



OPEN ACCESS

EDITED BY

Sudhir K. Shukla,
Bhabha Atomic Research Centre
(BARC), India

REVIEWED BY

Wanderson Marques Da Silva,
Consejo Nacional de Investigaciones
Científicas y Técnicas (CONICET),
Argentina
Gyan Sahukhal,
University of Arkansas for Medical
Sciences, United States

*CORRESPONDENCE

Ernesto Perez-Rueda
✉ ernesto.perez@iimas.unam.mx
Marisa Fabiana Nicolás
✉ marisa@lncc.br

SPECIALTY SECTION

This article was submitted to
Microbial Physiology and Metabolism,
a section of the journal
Frontiers in Microbiology

RECEIVED 21 September 2022

ACCEPTED 19 December 2022

PUBLISHED 10 January 2023

CITATION

Costa MOC, Nascimento APB,
Martins YC, Santos MT, Figueiredo
AMS, Perez-Rueda E and Nicolás MF
(2023) The gene regulatory network of
Staphylococcus aureus
ST239-SCCmecIII strain Bmb9393 and
assessment of genes associated with
the biofilm in diverse backgrounds.
Front. Microbiol. 13:1049819.
doi: 10.3389/fmicb.2022.1049819

COPYRIGHT

© 2023 Costa, Nascimento, Martins,
Santos, Figueiredo, Perez-Rueda and
Nicolás. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

The gene regulatory network of *Staphylococcus aureus* ST239-SCCmecIII strain Bmb9393 and assessment of genes associated with the biofilm in diverse backgrounds

Maiana de Oliveira Cerqueira e Costa¹,
Ana Paula Barbosa do Nascimento², Yasmin Cortes Martins¹,
Marcelo Trindade dos Santos¹, Agnes Marie de Sá Figueiredo³,
Ernesto Perez-Rueda^{4*} and Marisa Fabiana Nicolás^{1*}

¹Laboratório Nacional de Computação Científica (LNCC), Petrópolis, Brazil, ²Departamento de Análises Clínicas e Toxicológicas, Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, São Paulo, Brazil, ³Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, Unidad Académica Yucatán, Merida, Mexico, ⁴Laboratório de Biologia Molecular de Bactérias, Instituto de Microbiologia Paulo de Goês, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

Introduction: *Staphylococcus aureus* is one of the most prevalent and relevant pathogens responsible for a wide spectrum of hospital-associated or community-acquired infections. In addition, methicillin-resistant *Staphylococcus aureus* may display multidrug resistance profiles that complicate treatment and increase the mortality rate. The ability to produce biofilm, particularly in device-associated infections, promotes chronic and potentially more severe infections originating from the primary site. Understanding the complex mechanisms involved in planktonic and biofilm growth is critical to identifying regulatory connections and ways to overcome the global health problem of multidrug-resistant bacteria.

Methods: In this work, we apply literature-based and comparative genomics approaches to reconstruct the gene regulatory network of the high biofilm-producing strain Bmb9393, belonging to one of the highly disseminating successful clones, the Brazilian epidemic clone. To the best of our knowledge, we describe for the first time the topological properties and network motifs for the *Staphylococcus aureus* pathogen. We performed this analysis using the ST239-SCCmecIII Bmb9393 strain. In addition, we analyzed transcriptomes available in the literature to construct a set of genes differentially expressed in the biofilm, covering different stages of the biofilms and genetic backgrounds of the strains.

Results and discussion: The Bmb9393 gene regulatory network comprises 1,803 regulatory interactions between 64 transcription factors and the non-redundant set of 1,151 target genes with the inclusion of 19 new regulons compared to the N315 transcriptional regulatory network published in 2011. In the Bmb9393 network, we found 54 feed-forward loop motifs, where the most prevalent were coherent type 2 and incoherent type 2. The non-redundant set of differentially expressed genes in the biofilm consisted of

1,794 genes with functional categories relevant for adaptation to the variable microenvironments established throughout the biofilm formation process. Finally, we mapped the set of genes with altered expression in the biofilm in the Bmb9393 gene regulatory network to depict how different growth modes can alter the regulatory systems. The data revealed 45 transcription factors and 876 shared target genes. Thus, the gene regulatory network model provided represents the most up-to-date model for *Staphylococcus aureus*, and the set of genes altered in the biofilm provides a global view of their influence on biofilm formation from distinct experimental perspectives and different strain backgrounds.

KEYWORDS

Staphylococcus aureus, gene regulatory network, biofilm, motifs, transcriptional regulatory network, computational model

1. Introduction

Methicillin-resistant *Staphylococcus aureus* is a Gram-positive bacterium that causes a broad range of hospital infections contributing to increased morbidity and mortality. The presence of an extensive and diverse virulence repertoire, in addition to notable antimicrobial multiresistance phenotypes, make *S. aureus* one of the main public health concerns worldwide. A very alarming statistic unveiled by Bardi et al. (2021) was the increase of primary and secondary infections related to medical devices such as catheters and the incidence of pneumonia in patients with severe COVID-19 admitted to intensive care units, where *S. aureus* was one of the main causative agents in both situations. Bacterial biofilm growth provides advantages compared to the planktonic lifestyle, and it is an important mechanism of chronic infections, often found in medical devices (Figueiredo et al., 2017). Biofilm formation involves four main stages: attachment, proliferation, maturation, and detachment (Schilcher and Horswill, 2020). Regarding the composition, biofilm varies according to the strain and can be classified into *ica*-dependent and independent. The biofilm formation *ica*-dependent is mediated by polysaccharide intercellular adhesin (PIA), and is more common in methicillin-susceptible *S. aureus* (MSSA). The biofilm formation *ica*-independent is more common in methicillin-resistant *S. aureus* (MRSA), and the multifactorial mechanisms involved are not fully understood (O'Neill et al., 2007; Schilcher and Horswill, 2020). In addition, the same strain can produce different types of biofilm matrix depending on environmental cues (Figueiredo et al., 2017; Schilcher and Horswill, 2020).

Responses to environmental changes are triggered by modifications in the gene expression patterns. The most common modification occurs at the transcriptional level controlled by the transcription factors (TFs). TFs recognize conserved nucleotide sequences in the promoter region of target genes (TGs) called transcription factor binding sites (TFBSs), thus controlling the gene expression according to the

environmental changes (Mercatelli et al., 2020). The set of genes or operons regulated directly by the same TF is defined as a regulon, and the set of regulons constitutes the gene regulatory network (GRN) of an organism (Rodionov, 2007). In Systems Biology, the set of regulatory interactions between TFs and their target genes can be modeled as a directed graph, where regulators and targets are nodes, and interactions are edges whose direction depicts the information flow (TF → TG) (Barabási and Oltvai, 2004). A GRN can be decomposed into subgraphs of recurrent regulatory patterns called network motifs that act to define the regulatory state of a target gene (Alon, 2007; Stone et al., 2019). It is possible to obtain relevant information from biological networks by exploring their topological metrics and properties from graph theory.

GRN inference provides a systematic comprehension of molecular mechanisms that regulate metabolic phenotypes of organisms under distinct endogenous or environmental conditions. The methods of GRN reconstruction are diverse and vary according to the input data and prediction goal. Faria et al. (2014) proposed the classification of the methods as genomic-driven and data-driven, the first is based on comparative genomics approaches, and the latter is based on reverse engineering from expression data (Faria et al., 2014). A recent review of different GRN inference methodologies can be found in the work of Mercatelli et al. (2020). The work published by Ravcheev et al. (2011) applied the genomic-driven method implemented in the RegPredict web server to infer the first GRN for *S. aureus* strain N315, an intermediate biofilm-producer belonging to the clonotype II-A (SCCmecII), and other six species of the *Staphylococcaceae* family (Ravcheev et al., 2011). However, N315 GRN has not been updated since 2011, and several regulons essential to biofilm formation, virulence, and antimicrobial resistance are absent.

In our work, we reconstructed a GRN for the MRSA strain Bmb9393 by applying comparative genomics and literature-based approaches. Bmb9393 is a high biofilm-producing Brazilian strain possessing a resistance profile to a broad range

of antimicrobials. The Bmb9393 genome is the first strain belonging to ST239-SCC*mecIII* lineage from Latin America to be available in databases. For the first time, we described the topological characteristics and identified network motifs for an *S. aureus* GRN. In addition, we profiled a genomic landscape of differentially expressed genes (DEGs) in the biofilm condition by analyzing several transcriptomes available in the literature. Finally, we mapped the DEGs into the Bmb9393 GRN to portray how growth in the biofilm can alter the regulatory systems. Thus, our work not only provides an updated GRN based on a strain with a genetic and phenotypic background considerably distinct from the N315 strain but also offers a global view of biofilm formation from the regulatory perspective of different isolates.

2. Materials and methods

2.1. Identification of the transcription factor repertoire of Bmb9393 strain

The nucleotide and amino acid sequence of the complete genome of Bmb9393 strain were obtained from the NCBI FTP server. The first step toward the identification of the transcription factor (TF) repertoire was the Bmb9393 genome re-annotation, which was based on similarity searches against the UniProtKB/Swiss-Prot database (UniProt Consortium, 2019) using the BLASTp algorithm with an $e\text{-value} \leq 10^{-5}$ and a coverage $\geq 80\%$. We compared the previous annotation with BLAST results and updated the information accordingly. In addition, to identify the DNA-binding domains, we locally run InterProScan (version 5.39-77.0) with default parameters to map the InterPro families and Pfam assignments of the entire proteome (Jones et al., 2014). Next, we retrieved TF information from databases such as the *Predicted Prokaryotic Transcription Factor* (P2TF), one of the most relevant, because it provided the TF classification into regulatory families (Ortet et al., 2012). Besides, we submitted the Bmb9393 genome to the *Predicted Prokaryotic Regulatory Proteins* (P2RP) tool to identify regulatory proteins (Barakat et al., 2013). The P2TF and P2RP results were compared to the previous annotation, and information was updated accordingly. Then, we searched the literature with specific terms such as “transcriptional regulator,” “response regulator,” “activator,” and “repressor” to identify articles associated with new TFs in *S. aureus*. Once a TF was found, we included it in the Bmb9393 annotation after ortholog detection using ProteinOrtho (Lechner et al., 2011). To compare strains of the same species, we used more restrictive parameters for the alignment: $e\text{-value} \leq 10^{-10}$, identity $\geq 35\%$, and minimum coverage of 80%. Finally, we defined a protein as TF after manually checking the results obtained with BLASTp against UniProtKB/Swiss-Prot, PFAM assignments of DNA-binding domains, and the genes identified by P2TF and P2RP.

2.2. Gene regulatory network reconstruction of Bmb9393 strain

The GRN reconstruction of Bmb9393 strain was performed based on a hybrid methodology comprising a reference-based inference and the prediction of *cis*-regulatory elements. The identification of regulatory interactions between TFs and their target genes was obtained by: (i) comparative propagation of the GRN of *S. aureus* strain N315 (Ravcheev et al., 2011); (ii) identification of TFBSs in the upstream region of Bmb9393 coding sequences using the DNA-binding sequences from orthologous TFs of *S. aureus*, and the model Gram-positive bacterium *Bacillus subtilis* 168; and (iii) literature-based inference of regulons. In step (i), we retrieved TFs and target genes and the available TFBSs from the RegPrecise database, including those on the GRN of N315 published in 2011 (Ravcheev et al., 2011; Novichkov et al., 2013). To propagate the network, we detected the set of orthologous TFs and target genes between the N315 and Bmb9393 using the ProteinOrtho (Lechner et al., 2011) with the same parameters described above, thus obtaining the first draft of the GRN of Bmb9393.

In step (ii), we applied the concept of operon-based expansion used in the reconstruction of the GRN of *Mycobacterium tuberculosis* (Sanz et al., 2011). Briefly, if gene A regulates gene B, and gene B is part of the BCD operon, then we included the interactions $A \rightarrow C$ and $A \rightarrow D$ in the network. For TFBSs from *S. aureus*, we built a list for each TF with all TFBSs obtained from RegPrecise (Novichkov et al., 2013), as well as the experimentally validated databases CollectTF (Kiliç et al., 2014) and PRODORIC (Dudek and Jahn, 2022). Each file was given as input to the genome-scale-dna-pattern tool of the Regulatory Sequence Analysis Tools (RSAT) software suite (Santana-Garcia et al., 2022). We executed the application locally to search for TFBSs in the upstream region of Bmb9393 coding sequences obtained through the retrieve-sequence tool of RSAT, covering -400 to $+50$ bases from the start codon with and without the option `noorf` (prevent overlap with neighbor genes). Then, we filtered the outputs and mapped the results to retrieve the gene identifier and gene product in the Bmb9393 genome, comparing each list of possible target genes with those associated with the same TF already included in the Bmb9393 GRN. If there was a new interaction or regulon, we searched the literature to validate the predicted gene or at least its function as part of that regulon. For TFBSs from *B. subtilis*, we first retrieved a list of TFs from its most recent GRN (Faria et al., 2016) to identify orthologs in the Bmb9393 using ProteinOrtho. To compare different species, we considered the parameters described in Janga and Moreno-Hagelsieb (2004): $e\text{-value} \leq 10^{-6}$ and minimum coverage of 60%. Next, we aligned the TFBSs from the *B. subtilis* orthologous TFs with those recovered from *S. aureus* to quantify the variability between the binding sites according to the classification described in Ravcheev et al. (2011). Only the TFBSs classified as conserved (up to four

TABLE 1 Selection of transcriptome studies of *S. aureus* strains in biofilm and planktonic conditions.

Techniques	Strains	DEGs	References
Microarray	UAMS-1	451	Beenken et al., 2004
Microarray	strain 113	157	Resch et al., 2005
Microarray	MRSA-M2	20	Brady et al., 2006
Microarray	USA300	471	Stewart et al., 2022
Bulk RNA-seq	N315, MRSA252, LAC, MW2 and NRS385	1310	Tomlinson et al., 2021
DEGs Total		2,409	
Non-redundant DEGs		1,794	

The bold value represents the number of non-redundant DEGs among all transcriptome studies.

different bases) were considered when performing the searches in the upstream region of the Bmb9393 coding sequences using the `genome-scale-dna-pattern` tool. Again, if there was a new interaction or regulon, a manual curation based on literature and databases was the primary method to validate the interactions in *S. aureus*.

In step (iii), we enriched the GRN of Bmb9393 by searching for TFBSs of the new TFs identified in the previous section to infer regulons not yet available in public databases. To this end, we sought in the literature articles from 2010 to 2018 associated with TFBSs for new regulators, since the only *S. aureus* GRN is from 2011, and applied step (ii) for incorporation into the GRN of Bmb9393. After the reconstruction steps, we applied a text mining pipeline to retrieve evidence of transcription factor regulation from the literature using the NCBI PubMed Central API (*Access Programming Interface*) (more details about the pipeline can be found in [Supplementary Section 1.1](#)). We used the results obtained from the pipeline to input the potential mode of regulation of a TF upon their targets.

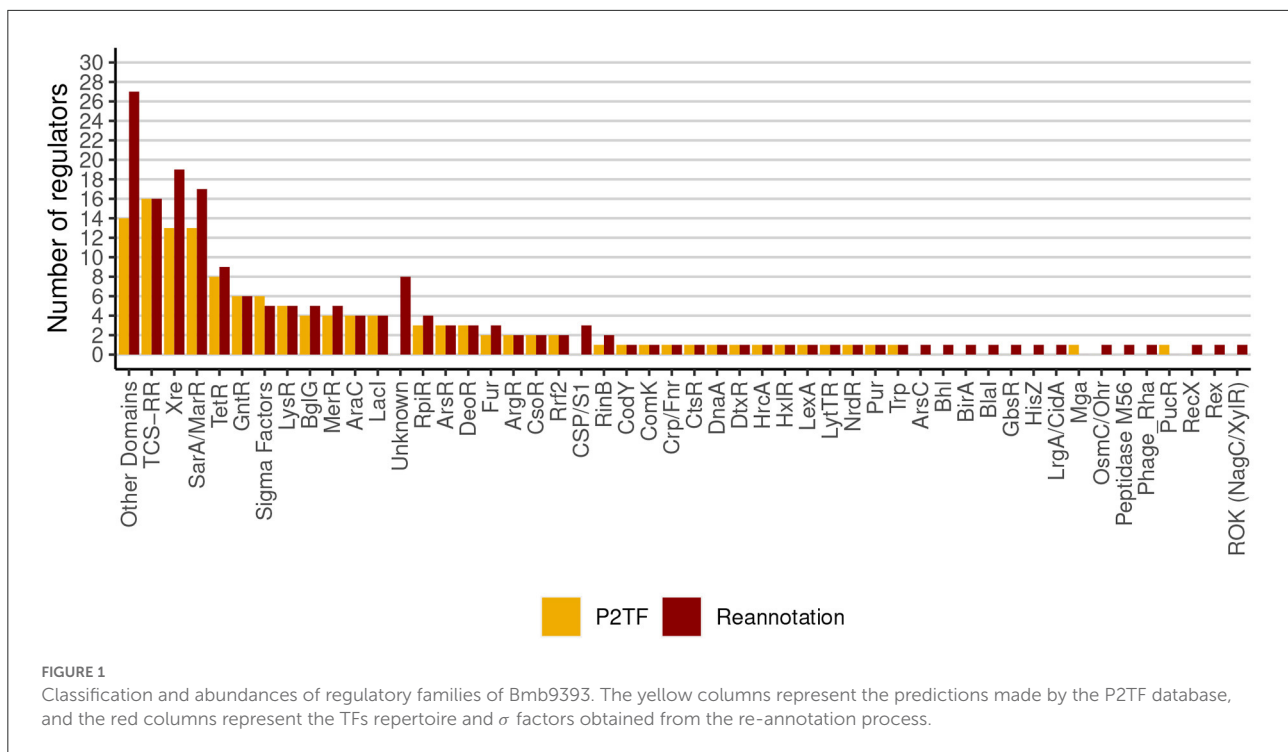
2.3. Topological analysis and comparison between the GRNs of Bmb9393 and N315 strains

We performed the network topological analysis using the `igraph` R package ([Csardi and Nepusz, 2006](#)). In addition, to detect *feed-forward loop* (FFL) motifs, we execute the `triad_census` function extracting the results from the corresponding pattern (more details about the FFL identification can be found in [Supplementary Section 1.1](#)) and classifying the types of FFL according to [Mangan and Alon \(2003\)](#). The statistical significance of the motif detection was calculated using the *Z-score* as described in

[Koutrouli et al. \(2020\)](#). The GRNs were manipulated with the `Cytoscape` software (version 3.9.1) ([Shannon et al., 2003](#)). We performed the Bmb9393 and N315 network overlapping using `DyNet App` ([Goenawan et al., 2016](#)) for `Cytoscape`.

2.4. Identification of differentially expressed genes in biofilm and overlapping with the GRN of Bmb9393

We obtained the differentially expressed genes (DEGs) from transcriptome analysis of different strains of *S. aureus* cultured under biofilm and planktonic conditions deposited in the former SAMMD database (current SATMD - *Staphylococcus aureus Transcriptome Meta-Database*) ([Nagarajan and Elasri, 2007](#)). In addition, we searched the literature on recent *S. aureus* transcriptome studies related to biofilm formation, resulting in the selection of four microarrays and one bulk RNA-seq data ([Table 1](#)). We considered for the analysis all DEGs regardless of their direction. This is due to several limiting factors as transcriptomes are from different *S. aureus* strains, collected in distinct timepoints and sequenced with different technologies. We determined one set of non-redundant DEGs from microarrays and another from the bulk RNA-seq (DEGs shared by two strains), and identified the orthologs in the Bmb9393 genome using `ProteinOrtho` ([Lechner et al., 2011](#)). The orthologous DEGs were included in the orthologous set of non-redundant DEGs from microarrays and from RNAseq lists. Next, we retrieved the amino acid sequences of the non-redundant genes from both lists using the `SeqKit` ([Shen et al., 2016](#)) and submitted them to `EggNOG-mapper` (available online at <http://eggno-mapper.embl.de/>) ([Huerta-Cepas et al., 2017](#)) to obtain a classification of the DEGs into functional categories of orthologs clusters (COGs) and ontology terms (GO) ([Tatusov et al., 2000](#); [The Gene Ontology Consortium, 2019](#)). We used default parameters except for *query* and *subject* minimum coverage of 60%. We used the `REVIGO` web server (available at <http://revigo.irb.hr/>) with default parameters to summarize the GO terms obtained from `EggNOG-mapper` ([Supek et al., 2011](#)). In addition, we performed a functional classification of the proteins into KEGG Orthology (KO) groups using `BlastKOALA` (available at <https://www.kegg.jp/blastkoala/>) to reconstruct the KEGG pathways using the `Reconstruct` tool from KEGG Mapper (available at <https://www.genome.jp/kegg/mapper/reconstruct.html>) ([Kanehisa et al., 2016, 2022](#)). Last, we mapped the orthologous set of non-redundant DEGs from microarrays and from RNAseq lists in the GRN of Bmb9393 to recover the regulons containing some or all differentially expressed target genes in biofilm.



3. Results and discussion

3.1. Transcription factor repertoire of the Bmb9393 strain

The *S. aureus* strain Bmb9393 is a high biofilm-producing bacterium isolated from nosocomial bacteremia. It is the first strain belonging to the Brazilian epidemic clone (BEC) from the ST239-SCC*mecIII* lineage to have its genome fully sequenced (GenBank accession number CP005288 for the chromosome, and CP005289 for the plasmid) (Costa et al., 2013). The re-annotation process of the Bmb9393 genome was important to include new information obtained since the first annotation in 2012. After the re-annotation, we identified a set of non-redundant transcriptional regulators consisting of 175 TFs and five sigma (σ) factors (Supplementary Table S1), classified into 48 regulatory families according to PFAM domains, the P2TF database and Ibarra et al. (2013) (Figure 1). The most represented families in Bmb9393 were Xre with 19 members, and MarR with 17 members, including the Sar proteins, followed by the TCS regulators with 16 members, and TetR family with 9 members. Twenty-seven regulators had domains characterized by PFAM (Finn et al., 2014) but no defined family (*Other Domains*), and 8 regulators had neither defined family nor domain according to PFAM assignments (*Unknown*) (Figure 1). Bmb9393 presented a larger transcription factor repertoire compared with the strains analyzed by Ibarra et al. (2013), whose numbers ranged from 126 to the RF122 bovine strain

up to 151 to the N315 strain (more results about the re-annotation process can be found in Supplementary Section 1.2 and Supplementary Figure S1).

Inside the Xre family, *xdrA* (SABB_01966) and *potR* (SABB_01062) are the only genes that have described functions. XdrA is an important transcriptional activator of the staphylococcal protein A (SpA), an important virulence factor (Cruz et al., 2021), whereas PotR is an activator of the *potABCD* operon, involved in the polyamine uptake in *S. aureus* (McCallum et al., 2010; Yao and Lu, 2014). Both genes play important roles in pathogenicity and virulence mechanisms, such as biofilm formation, quorum sensing, repression of capsule production, and homeostasis (DeFrancesco et al., 2017; Lei and Lee, 2018; Gimza et al., 2019). Unlike other genes encoding proteins of the Xre family in the Bmb9393 strain, *xdrA* and *potR* are not located inside or near phage-related elements. Of the 21 regulators, 14 are related to regions of genomic plasticity (RGPs), and 5 are conserved among all *S. aureus* strains analyzed by Ibarra et al. (2013): SABB_00224, SABB_02114, SABB_02258 (target gene of CodY), *xdrA* and *potR* (target gene of SrrA and ComK1).

The “Sigma Factors” is the sixth family most represented in the Bmb9393 genome (Figure 1) and includes the four known σ factors of *S. aureus*. While σ^A (*sigA* or SABB_00481) controls the expression of most genes, including housekeeping genes, the alternative σ factors regulate subsets of genes involved in specialized functions such as response and adaptation to environmental changes (Paget and Helmann, 2003). In this

context, σ^B (*sigB* or SABB_02394) plays an important role in the complex transcriptional network of *S. aureus* because it regulates the expression of virulence determinants such as biofilm, and increases the resistance to antimicrobials that target the cell wall (Bischoff et al., 2004; Pané-Farré et al., 2006). In addition, σ^B is related to the expression of genes involved with transport and stress response (Ishii et al., 2014; Guldimann et al., 2016). Another alternative σ factor, σ^H (*sigH* or SABB_02619), regulates genes related to natural competence for DNA transformation (Morikawa et al., 2012). Maree et al. (2022) demonstrated the occurrence of transformation in biofilm and showed that advantageous conditions for biofilm formation increase the transformation efficiency, identifying SrrA, SigH, and ComK as essential competency regulators (Maree et al., 2022). SigH also regulates the integration and excision of prophages and stability of the lysogenic state, playing an important role in horizontal gene transfer (HGT) (Tao et al., 2010). Last, the extracytoplasmic function (ECF) σ^S factor (*sigS* or SABB_01901) is involved in DNA repair response, cell wall stress, and amino acid starvation (Shaw et al., 2008; Miller et al., 2012). We did not include the σ factors in the Bmb9393 GRN but we reconstructed the σ^B and σ^H regulons due to their importance in the biofilm formation (Supplementary Table S2).

3.2. Gene regulatory network of the Bmb9393 strain

The initial step of the reconstruction consisted of the comparative propagation of the GRN of *S. aureus* strain N315 (Ravcheev et al., 2011). The N315 GRN comprises 49 regulons listed in the RegPrecise database (<http://regprecise.lbl.gov/>). Except for PdxR, YdfD/YisV, and RbsR2, the remaining 46 TFs were orthologous to Bmb9393 genes. For each orthologous TF, we propagated all target genes that were orthologous to Bmb9393 genes. It is important to mention that we choose not to include in the GRN truncated TFs annotated as pseudogenes. It was the case of the repressor MecI which has a truncated gene (*mecI**) in the genome of Bmb9393, and appears to be conserved and non-functional in the ST239-SCC*mecIII* lineage (Harris et al., 2010; Oliveira and de Lencastre, 2011).

After the propagation, we performed an *in silico* validation, and the expansion of the draft GRN based on searching for TFBSs from orthologous TFs in the upstream region of Bmb9393 coding sequences. Using the list of TFBSs recovered from *S. aureus* TF orthologs, we validated the presence of propagated target genes. In addition, we included three new regulons, CsoR, SarA, and MgrA identified by our approach. From the 72 *B. subtilis* orthologous TFs to Bmb9393, we retrieved information about 40 genes, of which 35 were already in the Bmb9393 GRN. We analyzed the variability level between the TFBSs of both species according to the classification applied

by Ravcheev et al. (2011), and identified 13 TFBSs classified as substantially different from the perspective motifs of *B. subtilis* (CggR, MntR, PurR, CsoR, CzrA, BraR, CstR, GraR, YtrA, HisR, GlvR, and NreC e RbsR), 7 as moderately different (ArgR, BglR/YydK, FapR, FruR, MtlR, and TreR e Zur), and 15 TFBSs were well conserved or only slightly variable (BirA, CcpA, CodY, CtsR, CymR, Fur, GltC, GlnR, GntR, HrcA, LexA, MurR, NrdR, and PerR e Rex) (Ravcheev et al., 2011). After the exclusion of the 13 TFBSs substantially different, the searching in Bmb9393 resulted in the addition of 12 new targets for CcpA (3 of them were experimentally validated, *manPA*, *drp35*, and *metK*), 2 for CodY (*nrdIEF*, and *vfrBA*), and one for Fur (SABB_01577). The moderately different TFBSs (2 to 4 non-conserved bases) did not return any potential target gene in Bmb9393 corroborating the results obtained by Ravcheev et al. (2011) that TFBSs are species-specific (Ravcheev et al., 2011). The remaining five orthologous TFs (CcpE, ComK, GbsR, MhqR, and SrrA, apart from SigB-Supplementary Table S2) were included as new regulons after the search in the intergenic regions of Bmb9393 for TFBSs recovered from *B. subtilis* to identify target genes, followed by manual curation of the results. The last step involved the GRN enrichment, based on searching in the literature for new TFs that were not submitted to any specific regulatory database. We retrieved a list of 51 TFs of which 23 were excluded because they are already in the draft GRN, do not have orthologs in Bmb9393, or were not characterized. Of the remaining 28, we identified TFBSs for 11 TFs (BraR, CstR, GraR, HypR, Rsp, SarR, WalR, SarT, QsrR, ArlR, and GbaA) and σ^H , and used those TFBSs to search in the upstream region of Bmb9393 coding sequences to identify new potential target genes.

The final GRN comprises 1,803 interactions between 64 TFs and 1,151 non-redundant target genes (Supplementary Figure S2 and Supplementary Table S3 available on https://github.com/maiolivei/GRN_Bmb9393/blob/main/Supplementary_Table_S3.xlsx). Figure 2 shows the main connected component of the Bmb9393 GRN, where we highlighted the regulons: SarA with 241 target genes, MgrA with 206, and CodY with 200 followed by CcpA with 195 target genes. SarA (SABB_00667) is an important TF that acts together with Agr as the known *sarA|agr* regulatory system, the principal regulator of the expression of genes involved in virulence. The system is related to the transition from exponential to post-exponential growth phase when the repression of adhesin biosynthesis and activation of exoprotein production occurs (Blevins et al., 2002; Jenul and Horswill, 2019; Oriol et al., 2021). SarA has a positive and/or negative effect on 241 target genes, including genes related to biofilm formation and several virulence factors such as adhesin, capsule clusters, leucotoxins, and proteases.

The second largest regulon corresponds to the pleiotropic dual transcriptional regulator MgrA (SABB_00733), which triggers several defensive pathways in response to oxidative stress in a SigB-dependent manner (Deochand and Grove, 2017).

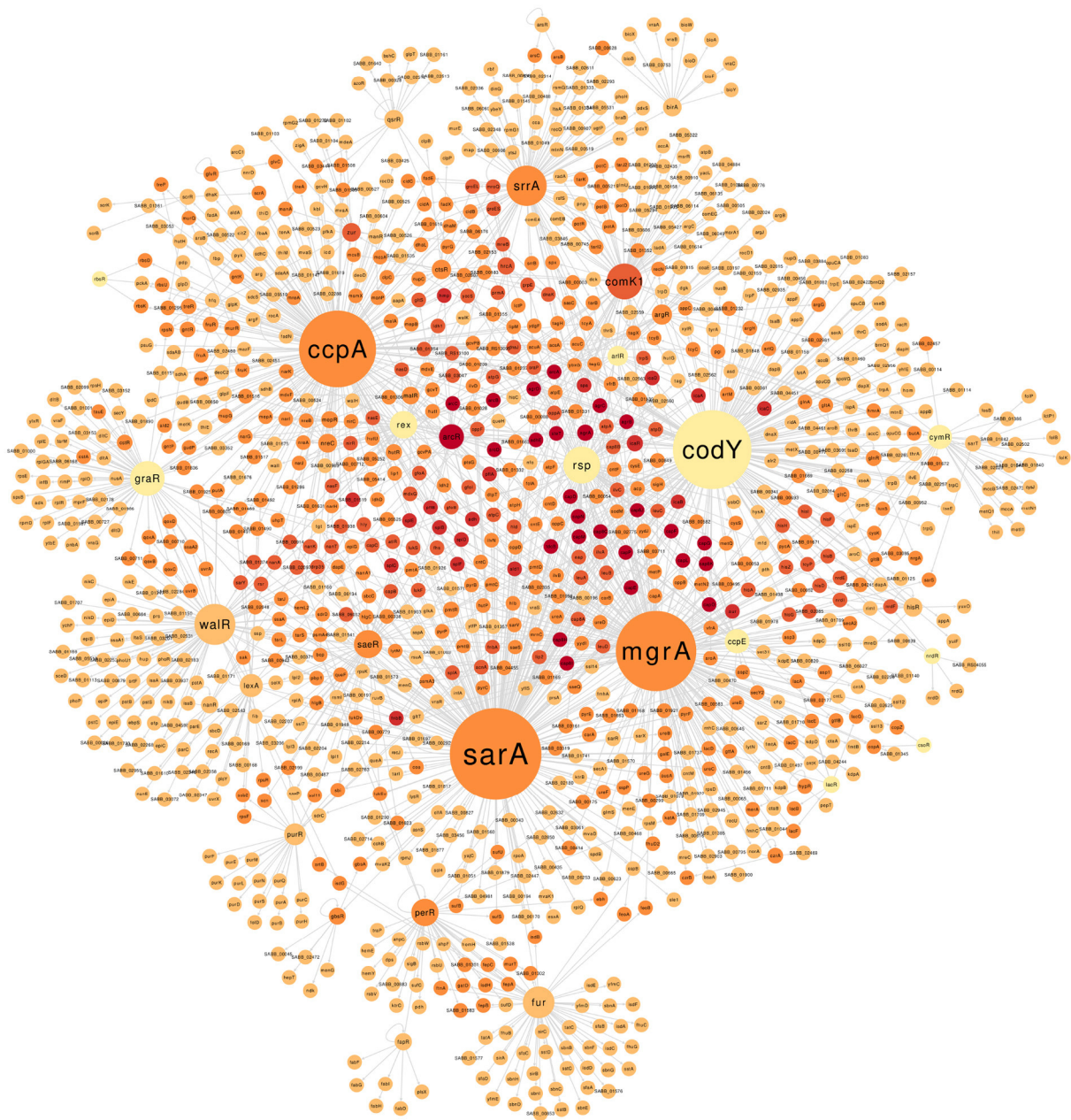
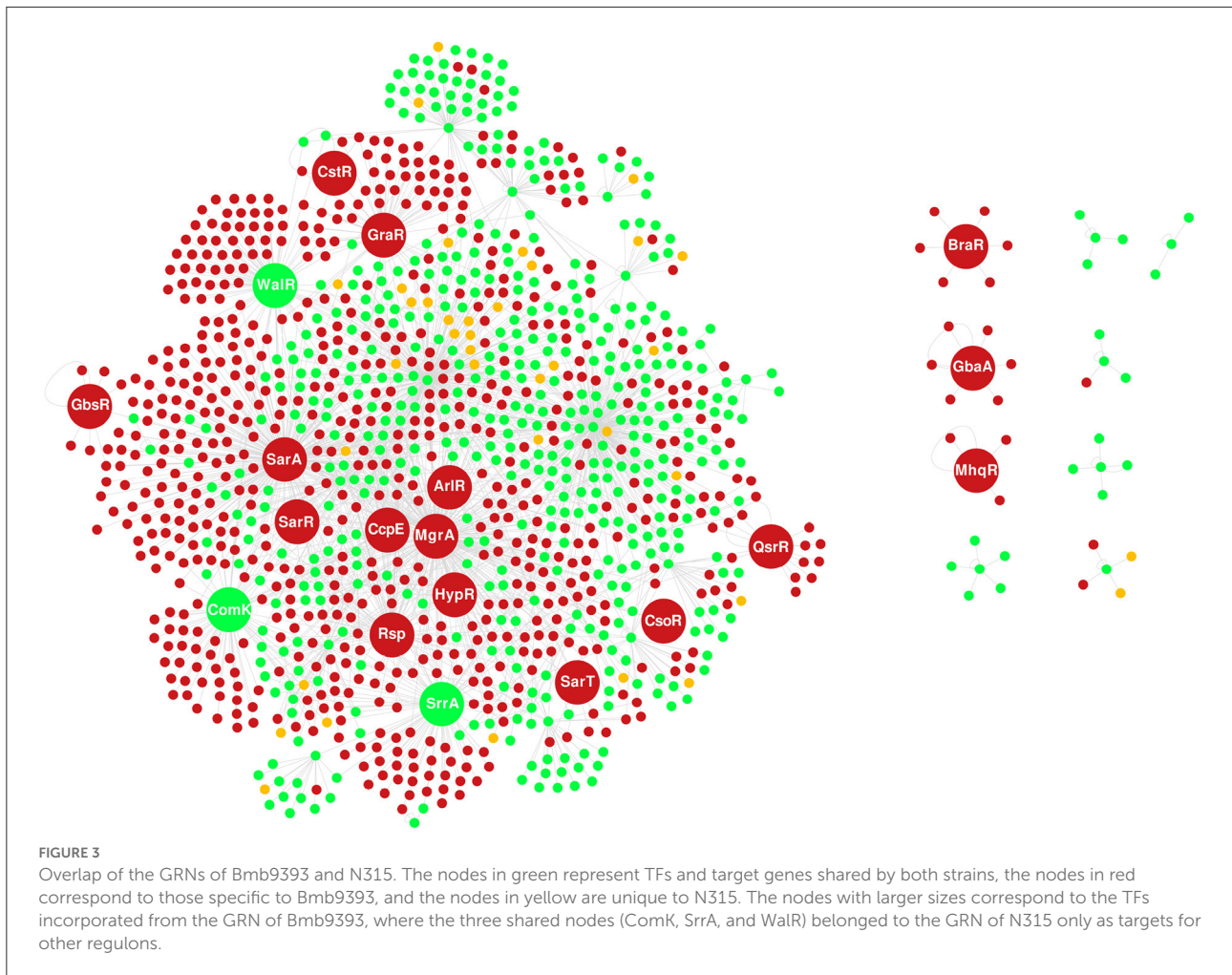


FIGURE 2
The largest connected component of the GRN Bmb9393. The size of the nodes is proportional to the outdegree values (number of edges leaving the node), with darker node colors indicating higher indegree (number of edges arriving at the node). The GRN of the Bmb9393 with all connected components can be found in [Supplementary Figure S2](#).

This TF modulates genes involved in several biological processes such as autolysis, biofilm formation, virulence, antibiotic resistance, staphylococcal capsule biosynthesis, and metabolism (Luong et al., 2006; Lei and Lee, 2020). MgrA has a similar expression pattern to *agr* and acts as a negative regulator of the biofilm as it inhibits the expression of adhesins and the autolysis process while activating extracellular enzymes that degrade adhesive matrix components (Luong et al., 2006; Crosby

et al., 2016; Jiang et al., 2018). The 206 target genes of the MgrA regulon in Bmb9393 comprise relevant protein classes with experimental evidence such as exotoxins, adhesins, and transporters.

The global regulator CodY (SABB_00251) controls the expression of several genes related to metabolism and virulence in response to the concentration of isoleucine, leucine, valine, and guanosine triphosphate (Waters et al., 2016). Thus, it works



as a link between the nutritional condition and virulence acting mostly as a repressor of its target genes (Pohl et al., 2009). Indeed, among the known modes of regulation present in the Bmb9393 GRN for CodY, 150 out of 200 are negative. In addition, the majority of target genes act on the metabolism, transport of amino acid and carbon, and virulence both indirectly by the repression of *agrB* and *saeR*, and directly through the binding on the promoter region of capsule-related genes, biofilm-related *ica* operon, and protease-encoding genes such as *aur*.

The fourth largest regulon belongs to CcpA (SABB_01861), the main regulator of the carbon catabolite repression phenomenon, where the expression of genes related to the consumption of different carbon sources is repressed by the presence of glucose or another preferred carbon source (Görke and Stülke, 2008). CcpA regulates several central metabolism pathways, amino acid biosynthesis, and virulence in a glucose-dependent and independent manner (Seidl et al., 2006, 2009; Troitzsch et al., 2021). In the Bmb9393 GRN, CcpA acts on the activation and repression of 195 target genes, highlighting the operons for consumption of alternative carbon sources, and

operons important for biofilm formation such as *cidABC* which is responsible for the cell lysis and eDNA release, an essential structural component of biofilm (Rice et al., 2007).

3.3. Comparison between the Bmb9393 and N315 GRNs

To the best of our knowledge, the reconstruction performed by Ravcheev et al. (2011) is the only *S. aureus* GRN published up to date. The N315 GRN comprises 49 TFs and the non-redundant set of 548 target genes. The overlap of N315 and Bmb9393 GRNs showed that the major variability resides in the target genes of shared global regulators, mainly CodY, CcpA, and Fur. The global TFs recognize a larger number of less conserved TFBSs, showing more flexibility than local TFs that bind with high affinity to one or few TFBSs (Balleza et al., 2009). It is important to mention that we choose to map the orthologous genes between Bmb9393 and N315 to avoid annotation bias. Of the 665 interactions present in the N315 GRN, we found 620 with orthologous genes to Bmb9393. Thus, the N315 GRN

contains 45 non-shared interactions with Bmb9393, apart from other differences shown in Figure 3. In addition, we found in Figure 3 the 19 new regulons included in the Bmb9393 GRN, of which 11 were obtained from the literature (ArlR, BraR, HypR, Rsp, SarT, SarR, CstR, WalR, GraR, QsrR, and GbaA), whereas eight were obtained from the expansion methodology (CcpE, ComK1, CsoR, GbsR, MgrA, MhqR, SarA, and SrrA).

SrrA (SABB_03639) is the response regulator of the SrrAB two-component system (TCS) orthologous to ResDE of *B. subtilis*, essential to the growth in anaerobic conditions controlling the respiratory shift from aerobiosis to anaerobiosis (Nakano et al., 1996). Studies with SrrAB (SrhSR) in *S. aureus* showed its relationship with the anaerobic respiratory shift is given by the activation of genes encoding fermentative enzymes (*ldh* and *adh*) and repression of citric acid cycle genes (*acnA* and *fumC*) (Throup et al., 2001). SrrA also controls the regulation of virulence factors in an *agr*-dependent and independent manner, besides the stimulation of biofilm formation (Pragman et al., 2004; Mashruwala et al., 2017). Among the 80 target genes predicted for SrrA, we highlight the *srrA* itself, *agrB*, *cidA*, *spa*, *adhE*, and *hmp*, all confirmed by the literature. Although not experimentally validated, we also mention the SasC adhesin and the regulator Rbf, both related to biofilm formation (Rice et al., 2007; Cue et al., 2009).

Rps (SABB_01294) is another newly included regulator that plays a role in the repression of biofilm formation by the inhibition of surface proteins while activating secreted proteins such as proteases (Lei et al., 2011). In addition, Rsp is relevant to the gene expression adaptation to acute infection development with severe tissue damage by promoting toxin production and limiting biofilm formation, being the latter a typical mechanism of chronic infections (Li et al., 2015). The Rps target genes were all experimentally validated with emphasis on those related to virulence, for instance, *agrB*, *hld*, *icaA*, *fnbA*, *fnbB*, and the cluster of capsule production.

Last, we highlight WalR (SABB_03186), the response regulator of WalKR TCS (YycG/YycF) that activates autolysin biosynthesis and biofilm formation. The WalKR system is a potential therapeutic due to its essentiality in the Firmicutes phylum, being specific to Gram-positive bacteria with low GC content. In *S. aureus*, apart from the effect on cell wall metabolism, WalKR is important to virulence and antimicrobial resistance, especially vancomycin (Delauné et al., 2012; Peng et al., 2017). Despite its importance, the nature of WalKR essentiality remains unclear. From the 82 target genes identified, we highlight those confirmed by the literature related to cell wall degradation and autolysis (*lytM*, *isaA*, *sceD*, and *ssaA*), and virulence (*sdrD*, *ebpS*, and *sak*). The target gene *ltaS* of the WalR regulon in Bmb9393, classified as a hypothetical protein in previous studies, drew our attention as a possible reason for the reported essentiality of the system. It is responsible for synthesizing lipoteichoic acid

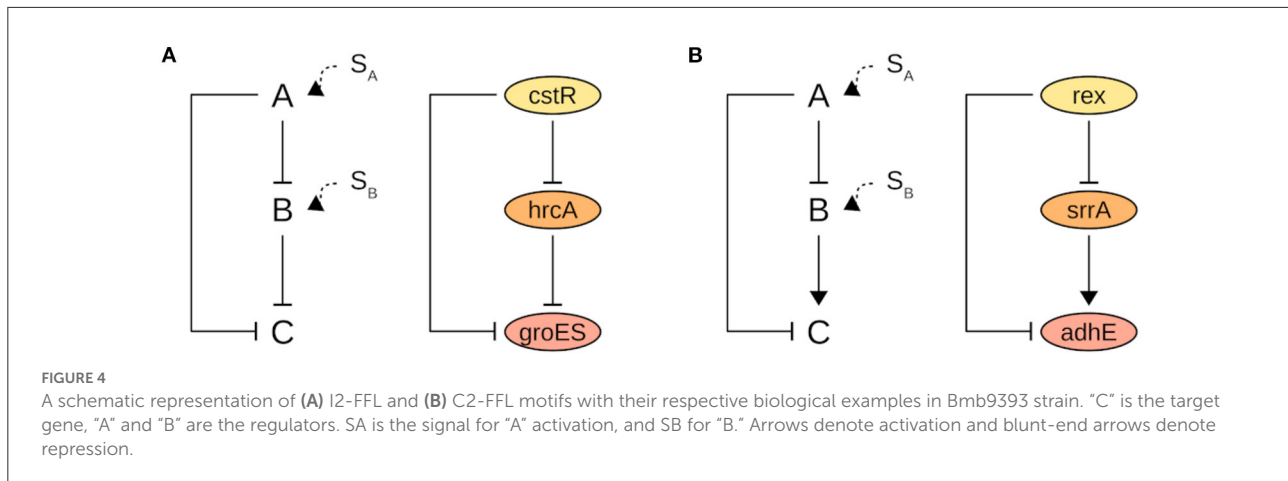
TABLE 2 Topological metrics between the GRNs of Bmb9393 and N315 strains.

Network Metrics	N315 GRN	Bmb9393 GRN
Nodes	563	1,162
Edges	665	1,803
Average network degree	2.36	3.10
Network density	$2.09e^{-3}$	$1.33e^{-3}$
Network diameter	3	4
Network radius	1	1
Average clustering coefficient	0.056	0.068
Connected components	21	10
Average shortest path	1.21	1.69
Average number of neighbors	2.23	3.02
FFL motifs (Z-score)	7 (4.03)	54 (2.13)

and required for bacterial growth and cell division, with experimental evidence of transposon library-based methods confirming *ltaS* as essential in *S. aureus* (Gründling and Schneewind, 2007; Chaudhuri et al., 2009; Santiago et al., 2015).

3.4. Comparative topological analysis between the Bmb9393 and N315 GRNs

The topological analysis of both networks reinforced the enrichment of Bmb9393 GRN (Table 2). The average network degree is the sum of all interactions (in and out-degrees) divided by the total number of nodes. The increased mean seen in the Bmb9393 GRN reflects the larger number of interactions present in the network. The network density is the ratio between the number of interactions in a network and the total number of possible interactions. Both Bmb9393 and N315 GRNs showed a low-density measure and are considered sparse networks. This is expected due to the biological nature of gene expression regulation where a small percentage of a genome is composed of regulatory genes and not all regulators control all genes (Cai et al., 2013; Koutrouli et al., 2020). However, the density of Bmb9393 GRN showed to be slightly smaller than N315 as a consequence of the higher number of nodes. Network diameter is the maximum number of edges to go through, considering the shortest possible path between two nodes. Unlike, the radius is the minimum distance between two nodes. In turn, the average shortest path is the mean of the shortest paths of all node pairs in the network. Except for radius, all those metrics are slightly higher in the Bmb9393 network once again reflecting the increased number of nodes and edges.



The average number of neighbors is another topological measure that reflects the enrichment of the network connectivity. A network can be expanded in the number of nodes without necessarily having an increase in the mean number of neighbors. The improved connectivity of Bmb9393 GRN can be seen by the higher expected number of neighbors for a node. Another measure showing this aspect more directly is the number of connected components that decreased considerably in the Bmb9393 GRN. A connected component is a set of nodes connected to each other by a path. The Bmb9393 GRN has only 10 connected components being the larger composed of 1,097 nodes, while the N315 has 21, the larger containing 449 nodes. A cluster is a set of nodes that are more connected with each other than they are with the rest of the network. The clustering coefficient of a node shows the tendency a node has to form a strongly connected group. Although only slightly, the Bmb9393 GRN showed an average clustering coefficient higher than N315 reinforcing the presence of larger connected communities.

Last, we searched the networks for motifs that can be interpreted as repeated patterns of interaction from which networks are composed. The most recurring pattern found in biological networks is the feed-forward loop (FFL). A FFL is composed of three nodes, which, in the context of gene regulatory networks, two of them are TFs and one target gene. The TF "A" regulates the TF "B", and the target gene "C" is regulated by both TFs. The biological importance of this type of motif is its contribution to fine-tuning gene expression regulation (Mangan and Alon, 2003; Alon, 2007). We found 7 FFL motifs in the N315 GRN in comparison to 54 in the Bmb9393 GRN (Supplementary Figures S3, S4). The FFLs motifs are classified into eight types depending on the mode of regulation that the TFs exert on target gene but because several interactions in both GRNs have an unknown or dual mode of regulation, we could not precisely

estimate which type is prevalent. Among those we could identify, coherent type 2 and incoherent type 2 were the most frequent.

Overall, the presence of a signal such as an inducer molecule, amino acid modification, or protein interaction activates the TFs which then exert their control upon the target gene. The control effect can be a delayed response in the case of a coherent FFL or a hastened response in the case of an incoherent FFL. The incoherent type 2 FFL (I2-FFL) is formed by two repressors "A" and "B", and the expression of target "C" requires the absence of the signal for both TFs simultaneously. If only the signal for "A" is present, "C" is repressed but if the signal fades "C" can be readily expressed. If the signal for "B" is present, "C" will be repressed independent of the presence or not of a signal for "A" (Figure 4) (Mangan and Alon, 2003; Alon, 2007). A practical example is the I2-FFL formed by the regulators CstR and HrcA, and the target *groES*. GroES is a chaperone protein that is expressed in response to heat shock stress. Indeed, Chastanet et al. (2003) demonstrated that both regulators act together to maintain the expression level of *groES* under control in the absence of stress (Figure 4A). The target genes composing the other I2-FFLs found are also related to some type of stress response.

The coherent type 2 FFL (C2-FFL) is formed by one repressor "A" and one activator "B". If "A" and "B" affect "C" expression independently, in the presence of a signal for "A" the repression of C will be delayed. Otherwise, if "A" and "B" affect "C" expression jointly, in the absence of a signal for "A" the activation of "C" will be delayed. A biological example of C2-FFL is the triad Rex, SrrA, and *adhE* (Figure 4B). AdhE is a fermentative enzyme induced under anaerobic conditions. Rex is inactivated when the NAD⁺/NADH ratio decreases leading to de-repression of *adhE*. However, (Pagels et al., 2010) demonstrated that expression of *adhE* is only strongly upregulated in the Rex mutant in anaerobiosis, suggesting that

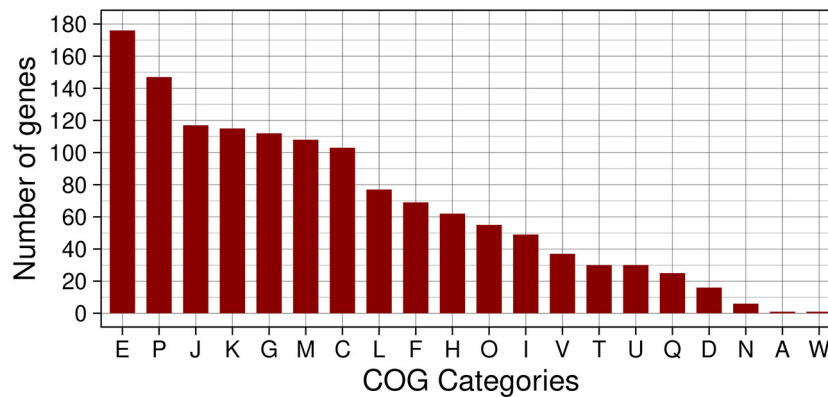


FIGURE 5

Histogram of COG functional classes of genes with altered expression in the biofilm in *S. aureus*. The COGs represented in the figure are: E → amino acid transport and metabolism, P → inorganic ion transport and metabolism, J → translation, ribosomal structure and biogenesis, K → transcription, G → carbohydrate transport and metabolism, M → cell wall/membrane/envelope biogenesis, C → energy production and conversion, L → replication, recombination and repair, F → nucleotide transport and metabolism, H → coenzyme transport and metabolism, O → posttranslational modification, protein turnover, chaperones, I → lipid transport and metabolism, V → defense mechanisms, T → signal transduction mechanisms, U → intracellular trafficking, secretion, and vesicular transport, Q → secondary metabolites biosynthesis, transport and catabolism, D → cell cycle control, cell division, chromosome partitioning, N → cell motility, A → RNA processing and modification, and W → extracellular structures.

an additional mechanism is required to fully induce *adhE* expression. Indeed, the activation of *adhE* is also dependent on SrrA which is also induced under oxygen limitation (Kinkel et al., 2013).

Incoherent FFLs are useful when a speedy response is required, and coherent FFLs when a specific response should only occur in the dependency of a persistent signal. Based on this assumption, motif detection can be useful to draw the attention of researchers to target genes composing the FFL triads that are annotated as hypothetical proteins to investigate their functions.

3.5. Assessment of genes associated with the biofilm in diverse backgrounds

Biofilm formation is a complex process with intricate regulation, and its composition is highly influenced by the genetic background of strains and environmental signals. To gain insights into the biofilm formation process from a regulatory point of view obtained from diverse backgrounds, we obtained two non-redundant sets of genes with altered expression in the biofilm: orthologous set of non-redundant DEGs from microarray experiments with 1,061 members, and orthologous set of non-redundant DEGs with 1,310 genes obtained from bulk RNA-seq. The overlap of the two sets revealed 576 shared genes in addition to 485 unique genes from the microarray data and 734 specific to the bulk RNA-seq. Then, we merged both shared and unique genes into a non-redundant set with 1,794 proteins referred to as DEGs in biofilm (Supplementary Table S4 available on https://github.com/maiolivei/GRN_Bmb9393/blob/main/Supplementary_Table_S4.xlsx).

[com/maiolivei/GRN_Bmb9393/blob/main/Supplementary_Table_S4.xlsx](https://github.com/maiolivei/GRN_Bmb9393/blob/main/Supplementary_Table_S4.xlsx)).

The functional characterization of the 1,794 proteins into COG categories, after removal of the S category assigned as unknown functions, resulted in the distribution shown in Figure 5. The most representative categories in the set of genes with altered expression in the biofilm were: E for amino acid transport and metabolism, P for inorganic ion transport and metabolism, J for translation, ribosomal structure and biogenesis, K for transcription, and G for carbohydrate transport and metabolism. Indeed, categories E and G encompass the function of target genes regulated by CodY and CcpA that respond to environmental condition variations by modulating the expression of metabolic and virulence genes. Furthermore, the over-representation of the categories K with a large diversity of transcriptional regulators, J with all the apparatus for protein translation, and P with transport systems for several essential metabolites is consistent with the nutritional gradient in the deeper biofilm layers. The chemical heterogeneity resulting from the bacterial metabolic activity and the formation of channels in the biofilm layers produces the oxygen, nutrients, and bacterial waste product gradients that require intense regulation to adapt to locally variable conditions throughout the biofilm formation process (Stewart and Franklin, 2008). Analysis of the ontology for the set of 1,794 proteins with altered expression in biofilm resulted in 2,504 GO terms with 24,718 occurrences. The representative set resulting from REVIGO contains biological processes associated with regulation, amino acid and carbohydrate metabolism, and diverse types of stress responses. The results corroborate those obtained by the COG analysis and highlight the importance of the stress response

produced by the microenvironments in the biofilm when the preferential nutrient sources are depleted in the deeper layers and bacteria metabolic shift to alternative sources. One source of accurate information from the literature resides in the KEGG metabolic pathways, so we mapped the set of genes with altered expression in the biofilm to obtain the most prevalent KEGG metabolic pathways and functional hierarchies. This analysis corroborates the previous results where the largest modules of complete pathways were amino acids metabolism followed by carbohydrate metabolism, and cofactors and vitamins. As expected, among the functional hierarchies of BRITE we highlight those related to TFs and TCSs, transport and secretion systems, and antimicrobial resistance genes, in addition to the one related to enzymes with the highest number of proteins.

After the functional annotation of DEGs in biofilm, we identified the presence of classes of proteins considered essential for biofilm formation in the literature, especially for *ica*-independent biofilm, as it is multifactorial and poorly understood. The first class comprises the surface proteins involved in the bacteria-host matrix interactions during colonization such as adhesion, invasion, and immune system evasion processes (Foster et al., 2014). Among these, fibronectin (Fn)-binding proteins A and B (FnBPA and FnBPB) are fundamental factors for *ica*-independent biofilm by enhancing bacterial aggregation and promoting biofilm maturation

(O'Neill et al., 2007; Shanks et al., 2008). The FnBPA is present in DEGs in biofilm along with other surface proteins relevant for different phases of biofilm formation, specially in the early to mature stages such as the multifunctional adhesin SpA, SasG, members of the Sdr family (SdrC, SdrD, and SdrE), fibrinogen-binding protein ClfA, proteins of the Isd system (IsdA, IsdB, and IsdC), secreted adhesins Emp and Eap, the elastin-binding protein EbpS and the non-protein adhesin PIA, the fundamental *ica*-dependent biofilm mediator. The absence of some surface-bound proteins in DEGs in biofilm may be explained by the different conditions of the transcriptomes or by the variation in *S. aureus* genomes since FnBPB and ClfB belong to a group that was absent in several strains (McCarthy and Lindsay, 2010). The second class is composed of secreted proteins that modulate biofilm integrity, where proteases are involved in dispersal from the mature biofilm and degradation of the *ica*-independent biofilm matrix (Mootz et al., 2013; Loughran et al., 2014). Indeed, the set of DEGs in biofilm contains all the protease operons and their respective inhibitors such as the serine proteases SspA and the SplA-F operon, the cysteine proteases ScpA and SspB, the metalloprotease Aur, and the intracellular inhibitors of the cysteine proteases ScpB and SspC (Mootz et al., 2013). Another secreted protein with altered expression in the biofilm was α -hemolysin (*hly/hla*) which contributes to the initial attachment steps and intercellular

TABLE 3 TFs with at least one of their target genes altered in the biofilm.

TF name	TF identifier	Functional role of the regulon	Target genes altered in the biofilm	Total target genes of regulon
<i>argR</i>	SABB_02844	Arginine metabolism	14	17
<i>arsR</i>	SABB_06187	Arsenic resistance	1	3
<i>codY</i>	SABB_00251	Amino acid metabolism	168	200
<i>fapR</i>	SABB_00278	Fatty acid biosynthesis	3	8
<i>gapR</i>	SABB_00821	Glycolysis	1	6
<i>graR</i>	SABB_00706	Virulence/Antimicrobial Resistance	46	62
<i>lacR</i>	SABB_01464	Lactose utilization	6	7
<i>malR</i>	SABB_00432	Maltose-maltodextrin utilization	5	9
<i>mepR</i>	SABB_01515	Multidrug resistance	6	7
<i>mgrA</i>	SABB_00733	Virulence	168	206
<i>mhqR</i>	SABB_03589	Response to oxidative stress	3	4
<i>mntR</i>	SABB_00682	Manganese homeostasis	3	5
<i>nanR</i>	SABB_01531	Utilization of sialic acid and N-acetylglucosamine	7	8
<i>nreC</i>	SABB_01284	Nitrate/nitrite respiration	11	13
<i>pmtR</i>	SABB_03881	Multidrug resistance	4	5
<i>purR</i>	SABB_05027	Purine metabolism	19	22
<i>rex</i>	SABB_02501	Energy metabolism	32	40
<i>srrA</i>	SABB_03639	Response to oxygen availability	54	80

adhesion (Caiazza and O'Toole, 2003). The last relevant group for the biofilm is the non-specific cytoplasmic proteins released and incorporated into the biofilm matrix during autolysis. Autolysis is mediated by hydrolases with emphasis on Atl which is essential to the early stages of biofilm by promoting primary adhesion to surfaces and is present in the set of DEGs in biofilm (Houston et al., 2011). Among the cytoplasmic proteins, we can highlight GAPDH and *eno*-encoded enolase as significant in biofilm formation, where both belong to the set of DEGs in biofilm (Foulston et al., 2014).

Finally, we mapped the set of 1,794 proteins with expression altered in the biofilm against the transcription factor repertoire of the Bmb9393 and its GRN to improve understanding of the regulatory circuits that modulate the biofilm formation process under diverse conditions. It is important to mention that all DEG analyses considered only the expression pattern shared among Bmb9393 and the other strains from the selected transcriptomes. Comparison with the transcription factor repertoire resulted in 113 proteins in common with the set of DEGs in biofilm, representing about 62% of the Bmb9393 regulators. As discussed in the functional analysis, high regulatory activity is necessary to respond to the recurrent environmental changes occurring in the heterogeneous microenvironments and well-defined stages of biofilm. The mapping of the 113 regulators against the Bmb9393 GRN revealed 45 shared TFs. Among them, we highlight the inducer WalR and the inhibitor Rsp, besides AgrA, which acts in the quorum-sensing of *S. aureus*, and the response regulator SaeR. The TCS SaeRS acts depending on the strain and environmental conditions to inhibit or induce biofilm formation. In the Newman strain, constitutive expression of SaeRS prevents the production of a robust biofilm, whereas in the USA300 strain is essential for induction of fermentative biofilm (Cue et al., 2015; Mashruwala et al., 2017). Importantly, DEGs in biofilm contains two of the largest regulons of Bmb9393, SarA, and CcpA. A recent paper by Bullock et al. (2022) with the UAMS-1 strain showed that inactivation of CcpA reduced eDNA incorporation into the biofilm matrix (Bullock et al., 2022). Comparison of the 1,794 DEGs in biofilm against the non-redundant set of 1,151 target genes of the GRN resulted in 876 targets with altered expression in the biofilm, of which 370 were regulated by more than one TF and σ factor (Table 3). Growth in biofilm modulated the gene expression of at least one target gene in each regulon of the Bmb9393 GRN, except for BlaI, showing the variability of regulatory response depending on the strain and/or condition.

4. Conclusions

The regulation of virulence factors in *S. aureus* is the result of a complex and intricate network of regulatory proteins that

respond to specific environmental stimuli in infection sites to integrate metabolism and virulence making *S. aureus* a pathogen remarkably adaptive and versatile. Our work described the reconstruction process and structural analysis of the larger *S. aureus* GRN published up to date, comprising 64 TFs connected to 1,151 target genes through 1,803 interactions, of which 610 were new predictions based on the Bmb9393 genome or whose regulatory roles were not yet characterized and await experimental validation. Bmb9393 GRN possesses 19 new regulons compared to N315 related to key cellular functions such as cell wall metabolism, biofilm formation, and virulence. To the best of our knowledge, this is the first GRN inference performed based on the genome of a Brazilian epidemic clone belonging to the ST239-SCCmecIII lineage with worldwide dispersion that poses relevant phenotypic traits such as the superior ability to accumulate *ica*-independent biofilm. To understand the regulatory impact of bacterial biofilm growth versus planktonic lifestyle, we performed the functional classification of a set of DEGs from several transcriptomes available in the literature and mapped them to the Bmb9393 GRN. The classification of 1,794 proteins showed classes relevant to biofilm formation, and the mapping revealed the extensive impact on regulators and targets, where 45 out of 64 TFs were considered DEGs in any experiment, and regulons had at least one target gene differentially expressed, except the BlaI. Our reconstruction represents the most recent and accurate GRN and serves as an essential starting point for the integration with other biological networks, such as metabolic or cell signaling, and/or experimental data aiming the generation of even more accurate models for this important pathogen.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Author contributions

MC performed the identification of the transcription factor repertoire, GRN reconstruction, identification and functional characterization of DEGs in biofilm, and the overlapping of DEGs with the GRN of Bmb9393. EP-R performed the classification into regulatory families. YM developed the text mining pipeline. AN performed the topological analysis and identification of motifs. MC, AN, and MN performed the network curation efforts. MN, EP-R, MS, and AF supervised the work. MC wrote the manuscript with inputs from AN, and critical feedback from YM, MS, AF, EP-R, and MN. All

authors have read and agreed to the published version of the manuscript.

Funding

MN was supported by a fellowship from CNPq (process no. 306894/2019-0) and grant by CAPES (process no. 88887.368759/2019-00) and FAPERJ (E-26/010.001675/2019). This work was supported to AF in part by (i) the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), 307672/2019; (ii) the Fundação Carlos Chagas Filho de Apoio à Ciência (FAPERJ), E-26/210.875/2016, E-26/210.110/2018, E-26/211.554/2019, and E-26/200.952/2021. EP-R was supported by the DGAPA of the Universidad Nacional Autónoma de México (IN-220523) and CONACYT (320012).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships

that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.1049819/full#supplementary-material>

References

- Alon, U. (2007). Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* 8, 450–461. doi: 10.1038/nrg2102
- Balleza, E., López-Bojorquez, L. N., Martínez-Antonio, A., Resendis-Antonio, O., Lozada-Chávez, I., Balderas-Martínez, Y. I., et al. (2009). Regulation by transcription factors in bacteria: beyond description. *FEMS Microbiol. Rev.* 33, 133–151. doi: 10.1111/j.1574-6976.2008.00145.x
- Barabási, A.-L., and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* 5, 101–113. doi: 10.1038/nrg1272
- Barakat, M., Ortet, P., and Whitworth, D. E. (2013). P2RP: a Web-based framework for the identification and analysis of regulatory proteins in prokaryotic genomes. *BMC Genomics* 14, 269. doi: 10.1186/1471-2164-14-269
- Bardi, T., Pintado, V., Gomez-Rojo, M., Escudero-Sanchez, R., Azzam Lopez, A., Diez-Remesal, Y., et al. (2021). Nosocomial infections associated to COVID-19 in the intensive care unit: clinical characteristics and outcome. *Eur. J. Clin. Microbiol. Infect. Dis.* 40, 495–502. doi: 10.1007/s10096-020-04142-w
- Beenken, K. E., Dunman, P. M., McAleese, F., Macapagal, D., Murphy, E., Projan, S. J., et al. (2004). Global gene expression in *Staphylococcus aureus* biofilms. *J. Bacteriol.* 186, 4665–4684. doi: 10.1128/JB.186.14.4665-4684.2004
- Bischoff, M., Dunman, P., Kormanec, J., Macapagal, D., Murphy, E., Mounts, W., et al. (2004). Microarray-based analysis of the *Staphylococcus aureus* sigmaB regulon. *J. Bacteriol.* 186, 4085–4099. doi: 10.1128/JB.186.13.4085-4099.2004
- Blevins, J. S., Beenken, K. E., Elasri, M. O., Hurlburt, B. K., and Smeltzer, M. S. (2002). Strain-dependent differences in the regulatory roles of sarA and agr in *Staphylococcus aureus*. *Infect. Immun.* 70, 470–480. doi: 10.1128/IAI.70.2.470-480.2002
- Brady, R. A., Leid, J. G., Camper, A. K., Costerton, J. W., and Shirtliff, M. E. (2006). Identification of *Staphylococcus aureus* proteins recognized by the antibody-mediated immune response to a biofilm infection. *Infect. Immun.* 74, 3415–3426. doi: 10.1128/IAI.00392-06
- Bulock, L. L., Ahn, J., Shinde, D., Pandey, S., Sarmiento, C., Thomas, V. C., et al. (2022). Interplay of CodY and CcpA in regulating central metabolism and biofilm formation in *Staphylococcus aureus*. *J. Bacteriol.* 204, e0061721. doi: 10.1128/jb.00617-21
- Cai, X., Bazerque, J. A., and Giannakis, G. B. (2013). Inference of gene regulatory networks with sparse structural equation models exploiting genetic perturbations. *PLoS Comput. Biol.* 9, e1003068. doi: 10.1371/journal.pcbi.1003068
- Caiazza, N. C., and O'Toole, G. A. (2003). Alpha-toxin is required for biofilm formation by *Staphylococcus aureus*. *J. Bacteriol.* 185, 3214–3217. doi: 10.1128/JB.185.10.3214-3217.2003
- Chastanet, A., Fert, J., and Msadek, T. (2003). Comparative genomics reveal novel heat shock regulatory mechanisms in *Staphylococcus aureus* and other Gram-positive bacteria. *Mol. Microbiol.* 47, 1061–1073. doi: 10.1046/j.1365-2958.2003.03355.x
- Chaudhuri, R. R., Allen, A. G., Owen, P. J., Shalom, G., Stone, K., Harrison, M., et al. (2009). Comprehensive identification of essential *Staphylococcus aureus* genes using transposon-mediated differential hybridisation (TMDH). *BMC Genomics* 10, 291. doi: 10.1186/1471-2164-10-291
- Costa, M. O. C., Beltrame, C. O., Ferreira, F. A., Botelho, A. M. N., Lima, N. C. B., Souza, R. C., et al. (2013). Complete genome sequence of a variant of the methicillin-resistant *Staphylococcus aureus* ST239 lineage, strain BMB9393, displaying superior ability to accumulate ica-independent biofilm. *Genome Announc.* 1, e00576-e00513. doi: 10.1128/genomeA.00576-13
- Crosby, H. A., Schlievert, P. M., Merriman, J. A., King, J. M., Salgado-Pabón, W., and Horswill, A. R. (2016). The *Staphylococcus aureus* global regulator MgrA modulates clumping and virulence by controlling surface protein expression. *PLoS Pathog* 12, e1005604. doi: 10.1371/journal.ppat.1005604
- Cruz, A. R., Boer, M. A. d., Strasser, J., Zwarthoff, S. A., Beurskens, F. J., et al. (2021). Staphylococcal protein A inhibits complement activation by interfering with IgG hexamer formation. *Proc. Natl. Acad. Sci. U.S.A.* 118, e201672118. doi: 10.1073/pnas.2016721118
- Csardi, G., and Nepusz, T. (2006). The igraph software package for complex network research. *Complex Syst. Interf.* 1695, 1–9.
- Cue, D., Junecko, J. M., Lei, M. G., Blevins, J. S., Smeltzer, M. S., and Lee, C. Y. (2015). SaeRS-dependent inhibition of biofilm formation in *Staphylococcus aureus* Newman. *PLoS ONE* 10, e0123027. doi: 10.1371/journal.pone.0123027
- Cue, D., Lei, M. G., Luong, T. T., Kuechenmeister, L., Dunman, P. M., O'Donnell, S., et al. (2009). Rbf promotes biofilm formation by *Staphylococcus aureus* via repression of icaR, a negative regulator of icaADBC. *J. Bacteriol.* 191, 6363–6373. doi: 10.1128/JB.00913-09
- DeFrancesco, A. S., Masloboeva, N., Syed, A. K., DeLoughery, A., Bradshaw, N., Li, G.-W., et al. (2017). Genome-wide screen for genes involved in eDNA release

- during biofilm formation by *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. U.S.A.* 114, E5969–E5978. doi: 10.1073/pnas.1704544114
- Delauné, A., Dubrac, S., Blanchet, C., Poupel, O., Mäder, U., Hiron, A., et al. (2012). The WalkR system controls major staphylococcal virulence genes and is involved in triggering the host inflammatory response. *Infect. Immun.* 80, 3438–3453. doi: 10.1128/IAI.00195-12
- Deochand, D. K., and Grove, A. (2017). MarR family transcription factors: dynamic variations on a common scaffold. *Crit. Rev. Biochem. Mol. Biol.* 52, 595–613. doi: 10.1080/10409238.2017.1344612
- Dudek, C.-A., and Jahn, D. (2022). PRODORIC: state-of-the-art database of prokaryotic gene regulation. *Nucleic Acids Res.* 50, D295–D302. doi: 10.1093/nar/gkab1110
- Faria, J. P., Overbeek, R., Taylor, R. C., Conrad, N., Vonstein, V., Goelzer, A., et al. (2016). Reconstruction of the regulatory network for bacillus subtilis and reconciliation with gene expression data. *Front. Microbiol.* 7, 275. doi: 10.3389/fmicb.2016.00275
- Faria, J. P., Overbeek, R., Xia, F., Rocha, M., Rocha, I., and Henry, C. S. (2014). Genome-scale bacterial transcriptional regulatory networks: reconstruction and integrated analysis with metabolic models. *Brief. Bioinform.* 15, 592–611. doi: 10.1093/bib/bbs071
- Figueiredo, A. M. S., Ferreira, F. A., Beltrame, C. O., and Côrtes, M. F. (2017). The role of biofilms in persistent infections and factors involved in ica-independent biofilm development and gene regulation in *Staphylococcus aureus*. *Crit. Rev. Microbiol.* 43, 602–620. doi: 10.1080/1040841X.2017.1282941
- Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., et al. (2014). Pfam: the protein families database. *Nucleic Acids Res.* 42, D222–D230. doi: 10.1093/nar/gkt1223
- Foster, T. J., Geoghegan, J. A., Ganesh, V. K., and Höök, M. (2014). Adhesion, invasion and evasion: the many functions of the surface proteins of *Staphylococcus aureus*. *Nat. Rev. Microbiol.* 12, 49–62. doi: 10.1038/nrmicro3161
- Foulston, L., Elsholz, A. K. W., DeFrancesco, A. S., and Losick, R. (2014). The extracellular matrix of *Staphylococcus aureus* biofilms comprises cytoplasmic proteins that associate with the cell surface in response to decreasing pH. *MBio* 5, e01667–e01614. doi: 10.1128/mBio.01667-14
- Gimza, B. D., Larias, M. I., Budny, B. G., and Shaw, L. N. (2019). Mapping the global network of extracellular protease regulation in *Staphylococcus aureus*. *mSphere* 4, e00676–e00619. doi: 10.1128/mSphere.00676-19
- Goenawan, I. H., Bryan, K., and Lynn, D. J. (2016). DyNet: visualization and analysis of dynamic molecular interaction networks. *Bioinformatics* 32, 2713–2715. doi: 10.1093/bioinformatics/btw187
- Görke, B., and Stülke, J. (2008). Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat. Rev. Microbiol.* 6, 613–624. doi: 10.1038/nrmicro1932
- Gründling, A., and Schneewind, O. (2007). Synthesis of glycerol phosphate lipoteichoic acid in *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 8478–8483. doi: 10.1073/pnas.0701821104
- Guldemann, C., Boor, K. J., Wiedmann, M., and Guariglia-Oropeza, V. (2016). Resilience in the face of uncertainty: sigma factor B fine-tunes gene expression to support homeostasis in gram-positive bacteria. *Appl. Environ. Microbiol.* 82, 4456–4469. doi: 10.1128/AEM.00714-16
- Harris, S. R., Feil, E. J., Holden, M. T. G., Quail, M. A., Nickerson, E. K., Chantratita, N., et al. (2010). Evolution of MRSA during hospital transmission and intercontinental spread. *Science* 327, 469–474. doi: 10.1126/science.1182395
- Houston, P., Rowe, S. E., Pozzi, C., Waters, E. M., and O’Gara, J. P. (2011). Essential role for the major autolysin in the fibronectin-binding protein-mediated *Staphylococcus aureus* biofilm phenotype. *Infect. Immun.* 79, 1153–1165. doi: 10.1128/IAI.00364-10
- Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., von Mering, C., et al. (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Mol. Biol. Evol.* 34, 2115–2122. doi: 10.1093/molbev/msx148
- Ibarra, J. A., Pérez-Rueda, E., Carroll, R. K., and Shaw, L. N. (2013). Global analysis of transcriptional regulators in *Staphylococcus aureus*. *BMC Genomics* 14, 126. doi: 10.1186/1471-2164-14-126
- Ishii, K., Adachi, T., Yasukawa, J., Suzuki, Y., Hamamoto, H., and Sekimizu, K. (2014). Induction of virulence gene expression in *Staphylococcus aureus* by pulmonary surfactant. *Infect. Immun.* 82, 1500–1510. doi: 10.1128/IAI.01635-13
- Janga, S. C., and Moreno-Hagelsieb, G. (2004). Conservation of adjacency as evidence of paralogous operons. *Nucleic Acids Res.* 32, 5392–5397. doi: 10.1093/nar/gkh882
- Jenul, C., and Horswill, A. R. (2019). Regulation of *Staphylococcus aureus* virulence. *Microbiol. Spectrum* 7, 2018. doi: 10.1128/microbiolspec.GPP3-0031-2018
- Jiang, Q., Jin, Z., and Sun, B. (2018). MgrA negatively regulates biofilm formation and detachment by repressing the expression of psm operons in *Staphylococcus aureus*. *Appl. Environ. Microbiol.* 84, e01008–e01018. doi: 10.1128/AEM.01008-18
- Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30, 1236–1240. doi: 10.1093/bioinformatics/btu031
- Kanehisa, M., Sato, Y., and Kawashima, M. (2022). KEGG mapping tools for uncovering hidden features in biological data. *Protein Sci.* 31, 47–53. doi: 10.1002/pro.4172
- Kanehisa, M., Sato, Y., and Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.* 428, 726–731. doi: 10.1016/j.jmb.2015.11.006
- Kiliç, S., White, E. R., Sagitova, D. M., Cornish, J. P., and Erill, I. (2014). CollecTF: a database of experimentally validated transcription factor-binding sites in Bacteria. *Nucleic Acids Res.* 42, D156–D160. doi: 10.1093/nar/gkt1123
- Kinkel, T. L., Roux, C. M., Dunman, P. M., and Fang, F. C. (2013). The *Staphylococcus aureus* SrrAB two-component system promotes resistance to nitrosative stress and hypoxia. *MBio* 4, e00696–e00613. doi: 10.1128/mBio.00696-13
- Koutrouli, M., Karatzas, E., Paez-Espino, D., and Pavlopoulos, G. A. (2020). A guide to conquer the biological network era using graph theory. *Front. Bioeng. Biotechnol.* 8, 34. doi: 10.3389/fbioe.2020.00034
- Lechner, M., Findeiss, S., Steiner, L., Marz, M., Stadler, P. F., and Prohaska, S. J. (2011). Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinform.* 12, 124. doi: 10.1186/1471-2105-12-124
- Lei, M. G., Cue, D., Roux, C. M., Dunman, P. M., and Lee, C. Y. (2011). Rsp inhibits attachment and biofilm formation by repressing fnbA in *Staphylococcus aureus* MW2. *J. Bacteriol.* 193, 5231–5241. doi: 10.1128/JB.05454-11
- Lei, M. G., and Lee, C. Y. (2018). Repression of capsule production by XdrA and CodY in *Staphylococcus aureus*. *J. Bacteriol.* 200, e00203–e00218. doi: 10.1128/JB.00203-18
- Lei, M. G., and Lee, C. Y. (2020). MgrA activates staphylococcal capsule via SigA-dependent promoter. *J. Bacteriol.* 203, e00495–e00420. doi: 10.1128/JB.00495-20
- Li, T., He, L., Song, Y., Villaruz, A. E., Joo, H.-S., Liu, Q., et al. (2015). AraC-type regulator Rsp adapts *Staphylococcus aureus* gene expression to acute infection. *Infect. Immun.* 84, 723–734. doi: 10.1128/IAI.01088-15
- Loughran, A. J., Atwood, D. N., Anthony, A. C., Harik, N. S., Spencer, H. J., Beenken, K. E., et al. (2014). Impact of individual extracellular proteases on *Staphylococcus aureus* biofilm formation in diverse clinical isolates and their isogenic sarA mutants. *Microbiologyopen* 3, 897–909. doi: 10.1002/mbo3.214
- Luong, T. T., Dunman, P. M., Murphy, E., Projan, S. J., and Lee, C. Y. (2006). Transcription Profiling of the mgrA regulon in *Staphylococcus aureus*. *J. Bacteriol.* 188, 1899–1910. doi: 10.1128/JB.188.5.1899-1910.2006
- Mangan, S., and Alon, U. (2003). Structure and function of the feed-forward loop network motif. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11980–11985. doi: 10.1073/pnas.2133841100
- Maree, M., Thi Nguyen, L. T., Ohniwa, R. L., Higashide, M., Msadek, T., and Morikawa, K. (2022). Natural transformation allows transfer of SCCmec-mediated methicillin resistance in *Staphylococcus aureus* biofilms. *Nat. Commun.* 13, 2477. doi: 10.1038/s41467-022-29877-2
- Mashruwala, A. A., Guchte, A. V. D., and Boyd, J. M. (2017). Impaired respiration elicits SrrAB-dependent programmed cell lysis and biofilm formation in *Staphylococcus aureus*. *Elife* 6, e23845. doi: 10.7554/eLife.23845
- McCallum, N., Hinds, J., Ender, M., Berger-Bächli, B., and Stutzmann Meier, P. (2010). Transcriptional profiling of XdrA, a new regulator of spa transcription in *Staphylococcus aureus*. *J. Bacteriol.* 192, 5151–5164. doi: 10.1128/JB.00491-10
- McCarthy, A. J., and Lindsay, J. A. (2010). Genetic variation in *Staphylococcus aureus* surface and immune evasion genes is lineage associated: implications for vaccine design and host-pathogen interactions. *BMC Microbiol.* 10, 173. doi: 10.1186/1471-2180-10-173
- Mercatelli, D., Scalambra, L., Triboli, L., Ray, F., and Giorgi, F. M. (2020). Gene regulatory network inference resources: a practical overview. *Bioch. Biophys. Acta Gene Regul. Mech.* 1863, 194430. doi: 10.1016/j.bbaggm.2019.194430
- Miller, H. K., Carroll, R. K., Burda, W. N., Krute, C. N., Davenport, J. E., and Shaw, L. N. (2012). The extracytoplasmic function sigma factor sigma(s) protects against both intracellular and extracytoplasmic stresses in *Staphylococcus aureus*. *J. Bacteriol.* 194, 4342–4354. doi: 10.1128/JB.00484-12

- Mootz, J. M., Malone, C. L., Shaw, L. N., and Horswill, A. R. (2013). Staphopains modulate *Staphylococcus aureus* biofilm integrity. *Infect. Immun.* 81, 3227–3238. doi: 10.1128/IAI.00377-13
- Morikawa, K., Takemura, A. J., Inose, Y., Tsai, M., Nguyen Thi, L. T., Ohta, T., et al. (2012). Expression of a cryptic secondary sigma factor gene unveils natural competence for DNA transformation in *Staphylococcus aureus*. *PLoS Pathog* 8, e1003003. doi: 10.1371/journal.ppat.1003003
- Nagarajan, V., and Elasm, M. O. (2007). SAMMD: *Staphylococcus aureus* microarray meta-database. *BMC Genomics* 8, 351. doi: 10.1186/1471-2164-8-351
- Nakano, M. M., Zuber, P., Glaser, P., Danchin, A., and Hulett, F. M. (1996). Two-component regulatory proteins ResD-ResE are required for transcriptional activation of *fmr* upon oxygen limitation in *Bacillus subtilis*. *J. Bacteriol.* 178, 3796–3802. doi: 10.1128/jb.178.13.3796-3802.1996
- Novichkov, P. S., Kazakov, A. E., Ravcheev, D. A., Leyn, S. A., Kovaleva, G. Y., Sutormin, R. A., et al. (2013). RegPrecise 3.0—a resource for genome-scale exploration of transcriptional regulation in bacteria. *BMC Genomics* 14, 745. doi: 10.1186/1471-2164-14-745
- Oliveira, D. C., and de Lencastre, H. (2011). Methicillin-resistance in *Staphylococcus aureus* is not affected by the overexpression in trans of the *mecA* gene repressor: a surprising observation. *PLoS ONE* 6, e23287. doi: 10.1371/journal.pone.0023287
- O'Neill, E., Pozzi, C., Houston, P., Smyth, D., Humphreys, H., Robinson, D. A., et al. (2007). Association between methicillin susceptibility and biofilm regulation in *Staphylococcus aureus* isolates from device-related infections. *J. Clin. Microbiol.* 45, 1379–1388. doi: 10.1128/JCM.02280-06
- Oriol, C., Cengher, L., Manna, A. C., Mauro, T., Pinel-Marie, M.-L., Felden, B., et al. (2021). Expanding the *Staphylococcus aureus* SarA regulon to small RNAs. *mSystems* 6, e0071321. doi: 10.1128/mSystems.00713-21
- Ortet, P., De Luca, G., Whitworth, D. E., and Barakat, M. (2012). P2TF: a comprehensive resource for analysis of prokaryotic transcription factors. *BMC Genomics* 13, 628. doi: 10.1186/1471-2164-13-628
- Pagels, M., Fuchs, S., Pané-Farré, J., Kohler, C., Menschner, L., Hecker, M., et al. (2010). Redox sensing by a Rex-family repressor is involved in the regulation of an anaerobic gene expression in *Staphylococcus aureus*. *Mol. Microbiol.* 76, 1142–1161. doi: 10.1111/j.1365-2958.2010.07105.x
- Paget, M. S. B., and Helmann, J. D. (2003). The sigma70 family of sigma factors. *Genome Biol.* 4, 203. doi: 10.1186/gb-2003-4-1-203
- Pané-Farré, J., Jonas, B., Förstner, K., Engelmann, S., and Hecker, M. (2006). The sigmaB regulon in *Staphylococcus aureus* and its regulation. *Int. J. Med. Microbiol.* 296, 237–258. doi: 10.1016/j.ijmm.2005.11.011
- Peng, H., Hu, Q., Shang, W., Yuan, J., Zhang, X., Liu, H., et al. (2017). Walk(S221P), a naturally occurring mutation, confers vancomycin resistance in VISA strain XN108. *J. Antimicrob. Chemother.* 72, 1006–1013. doi: 10.1093/jac/dkw518
- Pohl, K., Francois, P., Stenz, L., Schlink, F., Geiger, T., Herbert, S., et al. (2009). CodY in *Staphylococcus aureus*: a regulatory link between metabolism and virulence gene expression. *J. Bacteriol.* 191, 2953–2963. doi: 10.1128/JB.01492-08
- Pragman, A. A., Yarwood, J. M., Tripp, T. J., and Schlievert, P. M. (2004). Characterization of virulence factor regulation by SrrAB, a two-component system in *Staphylococcus aureus*. *J. Bacteriol.* 186, 2430–2438. doi: 10.1128/JB.186.8.2430-2438.2004
- Ravcheev, D. A., Best, A. A., Tintle, N., Dejongh, M., Osterman, A. L., Novichkov, P. S., et al. (2011). Inference of the transcriptional regulatory network in *Staphylococcus aureus* by integration of experimental and genomics-based evidence. *J. Bacteriol.* 193, 3228–3240. doi: 10.1128/JB.00350-11
- Resch, A., Rosenstein, R., Nerz, C., and Götz, F. (2005). Differential gene expression profiling of *Staphylococcus aureus* cultivated under biofilm and planktonic conditions. *Appl. Environ. Microbiol.* 71, 2663–2676. doi: 10.1128/AEM.71.5.2663-2676.2005
- Rice, K. C., Mann, E. E., Endres, J. L., Weiss, E. C., Cassat, J. E., Smeltzer, M. S., et al. (2007). The *cidA* murein hydrolase regulator contributes to DNA release and biofilm development in *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 8113–8118. doi: 10.1073/pnas.0610226104
- Rodionov, D. A. (2007). Comparative genomic reconstruction of transcriptional regulatory networks in bacteria. *Chem. Rev.* 107, 3467–3497. doi: 10.1021/cr068309+
- Santana-García, W., Castro-Mondragon, J. A., Padilla-Gálvez, M., Nguyen, N. T. T., Elizondo-Salas, A., Ksouri, N., et al. (2022). RSAT 2022: regulatory sequence analysis tools. *Nucleic Acids Res.* 2022, gkac312. doi: 10.1093/nar/gkac312
- Santiago, M., Matano, L. M., Moussa, S. H., Gilmore, M. S., Walker, S., and Meredith, T. C. (2015). A new platform for ultra-high density *Staphylococcus aureus* transposon libraries. *BMC Genomics* 16, 252. doi: 10.1186/s12864-015-1361-3
- Sanz, J., Navarro, J., Arbués, A., Martín, C., Marijuán, P. C., and Moreno, Y. (2011). The transcriptional regulatory network of *Mycobacterium tuberculosis*. *PLoS ONE* 6, e22178. doi: 10.1371/journal.pone.0022178
- Schilcher, K., and Horswill, A. R. (2020). Staphylococcal biofilm development: structure, regulation, and treatment strategies. *Microbiol. Mol. Biol. Rev.* 84, e00026-e00019. doi: 10.1128/MMBR.00026-19
- Seidl, K., Müller, S., François, P., Kriebitzsch, C., Schrenzel, J., Engelmann, S., et al. (2009). Effect of a glucose impulse on the CcpA regulon in *Staphylococcus aureus*. *BMC Microbiol.* 9, 95. doi: 10.1186/1471-2180-9-95
- Seidl, K., Stucki, M., Rugg, M., Goerke, C., Wolz, C., Harris, L., et al. (2006). *Staphylococcus aureus* CcpA affects virulence determinant production and antibiotic resistance. *Antimicrob. Agents Chemother.* 50, 1183–1194. doi: 10.1128/AAC.50.4.1183-1194.2006
- Shanks, R. M. Q., Meehl, M. A., Brothers, K. M., Martinez, R. M., Donegan, N. P., Graber, M. L., et al. (2008). Genetic evidence for an alternative citrate-dependent biofilm formation pathway in *Staphylococcus aureus* that is dependent on fibronectin binding proteins and the GraRS two-component regulatory system. *Infect. Immun.* 76, 2469–2477. doi: 10.1128/IAI.01370-07
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Shaw, L. N., Lindholm, C., Prajsnar, T. K., Miller, H. K., Brown, M. C., Golonka, E., et al. (2008). Identification and characterization of sigma, a novel component of the *Staphylococcus aureus* stress and virulence responses. *PLoS ONE* 3, e3844. doi: 10.1371/journal.pone.0003844
- Shen, W., Le, S., Li, Y., and Hu, F. (2016). SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS ONE* 11, e0163962. doi: 10.1371/journal.pone.0163962
- Stewart, P. S., and Franklin, M. J. (2008). Physiological heterogeneity in biofilms. *Nat. Rev. Microbiol.* 6, 199–210. doi: 10.1038/nrmicro1838
- Stewart, P. S., Williamson, K. S., Boegli, L., Hamerly, T., White, B., Scott, L., et al. (2022). Search for a shared genetic or biochemical basis for biofilm tolerance to antibiotics across bacterial species. *Antimicrob. Agents Chemother.* 66, e0002122. doi: 10.1128/aac.00021-22
- Stone, L., Simberloff, D., and Artzy-Randrup, Y. (2019). Network motifs and their origins. *PLoS Comput. Biol.* 15, e1006749. doi: 10.1371/journal.pcbi.1006749
- Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE* 6, e21800. doi: 10.1371/journal.pone.0021800
- Tao, L., Wu, X., and Sun, B. (2010). Alternative sigma factor sigmaH modulates prophage integration and excision in *Staphylococcus aureus*. *PLoS Pathog* 6, e1000888. doi: 10.1371/journal.ppat.1000888
- Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28, 33–36. doi: 10.1093/nar/28.1.33
- The Gene Ontology Consortium (2019). The gene ontology resource: 20 years and still GOing strong. *Nucleic Acids Res.* 47, D330-D338. doi: 10.1093/nar/gky1055
- Throup, J. P., Zappacosta, F., Lunsford, R. D., Annan, R. S., Carr, S. A., Lonsdale, J. T., et al. (2001). The *srhSR* gene pair from *Staphylococcus aureus*: genomic and proteomic approaches to the identification and characterization of gene function. *Biochemistry* 40, 10392–10401. doi: 10.1021/bi0102959
- Tomlinson, B. R., Malof, M. E., and Shaw, L. N. (2021). A global transcriptomic analysis of *Staphylococcus aureus* biofilm formation across diverse clonal lineages. *Microbial Genomics* 7, 598. doi: 10.1099/mgen.0.000598
- Troitzsch, A., Loi, V. V., Methling, K., Zühlke, D., Lalk, M., Riedel, K., et al. (2021). Carbon source-dependent reprogramming of anaerobic metabolism in *Staphylococcus aureus*. *J. Bacteriol.* 203, e00639-e00620. doi: 10.1128/JB.00639-20
- UniProt Consortium (2019). UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* 47, D506-D515. doi: 10.1093/nar/gky1049
- Waters, N. R., Samuels, D. J., Behera, R. K., Livny, J., Rhee, K. Y., Sadykov, M. R., et al. (2016). A spectrum of CodY activities drives metabolic reorganization and virulence gene expression in *Staphylococcus aureus*. *Mol. Microbiol.* 101, 495–514. doi: 10.1111/mmi.13404
- Yao, X., and Lu, C.-D. (2014). Functional characterization of the potRABCD operon for spermine and spermidine uptake and regulation in *Staphylococcus aureus*. *Curr. Microbiol.* 69, 75–81. doi: 10.1007/s00284-014-0556-1