Gene regulatory network inference and analysis of multidrug-resistant *Pseudomonas aeruginosa*

Fernando Medeiros Filho[1], Ana Paula Barbosa do Nascimento[1] ✉, Marcelo Trindade dos Santos[2], Ana Paula D'Alincourt Carvalho-Assef[3], and Fabricio Alves Barbosa da Silva[1]

[1]Programa de Computação Científica, Fundação Oswaldo Cruz, Rio de Janeiro, RJ, Brasil, CEP 21040-900
[2]Laboratório Nacional de Computação Científica, Petrópolis, RJ, Brasil
[3]Laboratório de Pesquisa em Infecção Hospitalar, Fundação Oswaldo Cruz, Rio de Janeiro, RJ, Brasil

Abstract

Background **- Healthcare-associated infections caused by bacteria such as *Pseudomonas aeruginosa* are a major public health problem worldwide. Gene regulatory networks computationally represent the interaction between regulatory genes and their targets, an important approach to understand bacterial behavior and to provide novel ways of overcoming scientific challenges, including the identification of potential therapeutic targets and the development of new drugs.**

Objectives **- Our goal in this manuscript is to present a reconstruction of multidrug-resistant *P. aeruginosa* gene regulatory network and to analyze its topological properties.**

Methods **- The methodology was based on gene orthology inference by the reciprocal best hit method. We used the genome of *P. aeruginosa* CCBH4851 as the basis of the reconstruction process. This multidrug-resistant strain is representative of an endemic outbreak in Brazilian territory belonging to ST277.**

Findings **- As the main finding, we obtained a network with a larger number of regulatory genes, target genes and interactions compared to previous work. Topological analysis results are accordant to complex networks representative of biological processes.**

Main conclusions **- The network properties are consistent with *P. aeruginosa* biological features. To the best of our knowledge, the *P. aeruginosa* gene regulatory network presented in this manuscript is the most complete version available to date.**

Pseudomonas aeruginosa **| gene regulatory network | multidrug resistance**
Correspondence: ***ana.pbn@gmail.com***

Introduction

Healthcare-associated infections (HAI) are one of the major public health problems worldwide, increasing the morbidity and mortality rates of hospitalized individuals. HAI infections are often caused by multidrug-resistant (MDR) bacteria such as *Pseudomonas aeruginosa*, especially in immunocompromised patients. In Brazil, *P. aeruginosa* was ranked as the fifth most common causative agent of HAI in patients hospitalized in adult and pediatric intensive care units, and nearly 35% of the reported strains are resistant to carbapenems, a class of antibiotics widely used in *P. aeruginosa* infections therapy[(1)]. In fact, individuals infected with MDR *P. aeruginosa* clones have a higher mortality rate (44.6%) compared to those with non-MDR infection (24.8%)[(2)].

P. aeruginosa is a versatile pathogen that cause several types of infections affecting the lower respiratory tract, skin, urinary tract, eyes, leading to bacteremia, endocarditis, and other complications. *P. aeruginosa* infections are difficult to treat as the therapeutic choices has becoming ever more limited. In addition to being intrinsically resistant to a broad range of antibiotics, this bacterium can turn multidrug-resistant through the acquisition of new resistance mechanisms[(3,4)].

In 2000, the genome sequence of *P. aeruginosa* PAO1 strain was published providing data concerning its genome sequence, genetic complexity and ecological versatility[(5)]. The PAO1 strain is sensitive to most clinically used antimicrobial agents and has been extensively studied ever since.

In 2003, the first clinical isolate of a MDR *P. aeruginosa* carrying the carbapenemase gene named blaSPM-1 was identified in Brazilian territory. The SPM-1 protein is a metallo-β-lactamase that confers resistance to almost all classes of beta-lactams[(6)]. Most of SPM-producing isolates belong to clone ST277, as indicated through multilocus sequence typing (MLST). This clone has been associated with hospital outbreaks in several Brazilian states, and have already been found in hospital sewage and rivers[(7–9)].

Over the past years, researchers have applied mathematical methods in order to generate computational models used to study several organisms' behavior, contributing to the development of new products, improvement and acceleration of existing health policies, and research of novel ways of overcoming scientific challenges. This approach is often based on the construction of biological networks and pathway analysis comprising gene regulatory, metabolic, signal transduction and/or protein-protein interactions[(10)].

A gene regulatory network (GRN) is a collection of transcription factors that interact with each other and with other molecules in the cell to regulate the levels of mRNA and protein expression. In 2011, Galán-Vásquez *et al.*[(11)] published the first *P. aeruginosa* GRN, analyzing its main topological properties and interactions between its regulatory components.

In this work, we reconstructed the *P. aeruginosa* GRN of a MDR strain, including as much curated biological data as available to date. We used as model the *P. aeruginosa* CCBH4851, a strain representative of an endemic outbreak in Brazilian territory caused by clones belonging to the ST277. This strain shows resistance to all antimicrobials of

clinical importance except for polymyxin B, has several mechanisms of resistance and mobile genetic elements[12]. The implications of the choice of a MDR strain as the basis of the GRN reconstruction presented in this manuscript are discussed. In addition, we analyzed GRN topological properties, characterizing regulators, target genes, transcription factors auto-activation mechanisms, influential genes and network motifs.

Materials and Methods

Bacterial strains - We present a gene regulatory network reconstruction for *P. aeruginosa* CCBH4851. This strain is deposited in the Culture Collection of Hospital-Acquired Bacteria (CCBH) located at the Laboratório de Pesquisa em Infecção Hospitalar - Instituto Oswaldo Cruz/Fiocruz (WDCM947; 39 CGEN022/2010) and its genome is available in the GenBank database (accession number JPSS00000000)[12]. In order to perform the orthology analysis, we used *P. aeruginosa* PAO1[5], *P. aeruginosa* PA7[13] and *P. aeruginosa* UCBPP-PA14 (PA14)[14] as reference strains.

Orthology-based model generation - Fitch[15] defines orthologs as genes diverging after a speciation event, sharing a common ancestor. The most common approach to find orthologs is the reciprocal best hits (RBH) method[16]. The regulatory interaction between a transcription factor (TF) and a target gene (TG) belonging to *P. aeruginosa* PAO1, *P. aeruginosa* PA14 and *P. aeruginosa* PA7 strains were propagated to *P. aeruginosa* CCBH4851 reconstructed network if both TF and TG form RBHs.
The criteria to define an orthology relationship is the existence of RBHs between the two genomes. A pair of genes (a, a') of the genomes A and A', respectively, are considered orthologs if it is also an RBH, that is, if aligning the sequence of a, against the gene list of A' we obtain a' as the best alignment, and if aligning the sequence of a' against genome-wide gene list A, we obtain a as the best hit. Once the complete set of genomes RBHs between A and A' is obtained, a regulatory interaction between a TF (the gene a) and a TG (the gene b) was propagated from the reference network to CCBH4851, if both genes have their respective RBHs in the CCBH4851 genome. The propagation of a regulatory interaction a-b from the reference genome A holds if there exists in genome A' a pair a'-b' such that both (a, a') and (b, b') are RBH pairs.
This test for the propagation of regulatory interactions was performed with all interactions known in PAO1, PA7 and PA14. For the alignments, the BLASTP program was used. Figure 1 presents an overview of the reconstruction processes.

Identification of RBHs - An algorithm was implemented in the Python language to automate and generate the list of RBHs in tabular format. The last step was to identify and separate the regulators and target genes in a single table, extending the work done by Galán-Vásquez *et al.*[11].

Data integration - The data integration process brings together biological information from all strains with the aim of organizing biological knowledge. The final network table is available as supplementary material. This table is organized into 6 columns: "Regulatory gene", "Ortholog of the regulatory gene", "Target gene", "Ortholog of the target gene", "Mode of regulation" and "Reference". The first column lists regulatory genes of *P. aeruginosa* CCBH4851, the second column contains orthologous of regulatory genes in the reference strain (PAO1, PA7 or PA14), the third column refers to the target gene in CCBH4851, the fourth column lists orthologous of target genes in the reference strain, the fifth column describes the mode of regulation, and the sixth column indicates the corresponding reference.

Curation process - Our group has developed a web application to support the curation of biological networks. This web application, called CurSystem[17] (available from: http://pseudomonas.procc.fiocruz.br:8185/CurSystem) provides support for distributed, asynchronous interaction among specialists. Through this tool the authors of this article could select specific gene interactions, discuss their main peculiarities and determine if they would be part of the network or not. This stage was fundamental to exclude doubtful biological information from the network.

Network generation and computational analysis - The R language and Rstudio free software were used in network generation and computational analysis[18]. Analysis of degree, centrality, clustering coefficient, connectivity, cycles, paths and hierarchical levels were made according to previous works[11,19]. We used the dplyr, tibble, readr, igraph and scales packages.

All data and code are available as supplementary files.

Results

General features of the gene regulatory network - The *P. aeruginosa* network reconstruction resulted in a total of 1049 genes, of which 42 behave as regulatory genes, 96 both as regulatory and target genes (*i.e.* a TF is influenced by another TF in the network), and 911 only as target genes. We found 1579 regulatory interactions between regulators and their target genes. Altogether, the genes represent approximately 16.52% of the *P. aeruginosa* CCBH4851 genome used as model in our work. Despite the apparent small coverage, we have included most transcription factors with described function among the 138 regulators in the *P. aeruginosa* CCBH4851 network. The number of regulatory genes, target genes and interactions represent an increase of 44.92%, 31.47% and 35.40% compared to previous work, respectively[11].
Network enrichment was not the only aspect observed in the *P. aeruginosa* CCBH4851 gene regulatory network reconstruction. As we based our reconstruction on the RBH method, comparing the CCBH4851 genome annotation with reference strains, we could not infer an orthology

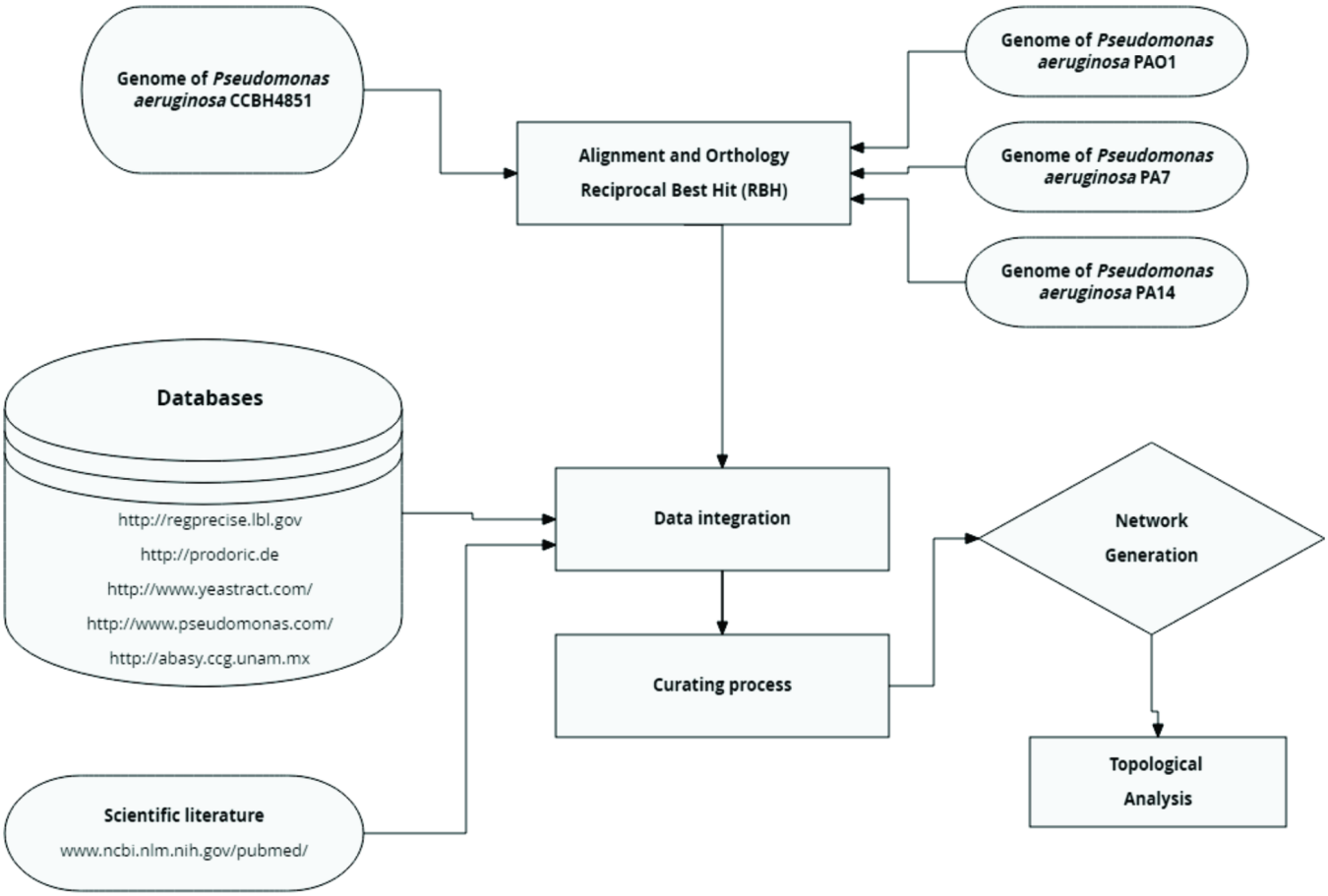

Fig. 1. Overview of general strategy for reconstruction of the *P. aeruginosa* GRN. The process started with the alignment of *P. aeruginosa* CCBH4851 genome and the three reference strains. Next, the RBH method was applied and the resulting genes were compared against gene regulatory databases listed in the "Databases" box. Then, data obtained from these databases were integrated and submitted to the curation process, which aims to solve network inconsistencies. Finally, the GRN was generated and its topology was analyzed.

relationship for some genes, particularly *oprD* and *mexZ*, which are genes involved in antibiotic resistance mechanisms. The curation process revealed that these genes were either fully absent or annotated as pseudogenes in CCBH4851. A pseudogene is a DNA sequence that resembles a gene from the reference genome, but has suffered modifications such as point mutations, insertions, deletions, premature stop codons or frameshifts, being impossible to attest if its product is still functional in the target organism without proper experimentation. The lack of orthology resulted in the exclusion of these genes from *P. aeruginosa* CCBH4851 GRN. In addition, some notations were kept as listed in the previous network[11], databases and/or scientific literature used. For example, *ihf* (for integration host factor) represents not a single gene, but a complex composed of the product of *himA* and *himD* genes that together act as a TF of several target genes. On the other hand, regulatory systems such as quorum sensing or two component systems are often formed by a pair of genes, but only one of them is able of binding in the promoter region, however both genes are listed as regulatory genes. This way we could maintain an equivalent notation to previous works[11].

Basic network topological analysis: number of vertices, number of edges and density - We identified 1579 edges in the CCBH4851 network. These interactions were classified in four types: activation ("+"), repression ("-"), dual ("d") (when the regulatory gene can act as an activator or repressor, depending on certain conditions) and unknown ("?"). Figure 2 is an illustration of the CCBH4851 GRN. Network density is a measure of interconnectivity between vertices. It is the ratio of the number of edges in the network by the maximum possible number of edges. The regulatory network of the CCBH4851 strain has a density (1.42e-03) which is slightly lower than the observed density for the PAO1 strain (2.15e-03) but maintains the same order of magnitude[11,20].

The degree k(i) of a vertex i is defined as its number of edges. Edges in directed networks can be of two types: they can "depart" from or "arrive" at node i, defining its "incoming" and "outgoing" degrees respectively. It was observed for the CCBH4851 GRN that, on average, each vertex is connected to 3 other vertices, same value reported for PAO1 GRN. Figure 3 illustrates incoming (A-B) and outgoing (C-D) degree distributions for the CCBH4851 GRN.

Scale-free is a common topology classification associated

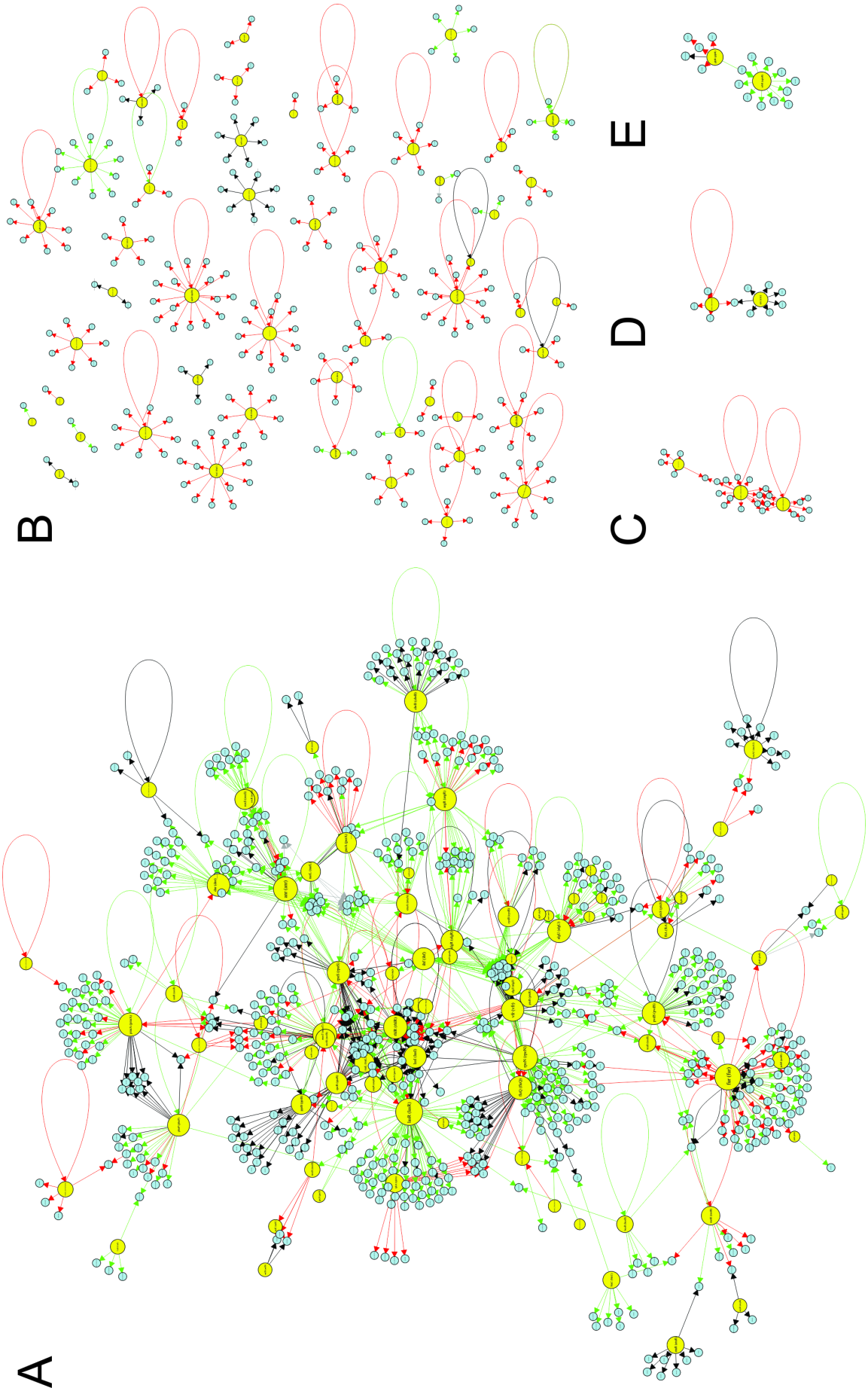

Fig. 2. Visualization of the *P. aeruginosa* CCBH4851 gene regulatory network. The yellow circles represent regulatory genes, light blue circles represent target genes, black lines indicate an unknown mode of regulation, green lines indicate activation, red lines indicate repression and gray lines a dual mode of regulation. (A) The GRN large highly connected network component; (B) All regulatory and target genes that have no connections with the component depicted in A; (C, D, E) Clusters of lower connectivity compared to the component depicted in A. All figures are available with better resolution in the supplementary material.

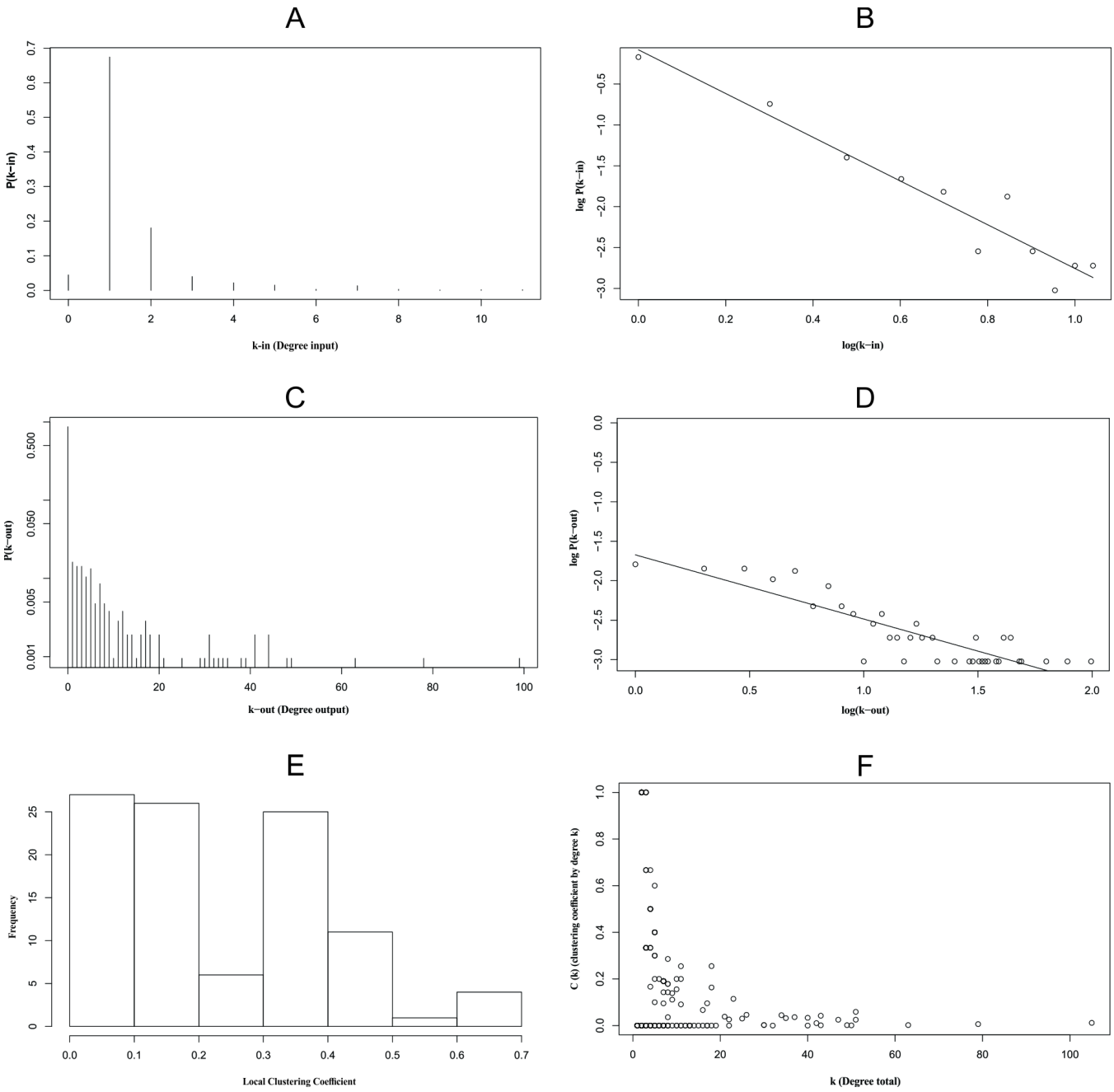

Fig. 3. (A) and (B) Incoming degree distribution of the *P. aeruginosa* CCBH4851 GRN. (C) and (D) Outgoing distribution of the *P. aeruginosa* CCBH4851 GRN. For clarity, the distributions are plotted both on a linear (A, C) and on logarithmic scale (B, D). (E) Local clustering coefficient distribution. (F) Clustering coefficient by degree.

with biological networks, corresponding to complex networks which degree distribution follows a power law. In scale-free networks, most nodes (vertices) have few connections and few nodes have a large number of connections. In this way, scale-free networks are dominated by a relatively small number of high degree nodes called hubs [21].

The degree distribution can be approximated by:

$$P(k) \sim k^{-\gamma} \tag{1}$$

Equation 1 corresponds to a power-law distribution and the exponent γ is its degree exponent [22]. The degree distribution in figures 3B and 3D is shown on double logarithmic axis, and the straight line is consistent to a power-law distribution. For the k-in, the estimated value for γ was 2.89, very close to the value reported by the reference work ($\gamma = 2.717$) [11].

Clustering coefficient distribution - Given a node i with m(i) neighbors in a directed network, the maximum number of edges connecting the elements of this neighborhood is given by $m_{max}(i) = m(i)(m(i)-1)$. The local clustering coefficient C(i) is defined as the ratio between the actual number of edges N(i) occurring in node i neighborhood and $m_{max}(i)$ [23]. We have $C(i)=N(i)/m_{max}(i)$. In GRNs, the local clustering coefficient is interpreted as the interaction

between genes forming regulatory groups. The distribution of local clustering coefficients can be seen in Figure 3E.
On the other hand, the global clustering coefficient is proportional to the number of triangles present in the network, disregarding the directionality of the edges. A triangle is a set of three nodes with at least two connections between them. We can have closed triangles, with three connections within the set, and open triangles, with only two edges. The global clustering coefficient C is the ratio between the number of closed triangles and the total number of triangles (closed or open) in the network. The CCBH4851 network has a global clustering coefficient equal to 2.7e-02.
Another interesting feature to observe is the correlation between the local clustering coefficient C(i) and the degree k(i), as shown by the scatter plot in Figure 3F. The observed correlation is negative, and the figure also shows that the vertices with high degree k correspond to the same vertices with null clustering coefficients, while the vertices that form clusters have low degrees. From this observation, it is confirmed that strongly cohesive groups are exceptions in the network and are formed by small number of genes. These results were obtained for both the CCBH4851 and the previously published *P. aeruginosa* GRN[11].

Connectivity - Network connectivity is a global network parameter that reflects whether the network multiple components are fully connected. Similar to the reference GRN[11], the CCBH4851 network is disconnected, *i.e.* has disjoint gene subsets. There is a total of 55 clusters present in the CCBH4851 network. This large number of clusters is a consequence of the disconnected nature of the network.

Dominant activity and self-regulation - The analysis of the frequency of the different modes of regulation indicated that activation is the predominant type of regulation in the CCBH4851 network, with frequency values very similar to those previously observed for the *P. aeruginosa* GRN[11]. 75.2% of the interactions are of the activation mode, while less than 3% is dual or unknown mode.

Motifs - The existence of cycles or motifs in biological networks is a necessary condition for the existence of multiple stationary states or attractors. The most important motifs, given their biological importance, are cycles with 2 or 3 genes. The motifs of 3 cycles identified as feedforward loops with only activation interactions (also known as coherent loops) were the most abundant in both strains, with 89 representatives in the PAO1 GRN and 218 in the CCBH4851 GRN. Of the triads, the second most abundant was the symmetrical feedforward loop, which totaled 32 in the reference strain and 45 in CCBH4851[11].

Hubs - Identifying the most influential genes in a gene transcription network is a key step in determining therapeutic targets against an infectious agent. One way to identify possible targets is to identify so-called network hubs. Different definitions for the word hub can be applied in the context of complex network theory: one of them is to verify which vertices have the highest k-out degrees in order to identify, in the case of a gene regulatory network, the genes with the greatest influence on target regulation. Table 1 shows the 30 most influential hubs in the *P. aeruginosa* GRN.

Table 1. The 30 most influential hubs of the *P. aeruginosa* CCBH4851 GRN.

GENE	k-out	GENE	k-out
lasR	99	*cbrB*	31
fur	78	*algR*	31
rpoN	63	*ihf*	30
anr	49	*phoP*	29
mexT	48	*qscR*	25
rhlR	44	*cysB*	21
fleQ	44	*exsA*	20
algU	41	*psrA*	20
pmrA	41	*pprB*	18
argR	39	*roxR*	18
pvdS	38	*rsaL*	17
rpoS	35	*roxS*	17
dnr	34	*np20*	17
vfr	33	*narL*	16
lasI	32	PA4851_19380	16

After pinpointing the hubs, we identified whether they are connected (through direct or indirect interactions) or not. We analyzed the interactions among the hubs and observed that only two are not interconnected: *np20* and PA4851_19380 (homologous to PA1520). The remaining hubs have a direct (when a hub affects the regulation of another hub) or indirect (when hubs affect the regulation of the same group of target genes) connection (see Figure 4) to other hubs. Interactions there are not common among hubs were hidden to better visualization in Figure 4.

Discussion

The importance of gene regulation on metabolic, adaptative, pathogenic and antibiotic resistance capabilities is well known. The GRN reconstruction and analysis of a versatile pathogen such as *P. aeruginosa*, in particular when based on a MDR strain, contribute to increase the knowledge of related cellular processes. Multidrug resistance can be conferred by a combination of factors varying according to the antimicrobial class. For instance, carbapenems resistance in *P. aeruginosa* is manly given by mutations in *oprD* and/or presence of MBLs. Mutations or differential expression of efflux system genes are also a contributing factor for both carbapenems and aminoglycosides resistance. However, aminoglycosides and fluoroquinolones resistance can also be provided by other mechanisms, including acquisition of genes through horizontal transfer and punctual mutations, in multiple combinations, when comparing several non-susceptible strains[24]. Data such as gene expression variation, point mutations or genes lacking experimental evidence are not eligible to be included in a gene regulatory network graph. However, we exclude from our network genes such as *oprD* and *mexZ*. Since *oprD* is a

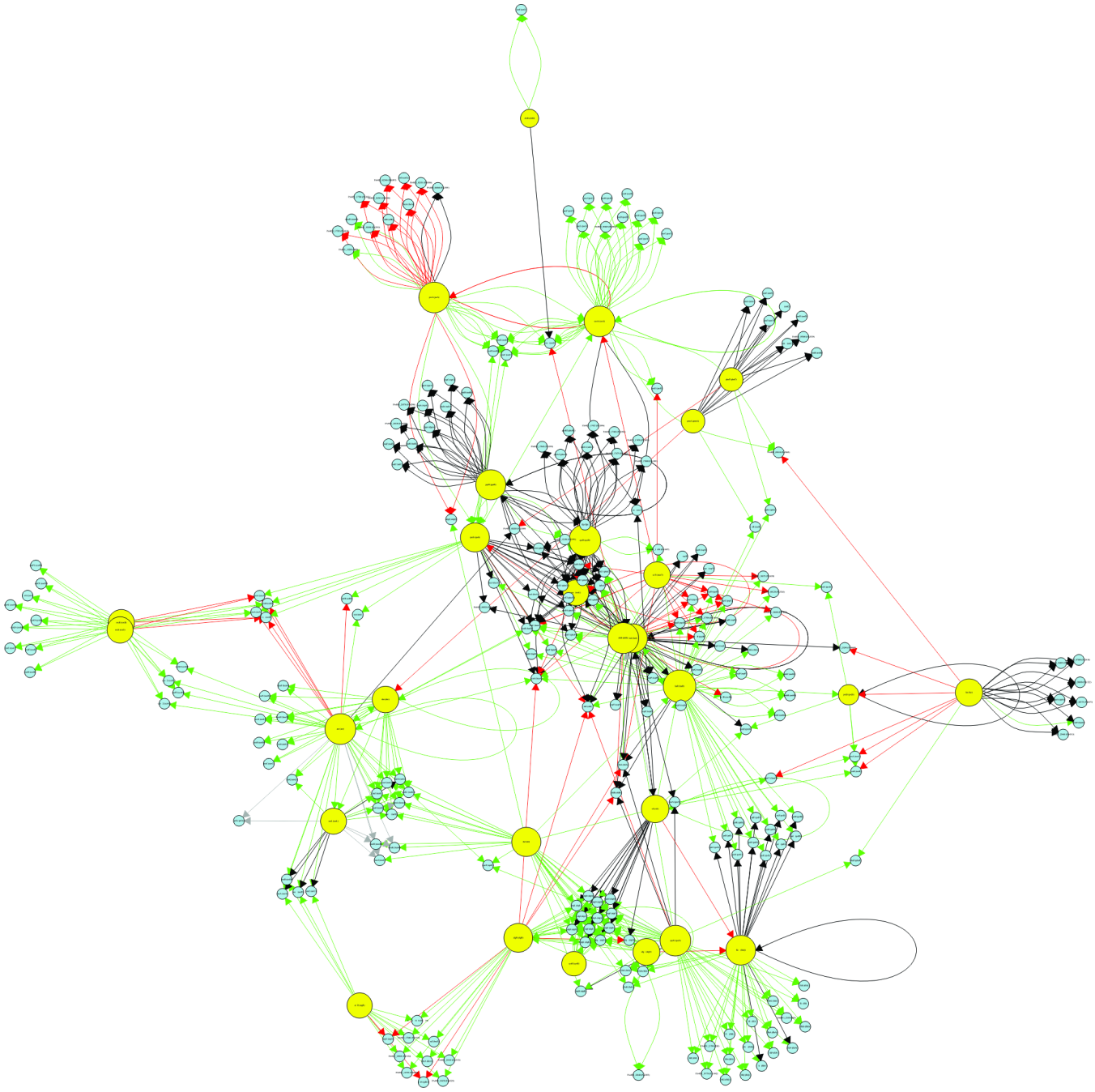

Fig. 4. Connectivity relationships among the 30 most influential hubs of the *P. aeruginosa* CCBH4851 GRN. The yellow circles represent regulatory genes, light blue circles represent target genes, black lines indicate an unknown mode of regulation, green lines indicate activation, red lines indicate repression and gray lines a dual mode of regulation.

target gene, its exclusion has a minor impact in the network topology, but it is extremely important for the cell since *oprD* codifies an outer membrane porin important for the absorption of carbapenems. The lack of OprD leads to low outer membrane permeability. Otherwise, *mexZ* is a regulatory gene and its exclusion from the network results in the exclusion of its node as well as the interactions (edges) with their target genes. The *mexZ* product represses the transcription of *mexX* and *mexY* genes. MexXY are part of an efflux pump system whose overexpression leads to aminoglycoside resistance through the extrusion of this family compounds. We know the MexXY overexpression needs to be experimentally established and cannot be represented in a graph. However, these alterations are common among MDR strains[25], and to have a network comprising these features could impact dynamics simulations designed to assess MDR bacteria behaviors based on *P. aeruginosa* CCBH4851 GRN. In addition, this reconstruction includes additional regulators, target genes, and new interactions described in literature or included in curated databases since the last *P. aeruginosa* GRN publication[11]. Several genes involved in virulence

mechanisms were identified, such those associated to the production of proteases and toxins, antimicrobial activity, iron uptake, antiphagocytosis, adherence and quorum sensing. Not only new nodes and connections were added but previously identified nodes were excluded (by the curation process or lack of homology) and interactions were revisited (due to genes that had regulatory effect recently elucidated). Altogether, these factors influenced directly in the network topological characteristics. The mathematical aspects of network topology *per se* were consistent with the type of network obtained. A concern is to make sure these aspects are consistent with the biological observations. The reconstructed network showed a low-density value compatible with the fact that networks representing natural phenomena often have low density, which is reflected by their structural and dynamic flexibility[20]. The low density observed in the CCBH4851 GRN means that the nodes are not all interconnected. Biologically, in an organism such as *P. aeruginosa* that has an average of 6,000 coding sequences, is not expected that all genes maintain an interaction since they are related to distinct biological process that are not all dependent on each other and are triggered in different growth phases, corroborating this low density. In the same way, the global clustering coefficient and connectivity parameters are affected by these biological behaviors, resulting in the large number of clusters found in the CCBH4851 GRN.

Results related to degree measurements suggest a scale-free network topology for the *P. aeruginosa* CCBH4851 GRN. Construction of several networks representing biological processes reveal the same topological characteristic[22,26]. In fact, *P. aeruginosa* shows a few regulatory genes controlling the expression of many genes, such as those described in CCBH4851 GRN: *lasR*, the main player on quorum-sensing regulatory system; *fur*, the global regulator for iron uptake; *rpoN*, an alternative sigma factor; *mexT*, the regulator of an efflux pump system and several virulence factors; *anr*, responsible for the regulation of anaerobic adaptation processes; and others.

The most common regulatory activity found was activation, with a predominance of positive self-regulation. As seen in other organisms, this mode is important to ensure continuity of biological processes from beginning to end. Adhesion, cell-to-cell signaling, production of virulence and resistance factors, biofilm formation, secretion of toxins, interaction host-pathogen factors are examples of processes that once started must reach a final stage in order to have the desired effect. In fact, we can observe that genes such as *lasR*, *rlhR*, *mexT*, *pvdS*, *anr*, *dnr*, *algU* and others involved in these types of process have demonstrated mostly a positive mode of regulation in our network. On the other hand, negative cycles are important to life-sustaining cyclic processes such as those involved in cell homeostasis. This is the case of metabolic process where we can observe genes such as *recA*, *lexA*, *hutC*, *iscR*, *desT*, *mvat* (although involved in virulence factors biosynthesis, this gene regulates arginine metabolism) and others which negative mode of regulation is the predominant effect[19].

Dominancy of activation mode was also revealed when looking to network motifs. Motifs are patterns of topological structures statistically overrepresented in the network. A common motif often related to transcriptional networks, called coherent feed-forward loop, is abundantly present in the CCBH4851 GRN. These coherent FFLs with three genes are characterized by a gene A that directly activates both gene B and gene C, and indirectly activates C through regulation of B over C. They act as sign-sensitive delays, *i.e.*, a circuit that responds rapidly to step-like stimuli in one direction (ON to OFF), and at a delay to steps in the opposite direction (OFF to ON)[27]. One last characteristic revealed by the topological analysis is the presence of hubs. Hubs are nodes showing a large number of connections, a concept that is inherent of scale-free networks. As expected, CCBH4851 GRN analysis pointed out among the most influential hubs genes such as *lasR*, *fur*, *anr*, *mexT*, *algU*, known to cause great impact in the gene regulatory systems of *P. aeruginosa*. They are involved in resistance, virulence, and pathogenicity mechanisms. LasR, for instance, directly activates the expression of 99 genes. LasR depends on presence and binding of *N*-3-oxo-dodecanoyl-L-homoserine lactone (C12) to act. Once bound, LasR-C12 coordinate the expression of target genes, including many genes encoding virulence factors and cell density[28]. We could observe that even though few hubs remained unconnected, most of the influential genes interact among them. This interaction can be direct as the positive effect of *lasR* on *rlhR* transcription, or indirect when hubs are regulating the same targets, *i.e.* involved in the regulation of the same processes, as *fur* and *algU* both affecting the expression of *phuR* that codifies a member of a heme uptake system to provide host iron acquisition[29,30]. Another example is the regulation of *algU*, *rpoN* and *cysB* over "*alg*" genes, not direct connected but related through the influence in the alginate biosynthesis, important to the mucoid phenotype of *P. aeruginosa* colonies[31]. Direct and indirect interactions reflect the importance of influential genes, not only in specific process but globally. Nevertheless, isolated hubs are equally important. In fact, they are related to process such as zinc uptake (*np20*) and purine metabolism (PA4851_19380), that are fundamental to bacterial survival but can be considered somehow independent of other process and are triggered under particular conditions.

A concept addressed by Csermely[32] is the plasticity of networks. Plastic networks have some interesting characteristics, such as diffuse core, overlapping modules, fewer hierarchies/more loops, large network entropy, and origin dominance, leading to many attractors. Csermely[32] states that biological plastic networks should be attacked by a "central impact" directed at their hubs, bridges and bottlenecks, since if they are attacked on their periphery the effect of the drug will never reach the center of the network due its efficient dissipation. For this reason, topological characteristics as cluster groups, motifs, hubs are important to determine the best approach to disturb a network in a way

to lead the cell to a desired phenotype. This reconstruction of *P. aeruginosa* gene regulatory network can contribute to increase our understanding of this bacterium behavior. As future work, we intend to construct a dynamic model of this network, aiming to help researchers working on experimental drug design and screening, to predict dynamical behaviors in order to have a better understanding of the bacteria lifestyle, also allowing the simulation of normal against stress conditions and eventually leading to the discovery of new potential therapeutic targets and the development of new drugs to combat *P. aeruginosa* infections.

Acknowledgements

The authors would like to acknowledge INOVA-FIOCRUZ, FAPERJ and CAPES for the financial support.

Author's Contribution

FMF performed the GRN reconstruction. APBN and APDCA coordinated the network curation effort. MTS and FABS designed the overall method. All authors have equally participated in the writing of this manuscript.

References

1. Ministério da Saúde. Boletim Segurança do Paciente e Qualidade em Serviços de Saúde nº 16: Avaliação dos indicadores nacionais das Infecções Relacionadas à Assistência à Saúde (IRAS) e Resistência microbiana do ano de 2016. Agência Nacional de Vigilância Sanitária;. Available from: http://portal.anvisa.gov.br/documents/33852/271855/Boletim+Seguran%C3%A7a+do+Paciente+e+Qualidade+em+Servi%C3%A7os+de+Sa%C3%BAde+n%C2%BA+16+-+Avalia%C3%A7%C3%A3o+dos+indicadores+nacionais+das+Infec%C3%A7%C3%B5es+Relacionadas+%C3%A0+Assist%C3%AAncia+%C3%A0+Sa%C3%BAde+%28IRAS%29+e+Resist%C3%AAncia+microbiana+do+ano+de+2016+%28REVISADO%29/e8ec4ea2-1832-489d-8354-0dbc7e3c2f7b.

2. Matos ECOd, Andriolo RB, Rodrigues YC, Lima PDLd, Carneiro ICdRS, Lima KVB. Mortality in patients with multidrug-resistant *Pseudomonas aeruginosa* infections: a meta-analysis.;51(4):415–420.

3. Lister PD, Wolter DJ, Hanson ND. Antibacterial-resistant *Pseudomonas aeruginosa*: clinical impact and complex regulation of chromosomally encoded resistance mechanisms;22(4):582–610.

4. Neves PR, Mamizuka EM, Levy CE, Lincopan N. *Pseudomonas aeruginosa* multirresistente: um problema endêmico no Brasil;47:409–420.

5. Stover CK, Pham XQ, Erwin AL, Mizoguchi SD, Warrener P, Hickey MJ, et al. Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen.;406(6799):959–964.

6. Toleman MA, Simm AM, Murphy TA, Gales AC, Biedenbach DJ, Jones RN, et al. Molecular characterization of SPM-1, a novel metallo-beta-lactamase isolated in Latin America: report from the SENTRY antimicrobial surveillance programme;50(5):673–679.

7. Carvalho APD, Albano RM, de Oliveira DN, Cidade DAdP, Teixeira LM, Marques EdA. Characterization of an epidemic carbapenem-resistant *Pseudomonas aeruginosa* producing SPM-1 metallo-beta-lactamase in a hospital located in Rio de Janeiro, Brazil.;12(2):103–108.

8. Nascimento APB, Ortiz MF, Martins WMBS, Morais GL, Fehlberg LCC, Almeida LGP, et al. Intraclonal Genome Stability of the Metallo-β-lactamase SPM-1-producing *Pseudomonas aeruginosa* ST277, an Endemic Clone Disseminated in Brazilian Hospitals;7.

9. Fontes LC, Neves PR, Oliveira S, Silva KC, Hachich EM, Sato MIZ, et al. Isolation of *Pseudomonas aeruginosa* coproducing metallo-beta-lactamase SPM-1 and 16S rRNA methylase RmtD1 in an urban river.;55(6):3063–3064.

10. Tatarinova TV, Nikolsky Y, editors. Biological Networks and Pathway Analysis. Humana Press;.

11. Galán-Vásquez E, Luna B, Martínez-Antonio A. The Regulatory Network of *Pseudomonas aeruginosa*;1(1):3–3.

12. Silveira M, Albano R, Asensi M, Assef APC. The draft genome sequence of multidrug-resistant *Pseudomonas aeruginosa* strain CCBH4851, a nosocomial isolate belonging to clone SP (ST277) that is prevalent in Brazil;109(8):1086–1087.

13. Roy PH, Tetu SG, Larouche A, Elbourne L, Tremblay S, Ren Q, et al. Complete genome sequence of the multiresistant taxonomic outlier *Pseudomonas aeruginosa* PA7;5(1):e8842.

14. Lee DG, Urbach JM, Wu G, Liberati NT, Feinbaum RL, Miyata S, et al. Genomic analysis reveals that *Pseudomonas aeruginosa* virulence is combinatorial;7(10):R90.

15. Fitch WM. Homology a personal view on some of the problems;16(5):227–231.

16. Kristensen DM, Wolf YI, Mushegian AR, Koonin EV. Computational methods for Gene Orthology inference;12(5):379–391.

17. Ramos TG. Reconstrução da Rede Metabólica da *Pseudomonas aeruginosa* CCBH4851;Available from: https://www.arca.fiocruz.br/handle/icict/29528.

18. RStudio Team. RStudio: Integrated development environment for R;. Available from: http://www.rstudio.com/.

19. Martínez-Antonio A, Janga SC, Thieffry D. Functional organisation of *Escherichia coli* transcriptional regulatory network;381(1):238–247.

20. Bales ME, Johnson SB. Graph theoretic modeling of large-scale semantic networks;39(4):451–464.

21. Amaral LA, Scala A, Barthelemy M, Stanley HE. Classes of small-world networks;97(21):11149–11152.

22. Barabási AL, Bonabeau E. Scale-free networks;288(5):60–69.

23. Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks;393(6684):440–442.

24. Kos VN, Déraspe M, McLaughlin RE, Whiteaker JD, Roy PH, Alm RA, et al. The resistome of *Pseudomonas aeruginosa* in relationship to phenotypic susceptibility;59(1):427–436.

25. Quale J, Bratu S, Gupta J, Landman D. Interplay of efflux system, *ampC*, and *oprD* expression in carbapenem resistance of *Pseudomonas aeruginosa* clinical isolates;50(5):1633–1641.

26. Zhu X, Gerstein M, Snyder M. Getting connected: analysis and principles of biological networks;21(9):1010–1024.

27. Mangan S, Alon U. Structure and function of the feed-forward loop network motif;100(21):11980–11985.

28. Kiratisin P, Tucker KD, Passador L. LasR, a transcriptional activator of *Pseudomonas aeruginosa* virulence genes, functions as a multimer;184(17):4912–4919.

29. Ochsner UA, Johnson Z, Vasil ML. Genetics and regulation of two distinct haem-uptake systems, *phu* and *has*, in *Pseudomonas aeruginosa*;146 (Pt 1):185–198.

30. Firoved AM, Boucher JC, Deretic V. Global Genomic Analysis of AlgU (σE)-Dependent Promoters (Sigmulon) in *Pseudomonas aeruginosa* and Implications for Inflammatory Processes in Cystic Fibrosis;184(4):1057–1064.

31. Ramsey DM, Wozniak DJ. Understanding the control of *Pseudomonas aeruginosa* alginate synthesis and the prospects for management of chronic infections in cystic fibrosis;56(2):309–322.

32. Csermely P. The Wisdom of Networks: A General Adaptation and Learning Mechanism of Complex Systems: The Network Core Triggers Fast Responses to Known Stimuli; Innovations Require the Slow Network Periphery and Are Encoded by Core-Remodeling;40(1).