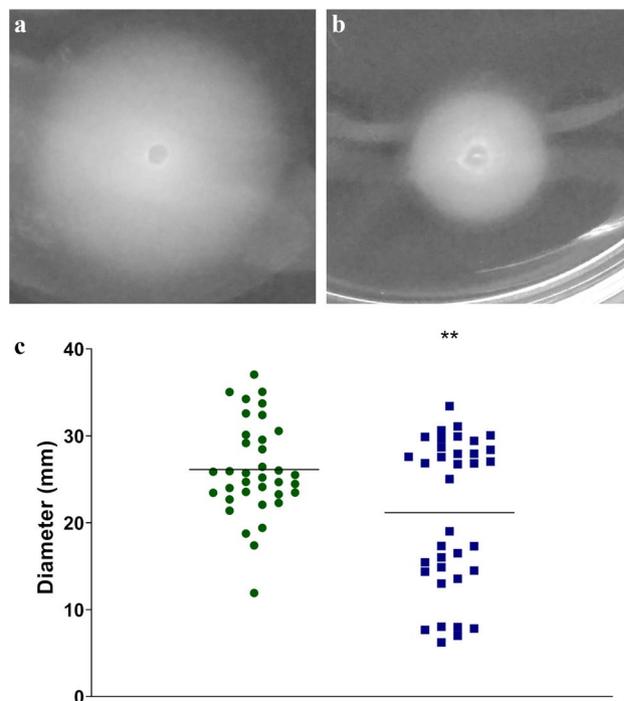


Figure 5. Differences in the swimming phenotype of *P. aeruginosa* PAO1 and CCBH4851 strains. Circular zone formed by (a) PAO1 and (b) CCBH4851 in 0.3% LB agar plates. (c) Measurement of the circular zone diameter of PAO1 (green circles) and CCBH4851 (blue squares). Horizontal lines are the mean of all replicates.

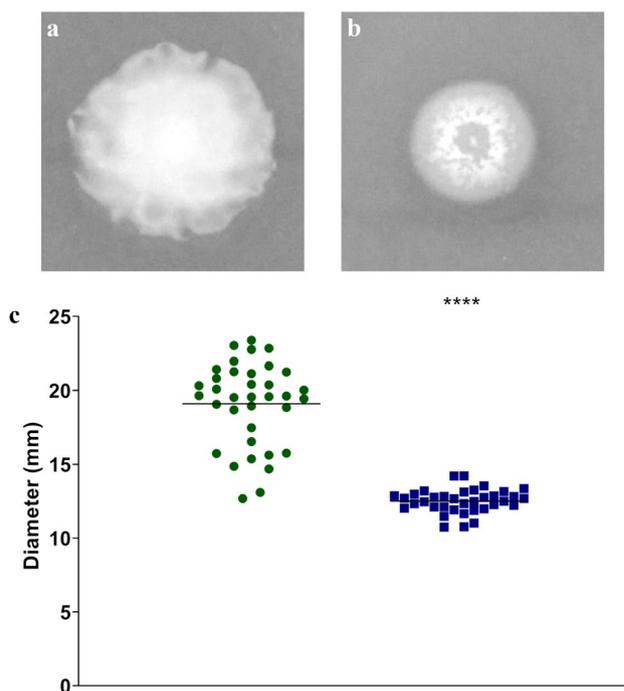


Figure 6. Differences in the swarming phenotype of *P. aeruginosa* PAO1 and CCBH4851 strains. Circular zone formed by (a) PAO1 and (b) CCBH4851 in 0.6% LB agar plates. (c) Measurement of the circular zone diameter of PAO1 (green circles) and CCBH4851 (blue squares). Horizontal lines are the mean of all replicates.

Mutations in *mexZ* are often present in clinical isolates overexpressing MexXY¹⁴. Aminoglycoside resistance is known to be produced not only by the overexpression of MexXY but the presence of aminoglycoside-modifying enzymes (AMEs) and rRNA methylases¹⁵. In addition to the possible efflux mechanism, CCBH4851 acquired

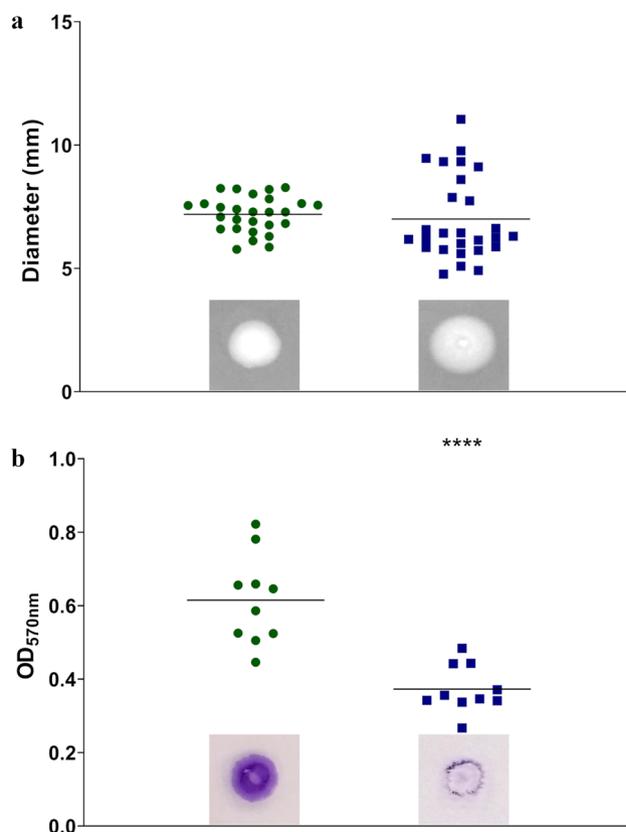


Figure 7. The twitching phenotype of *P. aeruginosa* PAO1 and CCBH4851 strains. **(a)** Measurement of the circular zone diameter of PAO1 (green circles) and CCBH4851 (blue squares), and visual aspect of the respective twitching zone in 1% LB agar plates. **(b)** Measurements of solubilized crystal violet absorbance of PAO1 (green circles) and CCBH4851 (blue squares), and visual aspect of respective stained cells area. **(c)** Quantification of biofilm formation of PAO1 (green bar) and CCBH4851 (blue bar). Horizontal lines are the mean of all replicates. Error bars represent the standard deviation of the mean.

Strain	No. of contigs	Total size (bp)	Sequence type	Accession no.
<i>- Group A</i>				
PAO1	1	6,264,404	549	NC_002516
UCBPP-PA14 (PA14)	1	6,537,648	253	NC_008463
<i>- Group B</i>				
PA1088	1	6,721,480	277	CP015001
PA3448	2	6,794,242	277	LVWC01000000
PA7790	1	7,018,690	277	CP014999
PA8281	1	6,928,736	277	CP015002
PA11803	1	7,006,578	277	CP015003
PA12117	3	6,643,782	277	LVXB01000000
<i>- Group C</i>				
PA7	1	6,588,339	1195	NC_009656
NCGM2_S1	1	6,764,661	235	AP012280
WH-SGI-V-07170	119	6,813,449	235	NZ_LLLZ00000000
WH-SGI-V-07694	106	6,776,438	235	NZ_LLSO00000000
AZPAE15021	129	6,686,424	244	NZ_JTNZ00000000
WH-SGI-V-07409	144	6,932,716	244	NZ_LLOW00000000
WH-SGI-V-07705	130	6,802,755	244	NZ_LLSZ00000000

Table 4. Bacterial strains used for comparative genome analysis.

two AME-encoding genes, *aac(6')-I* and *aadA7*, and a 16S rRNA methylase-encoding gene, *rmtD*, resulting in a MIC over 1024 $\mu\text{g}/\text{mL}$ for gentamicin.

Other differences revealed by the comparative analysis were the mutations of *algP* and *pilA* genes, both related to biofilm development. AlgP is a histone-like protein which activates the transcription of *algD*, responsible for the alginate precursor synthesis. Alginate is an important molecule for the structural stability of the biofilm matrix. The amino acid sequence of AlgP contains several repeated KPAA motifs which are mutation targets at high frequency in clinical isolates. According to this, the *algP* gene sequence of CCBH4851 had an insertion of 12 nucleotides, adding one KPAA motif to the AlgP protein sequence. The repeated 12-bp sequences of *algP* are considered a hot spot for DNA rearrangements and could provide a reversible switching mechanism between nonmucoid and mucoid phenotype, thus turning on and off virulence factors important for a successful infection^{16,17}. On the other hand, the *pilA* sequence is completely absent from the CCBH4851 genome. PilA is a pilin protein involved in type IV pilus (TFP) biogenesis, a major surface adhesin. Previous work demonstrated that the absence of *pilA* in carbapenem-resistant clinical isolates is not uncommon¹⁸.

P. aeruginosa is capable of switching between three types of motility depending on the environment: swimming, swarming and twitching. Swimming is mediated by flagella in aqueous environments. Swarming is mediated by both flagella and TFP in semisolid environments. Twitching is TFP-dependent and allows the cells to move across surfaces by extension, attachment and retraction of the pilus filament. Also, surface attachment mediated by TFP is required for biofilm initiation^{19–21}. As expected, our *in vitro* analyses showed that CCBH4851 is impaired in swarming, twitching, and biofilm development. *P. aeruginosa* CCBH4851 was also defective in its ability to swim. Although CCBH4851 presented no lack of known genes directly involved in flagellar biosynthesis, several SNPs were located in those genes (Supplementary Table S7 online), including *flhF*, *fliA*, *fleN*, *motB*, and others. During acute infection, motility is required to the colonization of environments, attachment to surfaces and initiation of biofilm formation, mechanisms strongly related to virulence. Motile bacteria often display a more virulent phenotype. Loss of motility is important to the switching for a sessile lifestyle and the establishment of chronic infection. Also, it is noteworthy that virulence and multidrug resistance are often inversely correlated, i.e. MDR strains are less or non-virulent whereas non-MDR strains are hypervirulent^{22,23}. In that sense, the phenotype displayed by CCBH4851 suggest a less virulent behavior in favor of a survival strategy under selective environmental pressures.

The acquisition of exogenous material by horizontal gene transfer is a common adaptive mechanism among *P. aeruginosa* strains, often related to the presence of genomic islands. Genomic islands are clusters of genes often encoding virulence factors, antimicrobial resistance proteins, toxins, secretion system proteins, transcriptional regulators, and other proteins²⁴. The CCBH4851 genome presented a high number of islands, all sharing homology with previously described PAGIs, except for PAGI-41. Although the majority of proteins encoded by these PAGIs are hypothetical proteins, some of them could be clustered into functional categories due to the presence of conserved domains. The outstanding categories comprised genes involved in mechanisms of replication, recombination and repair, and transcription. Indeed, the newly annotated PAGI-41 carries the gene AL347_09405 which encodes an XRE-type HTH domain-containing protein, suggesting a function as a transcriptional repressor. Downstream, the gene AL347_09410 encodes a protein containing a RelE/ParE toxin superfamily domain. In response to environmental stress, such as antimicrobial exposure, this type of toxin can inhibit replication or translation mechanisms, leading the cell to a dormant state. Cells in dormant state, also called persister cells, become multidrug-tolerant. The repressor binds to the toxin forming an inactive complex acting as an antitoxin, and is able to binding its own promoter region preventing toxin-antitoxin expression. Antitoxins could also regulate transcription of several virulence factors, including biofilm formation^{25,26}. CCBH4851 possess three additional toxin-antitoxin system in other PAGIs. Moreover, all acquired genes conferring resistance to a broad range of antimicrobial agents were located in PAGIs, including the two copies of the carbapenem resistance gene *bla_{SPM-1}*. Our *in vitro* analyses showed that CCBH4851 is not only resistant to several antimicrobial agents, but also presented MIC values over the resistance breakpoints.

In addition to the presence of PAGIs, distinct IS families were also detected in the CCBH4851 genome. The importance of ISs is not restricted to their role in horizontal gene transfer, but ISs movement along the chromosome can also affect antibiotic resistance by the activation of gene expression²⁷. Another player in shaping the bacterial genome is the CRISPR-Cas system. The type I-C found in the CCBH4851 genome was previously described in other ST-277 bacteria as well as in members of ST-235^{7,28}. A recent study suggests the protective effect of CRISPR-Cas systems is more evident at the population level than at an evolutionary scale²⁹. This could explain the intraclonal genome conservation already observed among Brazilian isolates belonging to ST-277 in previous work⁷ and corroborated here by our phylogenomic analysis.

During the infection course, *P. aeruginosa* strains tend to adapt to the selective pressures of the environment, often through mutations in intergenic and/or coding sequence regions. A classical example of adaptive mutation is the MexT protein, responsible for activation of the *mexEF-OprN* operon. MexEF-OprN is quiescent in wild-type cells and its expression occurs following mutations in MexT or MexT-related genes. Indeed, the *mexT* gene sequence of CCBH4851 had an 8-bp deletion known to render MexT active, causing the induction of *mexEF-oprN* transcription and a decrease in OprD levels, which characterizes the so-called *nfxC*-type mutants³⁰. However, later work showed that this 8-bp deletion alone was not sufficient to activate the transcription of *mexE*, and the generation of *nfxC*-type mutants seems to be multifactorial. Indeed, inactivation of *mexS*, a gene upstream of *mexT*, seems to be one of these additional factors³¹. The *mexS* gene of CCBH4851 suffered a deletion of 1 nucleotide in position 22 causing a frameshift which resulted in the loss of the stop codon. The gene product of the *mexS* pseudogene showed an alignment of only 7 aa with the MexS wild-type sequence of PAO1. This could suggest the absence of MexS in the CCBH4851 proteome, leading to the overexpression of *mexE* and other phenotypes related to *nfxC*-type mutants. Table 3 summarizes other SNPs in virulence and resistance-associated genes that are frequently mutated in *P. aeruginosa* clinical isolates. Among them, amino acid substitutions T83I

in GyrA and S87L in ParC are well-known to play an important role in fluoroquinolone resistance as well as the overexpression of efflux pump systems³². Once again, the antimicrobial susceptibility testing revealed a MIC of 16 µg/mL for ciprofloxacin, corroborating the observed genotype.

Quantification of efflux pump systems expression is required to confirm altered transcription levels, but amino acid substitutions in NalC (repressor of *mexA* gene), ArmZ (MexZ anti-repressor), PA2018 and PA2019 proteins suggest additional mechanisms leading to the overexpression of *mexA* and *mexX*. In addition to these factors, amino acid substitutions R79Q and T105A found in AmpC (Table 3) characterizes this protein as a variant type named PDC-5. Clinical isolates carrying this variant presented AmpC overexpression, increased beta-lactamase activity and reduced susceptibility to ceftazidime, cefepime, ceftiofime, aztreonam, imipenem, and meropenem^{14, 32–34}. Although SPM-1 is an enzyme capable of hydrolyzing a broad range of beta-lactams, it is known that MBLs has little or no effect against monobactams^{2, 3}. However, CCBH4851 was resistant to aztreonam⁶, a monobactam widely available for clinical use. This phenotype can be explained by the presence of PDC-5, depicting a synergistic effect between AmpC and SPM-1 on beta-lactam resistance mechanisms of CCBH4851.

Not all SNPs found in this work were previously described; however, different mutations in the same genes listed in Table 3 were observed in other clinical isolates, suggesting a recurrence in the mutations' location caused by selective pressure. The variants found in CCBH4851 could be intraclonal, but could still change gene function, thus contributing to the enhanced resistance of ST-277 clones. In fact, the *in silico* analysis suggests some of these SNPs could affect protein function due to the introduction of amino acids with different properties, such as size, charge, hydrophobicity, often located in protein domains. The prediction indicates noteworthy mutations occurring in genes such as *ampD* (negative regulator of *ampC* expression), *mucD* (repressor of *algU* alternative sigma factor transcription), *oprJ* (member of *mexCD* efflux system), and others, which could affect protein folding and/or function (Supplementary Table S8 online).

The features found in the *P. aeruginosa* CCBH4851 complete genome are somehow related to its pathoadaptive behavior, since part of them are common among clinical isolates and some are unique of ST-277 clones. The number of unique genes, the number of genes contained in the acquired genomic islands, and the larger size of the CCBH4851 chromosome are consistent with this observation. However, most of the genome is shared with more susceptible strains, suggesting the high number of mutations found in conserved genes are contributing to the success of this MDR strain. Further validation of uncharacterized polymorphisms revealed by this study should help increase the understanding of MDR strains phenotype. The results provided in the present study will stimulate and contribute to future works that lead to the discovery of new antimicrobial targets and to the development of effective therapeutic choices against *P. aeruginosa* infections.

Methods

Bacterial strains. The focus of the present study was the bacterial strain *Pseudomonas aeruginosa* CCBH4851. The isolate is available in the Culture Collection of Hospital-Acquired Bacteria (CCBH), located at Fundação Oswaldo Cruz (WDCM947; CGEN022/2010). Other strains used to perform comparative genome analysis are listed in Table 4.

Genome sequencing, reassembly and re-annotation. Whole-genome sequencing of CCBH4851 was previously performed using the Illumina MiSeq platform⁶. Genome assembly resulted in a draft genome comprising 150 contigs, which were annotated as described in previous work. Later, an additional whole-genome sequencing was performed using the PacBio platform. Here, *de novo* assembly was carried out using the MaSuRCA assembler, which allowed the combination of Illumina short reads with PacBio long reads resulting in an improvement of the original assembly³⁵. A single contig was obtained and was annotated using a customized pipeline. Briefly, annotation from the closely related strain *P. aeruginosa* PAO1 was transferred to the CCBH4851 genome using the Rapid Annotation Transfer Tool software³⁶. Then, CDSs were predicted by GeneMarkS³⁷. Prediction of rRNAs and tRNAs were performed by RNAmmer and tRNAscan-SE software, respectively^{38, 39}. The predicted CDSs were functionally annotated based on homology searches against public databases, such as RefSeq⁴⁰, COG⁴¹, KEGG⁴², UniProt⁴³ and InterPro⁴⁴. Finally, the annotations from previous steps were compared, combined and manually curated. Additionally, PRIAM software was used to assign an enzyme commission number to CDSs⁴⁵. All genome visualization was made using the Artemis software⁴⁶. The circular plot of CCBH4851 genome was produced by Circos software⁴⁷.

Comparative genome analysis. Whole-genome comparison was performed to identify similarities and differences between CCBH4851 and each group of strains listed in Table 4. The genomes were analyzed using the bidirectional best-hit (BBH) clustering method based on homology searches using the BLAST algorithm to identify pairs of corresponding genes that are each other's best hit when different genomes are compared⁴⁸. All-against-all BLAST alignments were performed using the following parameters: ≥ 90% coverage, ≥ 90% similarity and E value cut-off of 1e-10. A Python algorithm was applied over BLAST results to identify the BBHs⁴⁹. Core, accessory, and unique genomes were analyzed using a MySQL database created with the data generated in previous steps. The CDSs were classified into Clusters of Orthologous Groups (COG) functional categories using the eggNOG-Mapper web application⁵⁰. The identification of genomic islands was carried out by IslandViewer⁵¹. Insertion sequences were identified using the ISSaga web application⁵². Presence of a CRISPR-Cas system was assessed by the CRISPRCasFinder web application⁵³. In order to detect genes involved with antimicrobial resistance mechanisms, protein sequences of the CCBH4851 genome were used to perform BLAST searches against CARD using the Resistance Gene Identifier web application⁵⁴. Also, protein sequences were searched against the Virulence Factors Database of Pathogenic Bacteria⁵⁵ using BLASTP with ≥ 75% coverage, ≥ 50% identity and E

value cut-off of $1e-10$. Regulatory proteins were predicted using the Predicted Prokaryotic Regulatory Proteins web application⁵⁶. Detection of SNPs was performed using the Snippy software based on the alignment of unassembled reads with the *P. aeruginosa* PAO1 genome sequence using default parameters⁹. The hypothetical effect of non-synonymous variants were analyzed using the HOPE web application⁵⁷.

Phylogenomic analysis. Genome sequences of CCBH4851 and strains listed in Table 4 were used to perform whole-genome phylogeny analysis. The analysis was performed by CSI Phylogeny web application using default parameters⁵⁸. A core genome phylogeny tree was inferred and edited with FigTree (available at <http://tree.bio.ed.ac.uk/software/figtree/>).

Antimicrobial susceptibility testing. Antimicrobial susceptibility testing was performed according to the EUCAST guidelines using the broth microdilution method¹¹. Four major antimicrobial classes were assessed: aminoglycosides (gentamicin), quinolones (ciprofloxacin), beta-lactams (imipenem) and lipopeptides (polymyxin E or colistin). The MIC was determined as the lowest concentration of antimicrobial agent that prevented visible growth of the microorganism. Classification as susceptible, intermediate or resistant was based on EUCAST criteria⁵⁹. Briefly, resistance breakpoints for ciprofloxacin, imipenem and colistin are $>0.5 \mu\text{g/mL}$, $>4 \mu\text{g/mL}$ and $>2 \mu\text{g/mL}$, respectively. Gentamicin has an “IE” status, meaning there is no sufficient evidence that the organism is a good target for therapy with the agent, MIC may be reported but without an S, I or R assignment. For the broth microdilution method, a total of six biological replicates with two technical replicates each were performed divided in three independent experiments. With exception of colistin susceptibility testing, that showed growth in skipped wells with no pattern, the same visual reading was observed in all experiments.

Motility assays. Motility assays were performed as described previously with slight modifications⁶⁰. Luria–Bertani (LB) agar was used as growth medium. Swimming plates (0.3% agar) were inoculated with a sterile toothpick and incubated for 16 h at 37°C . Swarming plates (0.6% agar) were inoculated with a sterile toothpick and the plates were incubated for 24 h at 37°C . For twitching, cells were inoculated with a toothpick through a LB layer (1% agar) to the bottom of the petri dish. Plates were incubated for 24 h at 37°C . After incubation, the agar was discarded, the plates were washed to remove unattached cells, and stained with crystal violet (1% [wt/vol] solution). All motilities were measured as the circular zone formed by bacterial migration away from the inoculation point. Additionally, the stain in the twitching plates was solubilized in ethanol (95% [vol/vol] solution) and the absorbance was measured at 570 nm using an ethanol solution as blank. With exception of the stain quantification in the twitching plates, at least three biological replicates with three technical replicates each were performed to three independent experiments. For stain quantification in the twitching plates, a total of five biological replicates with two technical replicates each were performed divided in two independent experiments.

Biofilm formation assay. The ability of *P. aeruginosa* to form biofilms was assessed as described previously⁶¹. Briefly, cells were inoculated in 96-well plates containing LB or Mueller-Hinton medium. Plates were incubated for 20–24 h at 37°C . After incubation, the medium was discarded, the plates were washed with water to remove unattached cells and stained with crystal violet (1% [wt/vol] solution). Then, the crystal violet was solubilized in ethanol (95% [vol/vol] solution), and the absorbance was measured at 570 nm using an ethanol solution as blank. *P. aeruginosa* PAO1 was used as positive control. A total of seven biological replicates with at least three technical replicates each were performed divided in two independent experiments.

Statistical analysis. Statistics was performed in GraphPad Prism software. Student’s t test was used to test whether the results of phenotypic experiments were significantly different (P value < 0.05). Data were expressed as mean \pm standard error of the mean. In the figures, * equals P value < 0.05 , ** equals P value < 0.01 , *** equals P value < 0.001 and **** equals P value < 0.0001 .

Data availability

The genome sequence and annotation data generated in this study have been submitted to the GenBank database (<https://www.ncbi.nlm.nih.gov/genbank/>) under accession number CP021380.

Received: 13 March 2020; Accepted: 16 July 2020

Published online: 06 August 2020

References

- Hong, D. J. *et al.* Epidemiology and characteristics of metallo-beta-lactamase-producing *Pseudomonas aeruginosa*. *Infect. Chemother.* **47**, 81–97. <https://doi.org/10.3947/ic.2015.47.2.81> (2015).
- Sawa, T., Kooguchi, K. & Moriyama, K. Molecular diversity of extended-spectrum β -lactamases and carbapenemases, and antimicrobial resistance. *J. Intensive Care* **8**, 13. <https://doi.org/10.1186/s40560-020-0429-6> (2020).
- Tooke, C. L. *et al.* β -Lactamases and β -Lactamase Inhibitors in the 21st Century. *J. Mol. Biol.* **431**, 3472–3500. <https://doi.org/10.1016/j.jmb.2019.04.002> (2019).
- Agência Nacional de Vigilância Sanitária. Boletim segurança do paciente e qualidade em serviços de saúde no. 17: Avaliação dos indicadores nacionais das infecções relacionadas à assistência à saúde (IRAS) e resistência microbiana do ano de 2017 (2019).
- Escandón-Vargas, K., Reyes, S., Gutiérrez, S. & Villegas, M. V. The epidemiology of carbapenemases in Latin America and the caribbean. *Expert Rev. Anti-Infective Ther.* **15**, 277–297. <https://doi.org/10.1080/14787210.2017.1268918> (2017).
- Silveira, M., Albano, R., Asensi, M. & Assef, A. P. C. The draft genome sequence of multidrug-resistant *Pseudomonas aeruginosa* strain CCBH4851, a nosocomial isolate belonging to clone SP (ST277) that is prevalent in Brazil. *Memórias do Inst. Oswaldo Cruz* **109**, 1086–1087. <https://doi.org/10.1590/0074-0276140336> (2014).

7. Nascimento, A. P. B. *et al.* Intracolon genome stability of the Metallo- β -lactamase SPM-1-producing *Pseudomonas aeruginosa* ST277, an Endemic Clone Disseminated in Brazilian Hospitals. *Front. Microbiol.* **7**, 1946. <https://doi.org/10.3389/fmicb.2016.01946> (2016).
8. Silveira, M. C., Albano, R. M., Asensi, M. D. & Carvalho-Assef, A. P. D. Description of genomic islands associated to the multidrug-resistant *Pseudomonas aeruginosa* clone ST277. *Infect. Genet. Evol.: J. Mol. Epidemiol. Evolut. Genet. Infect. Dis.* **42**, 60–65. <https://doi.org/10.1016/j.meegid.2016.04.024> (2016).
9. Seemann, T. Snippy: Rapid haploid variant calling and core genome alignment (2019).
10. Roy, P. H. *et al.* Complete genome sequence of the multiresistant taxonomic outlier *Pseudomonas aeruginosa* PA7. *PLoS ONE* **5**, e8842. <https://doi.org/10.1371/journal.pone.0008842> (2010).
11. The European Committee on Antimicrobial Susceptibility Testing. EUCAST reading guide for broth microdilution (2019).
12. Silby, M. W., Winstanley, C., Godfrey, S. A. C., Levy, S. B. & Jackson, R. W. *Pseudomonas* genomes: diverse and adaptable. *FEMS Microbiol. Rev.* **35**, 652–680. <https://doi.org/10.1111/j.1574-6976.2011.00269.x> (2011).
13. Shu, J.-C. *et al.* Development of carbapenem resistance in *Pseudomonas aeruginosa* is associated with OprD polymorphisms, particularly the amino acid substitution at codon 170. *J. Antimicrob. Chemother.* **72**, 2489–2495. <https://doi.org/10.3947/ic.2015.47.2.810> (2017).
14. Quale, J., Bratu, S., Gupta, J. & Landman, D. Interplay of efflux system, ampC, and oprD expression in carbapenem resistance of *Pseudomonas aeruginosa* clinical isolates. *Antimicrob. Agents Chemother.* **50**, 1633–1641. <https://doi.org/10.3947/ic.2015.47.2.811> (2006).
15. Poole, K. Aminoglycoside resistance in *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* **49**, 479–487. <https://doi.org/10.3947/ic.2015.47.2.812> (2005).
16. Chung, J. C. S. *et al.* Genomic variation among contemporary *Pseudomonas aeruginosa* isolates from chronically infected cystic fibrosis patients. *J. Bacteriol.* **194**, 4857–4866. <https://doi.org/10.3947/ic.2015.47.2.813> (2012).
17. Mann, E. E. & Wozniak, D. J. *Pseudomonas* biofilm matrix composition and niche biology. *FEMS Microbiol. Rev.* **36**, 893–916. <https://doi.org/10.3947/ic.2015.47.2.814> (2012).
18. Haghi, F. *et al.* Diversity of virulence genes in multidrug resistant *Pseudomonas aeruginosa* isolated from burn wound infections. *Microb. Pathog.* **115**, 251–256. <https://doi.org/10.3947/ic.2015.47.2.815> (2018).
19. Persat, A., Inclan, Y. F., Engel, J. N., Stone, H. A. & Gitai, Z. Type IV pili mechanochemically regulate virulence factors in *Pseudomonas aeruginosa*. *Proc. Nat. Acad. Sci. USA* **112**, 7563–7568. <https://doi.org/10.3947/ic.2015.47.2.816> (2015).
20. Köhler, T., Curty, L. K., Barja, F., van Delden, C. & Pechère, J.-C. Swarming of *Pseudomonas aeruginosa* is dependent on cell-to-cell signaling and requires Flagella and Pili. *J. Bacteriol.* **182**, 5990–5996 (2000).
21. O’Toole, G. A. & Kolter, R. Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Molecular Microbiology* **30**, 295–304. <https://doi.org/10.1046/j.1365-2958.1998.01062.x> (1998).
22. Hwang, W. & Yoon, S. S. Virulence characteristics and an action mode of antibiotic resistance in multidrug-resistant *Pseudomonas aeruginosa*. *Sci. Rep.* **9**, 1–15. <https://doi.org/10.3947/ic.2015.47.2.817> (2019).
23. Winstanley, C., O’Brien, S. & Brockhurst, M. A. *Pseudomonas aeruginosa* evolutionary adaptation and diversification in cystic fibrosis chronic lung infections. *Trends Microbiol.* **24**, 327–337. <https://doi.org/10.1016/j.tim.2016.01.008> (2016).
24. Kung, V. L., Ozer, E. A. & Hauser, A. R. The accessory genome of *Pseudomonas aeruginosa*. *Microbiol. Mol. Biol. Rev.: MMBR* **74**, 621–641. <https://doi.org/10.3947/ic.2015.47.2.818> (2010).
25. Lewis, K. Persister cells. *Annu. Rev. Microbiol.* **64**, 357–372. <https://doi.org/10.3947/ic.2015.47.2.819> (2010).
26. Liu, Y. *et al.* Structural insights into the transcriptional regulation of HigBA toxin-antitoxin system by antitoxin HigA in *Pseudomonas aeruginosa*. *Front. Microbiol.* **10**, 3158. <https://doi.org/10.1186/s40560-020-0429-60> (2019).
27. Siguier, P., Gourbeyre, E. & Chandler, M. Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol. Rev.* **38**, 865–891. <https://doi.org/10.1186/s40560-020-0429-61> (2014).
28. van Belkum, A. *et al.* Phylogenetic distribution of CRISPR-Cas systems in antibiotic-resistant *Pseudomonas aeruginosa*. *mBio* **6**, e01796–01715. <https://doi.org/10.1128/mBio.01796-15> (2015).
29. Gophna, U. *et al.* No evidence of inhibition of horizontal gene transfer by CRISPR-Cas on evolutionary timescales. *ISME J.* **9**, 2021. <https://doi.org/10.1186/s40560-020-0429-62> (2015).
30. Maseda, H., Saito, K., Nakajima, A. & Nakae, T. Variation of the mexT gene, a regulator of the MexEF-oprN efflux pump expression in wild-type strains of *Pseudomonas aeruginosa*. *FEMS Microbiol. Lett.* **192**, 107–112. <https://doi.org/10.1186/s40560-020-0429-63> (2000).
31. Tian, Z.-X. *et al.* MexT modulates virulence determinants in *Pseudomonas aeruginosa* independent of the MexEF-OprN efflux pump. *Microb. Pathog.* **47**, 237–241. <https://doi.org/10.1186/s40560-020-0429-64> (2009).
32. López-Causapé, C., Cabot, G., Del Barrio-Tofiño, E. & Oliver, A. The versatile mutational resistome of *Pseudomonas aeruginosa*. *Front. Microbiol.* **9**, 685. <https://doi.org/10.1186/s40560-020-0429-65> (2018).
33. Ramanathan, B. *et al.* Next generation sequencing reveals the antibiotic resistant variants in the genome of *Pseudomonas aeruginosa*. *PLoS ONE* **12**, e0182524. <https://doi.org/10.1186/s40560-020-0429-66> (2017).
34. Rodriguez-Martinez, J.-M., Poiriel, L. & Nordmann, P. Extended-spectrum cephalosporinases in *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* **53**, 1766–1771. <https://doi.org/10.1186/s40560-020-0429-67> (2009).
35. Zimin, A. V. *et al.* The MaSuRCA genome assembler. *Bioinformatics* **29**, 2669–2677. <https://doi.org/10.1186/s40560-020-0429-68> (2013).
36. Otto, T. D., Dillon, G. P., Degraeve, W. S. & Berriman, M. RATT: rapid annotation transfer tool. *Nucleic Acids Res.* **39**, e57. <https://doi.org/10.1186/s40560-020-0429-69> (2011).
37. Besemer, J., Lomsadze, A. & Borodovsky, M. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* **29**, 2607–2618 (2001).
38. Lagesen, K. *et al.* RNAMmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **35**, 3100–3108. <https://doi.org/10.1093/nar/gkm160> (2007).
39. Lowe, T. M. & Chan, P. P. tRNAscan-SE on-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res.* **44**, W54–W57. <https://doi.org/10.1093/nar/gkw413> (2016).
40. Pruitt, K. D., Tatusova, T. & Maglott, D. R. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **35**, D61–D65. <https://doi.org/10.1093/nar/gkl842> (2007).
41. Tatusov, R. L., Galperin, M. Y., Natale, D. A. & Koonin, E. V. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **28**, 33–36 (2000).
42. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–462. <https://doi.org/10.1093/nar/gkv1070> (2016).
43. The UniProt Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **45**, D158–D169. <https://doi.org/10.1093/nar/gkw1099> (2017).
44. Mitchell, A. L. *et al.* InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* **47**, D351–D360. <https://doi.org/10.1093/nar/gky1100> (2019).
45. Claudel-Renard, C., Chevalet, C., Faraut, T. & Kahn, D. Enzyme-specific profiles for genome annotation: PRIAM. *Nucleic Acids Res.* **31**, 6633–6639 (2003).
46. Rutherford, K. *et al.* Artemis: sequence visualization and annotation. *Bioinformatics (Oxford, England)* **16**, 944–945 (2000).

47. Krzywinski, M. *et al.* Circos: an information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645. <https://doi.org/10.1101/gr.092759.109> (2009).
48. Overbeek, R., Fonstein, M., D'Souza, M., Pusch, G. D. & Maltsev, N. The use of gene clusters to infer functional coupling. *Proc. Natl. Acad. Sci.* **96**, 2896–2901. <https://doi.org/10.1073/pnas.96.6.2896> (1999).
49. Medeiros Filho, F., do Nascimento, A. P. B., Dos Santos, M. T., Carvalho-Assef, A. P. D. & da Silva, F. A. B. Gene regulatory network inference and analysis of multidrug-resistant *Pseudomonas aeruginosa*. *Memorias Do Instituto Oswaldo Cruz* **114**, e190105. <https://doi.org/10.1590/0074-02760190105> (2019).
50. Huerta-Cepas, J. *et al.* Fast genome-wide functional annotation through orthology assignment by eggno-mapper. *Mol. Biol. Evol.* **34**, 2115–2122. <https://doi.org/10.1093/molbev/msx148> (2017).
51. Bertelli, C. *et al.* Islandviewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res.* **45**, W30–W35. <https://doi.org/10.1093/nar/gkx343> (2017).
52. Varani, A. M., Siguier, P., Gourbeyre, E., Charneau, V. & Chandler, M. ISSaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. *Genome Biol.* **12**, R30. <https://doi.org/10.1186/gb-2011-12-3-r30> (2011).
53. Couvin, D. *et al.* CRISPRCasFinder, an update of CRISPRFinder, includes a portable version, enhanced performance and integrates search for cas proteins. *Nucleic Acids Res.* **46**, W246–W251. <https://doi.org/10.1093/nar/gky425> (2018).
54. Jia, B. *et al.* CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **45**, D566–D573. <https://doi.org/10.1093/nar/gkw1004> (2017).
55. Liu, B., Zheng, D., Jin, Q., Chen, L. & Yang, J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res.* **47**, D687–D692. <https://doi.org/10.1093/nar/gky1080> (2019).
56. Barakat, M., Ortet, P. & Whitworth, D. E. P2RP: a web-based framework for the identification and analysis of regulatory proteins in prokaryotic genomes. *BMC Genom.* **14**, 269. <https://doi.org/10.1186/1471-2164-14-269> (2013).
57. Venselaar, H., Te Beek, T. A. H., Kuipers, R. K. P., Hekkelman, M. L. & Vriend, G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinform.* **11**, 548. <https://doi.org/10.1186/1471-2105-11-548> (2010).
58. Kaas, R. S., Leekitcharoenphon, P., Aarestrup, F. M. & Lund, O. Solving the problem of comparing whole bacterial genomes across different sequencing platforms. *PLoS ONE* **9**, e104984. <https://doi.org/10.1371/journal.pone.0104984> (2014).
59. The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters (2020).
60. Déziel, E., Comeau, Y. & Villemur, R. Initiation of biofilm formation by *Pseudomonas aeruginosa* 57rp correlates with emergence of hyperpilated and highly adherent phenotypic variants deficient in swimming, swarming, and twitching motilities. *J. Bacteriol.* **183**, 1195–1204. <https://doi.org/10.1128/JB.183.4.1195-1204.2001> (2001).
61. O'Toole, G. A. Microtiter dish biofilm formation assay. *Journal of Visualized Experiments: JoVE*. <https://doi.org/10.3791/2437> (2011).
62. Lee, J.-Y. & Ko, K. S. Mutations and expression of PmrAB and PhoPQ related with colistin resistance in *Pseudomonas aeruginosa* clinical isolates. *Diagn. Microbiol. Infect. Dis.* **78**, 271–276. <https://doi.org/10.1016/j.diagmicrobio.2013.11.027> (2014).

Acknowledgements

This work was supported by Programa Fiocruz de Fomento à Inovação - Inova Fiocruz (VPPCB-007-FIO-18-2-29) and Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (E-26/202.554/2019). This study was financed in part by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. Hermes Senger is supported by Fundação de Apoio à Pesquisa do Estado de São Paulo (2018/00452-2).

Author Contributions

A.N. designed the study, performed computational analyses and biological experiments, and wrote the manuscript. H.P. performed biological experiments. F.F., H.S., H.S. and R.A. provided computational support. L.A., M.S., A.C., F.S. supervised the study. All authors provided critical feedback and reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-69944-6>.

Correspondence and requests for materials should be addressed to A.P.B.N. or F.A.B.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.