



Joint Genomic and Proteomic Analysis Identifies Meta-Trait Characteristics of Virulent and Non-virulent *Staphylococcus aureus* Strains

OPEN ACCESS

Edited by:

Francois Vandenesch,
Université de Lyon, France

Reviewed by:

Nick Wheelhouse,
Edinburgh Napier University,
United Kingdom
Stefano Giulieri,
Université de Lausanne, Switzerland

*Correspondence:

Benedykt Wladyka
benedykt.wladyka@uj.edu.pl

†These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Molecular Bacterial Pathogenesis,
a section of the journal
Frontiers in Cellular and Infection
Microbiology

Received: 10 May 2018

Accepted: 16 August 2018

Published: 06 September 2018

Citation:

Bonar EA, Bukowski M, Hydzik M,
Jankowska U, Kedracka-Krok S,
Groborz M, Dubin G, Akkerboom V,
Miedzobrodzki J, Sabat AJ,
Friedrich AW and Wladyka B (2018)
Joint Genomic and Proteomic Analysis
Identifies Meta-Trait Characteristics of
Virulent and Non-virulent
Staphylococcus aureus Strains.
Front. Cell. Infect. Microbiol. 8:313.
doi: 10.3389/fcimb.2018.00313

Emilia A. Bonar^{1†}, Michal Bukowski^{1†}, Marcin Hydzik¹, Urszula Jankowska²,
Sylvia Kedracka-Krok³, Magdalena Groborz¹, Grzegorz Dubin^{2,4}, Viktoria Akkerboom⁵,
Jacek Miedzobrodzki⁴, Artur J. Sabat⁵, Alexander W. Friedrich⁵ and Benedykt Wladyka^{1*}

¹ Department of Analytical Biochemistry, Faculty of Biochemistry, Biophysics and Biotechnology, Jagiellonian University, Krakow, Poland, ² Malopolska Centre of Biotechnology, Jagiellonian University, Krakow, Poland, ³ Department of Physical Biochemistry, Faculty of Biochemistry, Biophysics and Biotechnology, Jagiellonian University, Krakow, Poland, ⁴ Department of Microbiology, Faculty of Biochemistry, Biophysics and Biotechnology, Jagiellonian University, Krakow, Poland, ⁵ Department of Medical Microbiology, University Medical Center Groningen, University of Groningen, Groningen, Netherlands

Staphylococcus aureus is an opportunistic pathogen of humans and warm-blooded animals and presents a growing threat in terms of multi-drug resistance. Despite numerous studies, the basis of staphylococcal virulence and switching between commensal and pathogenic phenotypes is not fully understood. Using genomics, we show here that *S. aureus* strains exhibiting virulent (VIR) and non-virulent (NVIR) phenotypes in a chicken embryo infection model genetically fall into two separate groups, with the VIR group being much more cohesive than the NVIR group. Significantly, the genes encoding known staphylococcal virulence factors, such as clumping factors, are either found in different allelic variants in the genomes of NVIR strains (compared to VIR strains) or are inactive pseudogenes. Moreover, the pyruvate carboxylase and gamma-aminobutyrate permease genes, which were previously linked with virulence, are pseudogenized in NVIR strain ch22. Further, we use comprehensive proteomics tools to characterize strains that show opposing phenotypes in a chicken embryo virulence model. VIR strain CH21 had an elevated level of diacylglycerol kinase involved in staphyloxanthin production (protection against free radicals) and expressed a higher level of immunoglobulin-binding protein Sbi on its surface compared to NVIR strain ch22. Furthermore, joint genomic and proteomic approaches linked the elevated production of superoxide dismutase and DNA-binding protein by NVIR strain ch22 with gene duplications.

Keywords: genome, genomics, pathogen, proteome, proteomics, *Staphylococcus aureus*, virulence

INTRODUCTION

The genetic determinants and protein effectors that are responsible for the virulence of *Staphylococcus aureus* escape our full understanding despite a number of comprehensive studies. In homeostasis, *S. aureus* coexists with its host without distinguished adverse effects. However, in an imbalanced state, the nature of which is poorly understood, this opportunistic pathogen may cause infection and pose a significant health threat. Thus, the Janus-face bacteria constantly balances commensal and virulent phenotypes, coping with different levels of host defenses (Rasigade and Vandenesch, 2014). Indeed, it was recently demonstrated that within the same clonal complex, phenotypic differences may be linked with the severity of infections. Moreover, factors correlated with high pathogenicity in the group of genetically related *S. aureus* had little effect on the mortality rates associated with infections caused by bacteria from other clonal complexes (Recker et al., 2017). This finding indicates both the genetic and phenotypic basis of staphylococcal virulence. Aside from maintaining host/pathogen balance in a single host species, staphylococci have been demonstrated to switch between animal and human hosts. Such switching is associated with the exchange of host-specific virulence factors that are responsible for colonization and spread (Lowder et al., 2009). This plasticity significantly complicates studies on virulence determinants, especially in terms of likely human specific factors that can be experimentally tested exclusively in animal models. Genetic methods have been successfully used to predict antibiotic resistance with high credibility and the recent advent of massive parallel sequencing promises clinical utility (Aanensen et al., 2016). However, only a few genetic markers, whose mechanism of action has been determined at the molecular level, have been convincingly linked with successful colonization and virulence [e.g., arginine catabolism mobile element, ACME (Diep et al., 2008; Thurlow et al., 2013), exfoliative toxins (Bukowski et al., 2010)]. Other genetic markers, based on statistical analysis of whole genome sequencing and DNA microarray assays data, were pointed to positively correlate with severity of infections, including bacteremia and infective endocarditis (Gill et al., 2011; Calderwood et al., 2014; Bouchiat et al., 2015). Nevertheless, a palette of genetic predictors of staphylococcal virulence is still limited. A number of studies have used proteomics approaches to try to identify the protein effectors of virulent phenotypes (Bonar et al., 2015). Earlier, we compared exoproteomes in a set of strains exhibiting high and low virulence in a chicken embryo infection model. Despite the high heterogeneity of the analyzed proteomes, we were nevertheless able to identify alpha-hemolysin and bifunctional autolysin as indicators of virulence, whereas glutamylendopeptidase production was characteristic of non-virulent strains (Bonar et al., 2016). This prior study, however, did not take into account surface-attached proteins, which may contain additional virulence factors. The intracellular proteome was not taken into consideration either, but it may potentially contain regulatory proteins.

However, another source of information on staphylococcal virulence factors traditionally comes from the investigation of knock-out and knock-in strains devoid of or supplemented with

tested factors and subsequently challenged in animal models (Kim et al., 2014b). Unfortunately, a lack of appropriate models has detrimentally affected the results (Polakowska et al., 2012). Moreover, this approach does not allow to distinguish “true” virulence factors from those affecting the overall fitness of bacteria.

Given the limitations of these current approaches, here, we designed and applied a new, combined workflow to successfully identify virulence- and commensalism-related meta-traits within the genomes and proteomes of virulent (VIR) and non-virulent (NVIR) *S. aureus* strains. Two belonging to the same sequence type wild-type strains that have been well-characterized in terms of virulence in an *in vivo* model were compared and contrasted using a combined genomic and proteomic methodology. We show that the non-virulent strain ch22 is characterized by a more complex exoproteome than its virulent counterpart CH21. This finding is associated with the smaller genome of CH21 than ch22. Interestingly, CH21 is not characterized by the production of any classical virulence factors compared to ch22. It is rather the combined differential expression of multiple factors that determines the virulence of CH21; the rationale behind this conclusion is discussed in our communication.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

Poultry-isolated *S. aureus* strains exhibiting either high (CH3, CH5, CH9, CH21, and CH23) or low (ch22, ch24, pa3, and ph2) virulence (VIR and NVIR, respectively) in a chicken embryo experimental infection model were used in the study (Supplementary Table 1). Strain origin and general genetic and phenotypic characteristics, including basic phylogenetic relationships and virulence, were described previously (Lowder et al., 2009; Polakowska et al., 2012; Bonar et al., 2016). The bacteria were cultured in tryptic soy broth (TSB) for 16 h at 37°C with vigorous shaking unless indicated otherwise.

Genome Sequencing and Assembly

Whole Genome Sequencing

Genomic DNA was isolated using a DNeasy Blood and Tissue Kit (Qiagen) from an overnight culture derived from a single colony. Purified DNA was quantified with a Qubit 2.0 Fluorometer (Life Technologies). Whole genome sequencing was performed using an Illumina MiSeq system with DNA fragment libraries prepared using a Nextera XT v3 kit (Illumina) according to the manufacturer’s protocol. The samples were sequenced to obtain a minimum of 100-fold coverage. Reads were *de novo* assembled into contigs using CLC Genomics Workbench (version 8.5.1). Contigs were ordered on a template of the *S. aureus* ED98 complete chromosome sequence (GenBank CP001781.1) using self-developed Python scripts, which utilized nucleotide BLAST from the NCBI BLAST+ toolkit [version 2.3.0 (Camacho et al., 2009)]. The complete genomic sequences of the CH21 and ch22 strains were obtained by closing the remaining gaps using PCR amplification and Sanger sequencing. Automated genome annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline (<http://www.ncbi.nlm.nih.gov/>)

genome/annotation_prok/). The sequences were deposited in GenBank with the accession numbers: CH3, MOYG00000000; CH5, MSGQ00000000; CH9, MOYH00000000; CH21, CP017804, CP017804, CP017806; ch22, CP017807, CP017808, CP017809; ch23, MOYI00000000; ch24, MOYJ00000000; pa3, MOXP00000000; ph2, MOYK00000000. Detailed information may be found in the **Supplementary Table 1**.

Identification of Mobile Genetic Elements (MGEs)

The contigs that did not match the chromosome sequence were examined for possible plasmid origin by evaluating their similarity to existing plasmids or fragments thereof. Whole genomic sequences were examined for known staphylococcal phages and pathogenicity islands using an exhaustive set of reference sequences obtained from GenBank (**Supplementary Table 2**). Short sequence fragments of at least 1 kb showing similarity (but not identical) to known phages and pathogenicity islands were classified as putatively novel.

Construction of Phylogenetic Trees

Phylogenetic trees of the identified prophages and pathogenicity islands were constructed using CLC Main Workbench (version 7.7.2) and the k-mer based tree construction method with the Neighbor Join algorithm (k-mer size 15; distance measure: fractional common k-mer count). SNP analysis was performed using the CSI Phylogeny 1.4 server (Kaas et al., 2014). As the chromosome sequences were submitted to the server minimum depth at SNP positions, minimum relative depth at SNP positions, minimum distance between SNPs and minimum SNP quality as input parameters were disabled during analysis. The read mapping quality was set to minimum 25 and the z-score to 1.96. The maximum likelihood tree produced by CSI Phylogeny 1.4 server was visualized in MEGA6 software (Tamura et al., 2013).

In silico MLST and ST5 Group Phylogenetic Analysis

A pipeline was developed in Python 3 for *in silico* MLST and phylogenetic analysis of strains belonging to ST5 group. All 8,688 *S. aureus* genomes, complete genomic sequences as well as shotgun sequencing results, were fetched from NCBI GenBank database (<ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/bacteria/>, accessed on 2018-07-06). 289 genomes which did not provide a cross-reference to NCBI Biosample database were omitted. Loci profiles from BISGs database were used to classify the genomes according to their sequence type (<https://pubmlst.org/saureus/>, accessed on 2018-07-06 [Jolley and Maiden, 2010]). Loci sequences used for *S. aureus* MLST typing were obtained from USA300 strain genomic sequence (accession version CP014420.1) and used as queries for nucleotide BLAST search of the remaining genomes with *E*-value threshold 0.0001 and required query coverage 100% (ver. 2.7.1+ [Camacho et al., 2009]). For 2,129 assemblies, classified as ST5, cross-references to NCBI Biosample database (<ftp://ftp.ncbi.nlm.nih.gov/biosample/>, accessed on 2018-07-06) were used to obtain information on the host, year of isolation and country of origin (attributes of “harmonized_name” property equal to “host,” “geo_loc_name,” and “collection_date,” respectively).

1,635 assemblies contained such information. The information on the reference strains ED98 and N315, missing in NCBI Biosample database, was added manually. These genomes were searched with nucleotide BLAST with *E*-value threshold 0.0001 and minimal query coverage 95% for loci sequences obtained from N315 strain genome (accession version BA000018.3), which were utilized before for ST5 phylogenetic analysis by Nubel et al. (2008). However, to avoid false positive hits, only loci of length equal or greater than 400 bp were selected (in total 97 of 126, **Supplementary Table 10**). For a significant over-representation of strains isolated from human host and originating from the USA (1,248 of 1,474, 85%), a random sample of size 50 was obtained for this group. Selected in such a way 279 strains together with other six being analyzed in this research and belonging to ST5 group (**Supplementary Table 11**) were subjected to phylogenetic analysis. For every strain the sequences corresponding to the aforementioned N315 strain loci were ordered, concatenated and aligned using Clustal Omega with default parameters [ver. 1.2.1 (Sievers and Higgins, 2014)]. Subsequently every column containing a gap was removed from the alignment, which shortened it from 44,295 to 43,872 bp. The alignment was converted to PHYLIP format and used to create a phylogenetic tree with RAxML [ver. 8.2.9 (Stamatakis, 2014)] based on GTRGAMMAI model [general time-reversible, GTR (Tavaré, 1986)]. The tree was subsequently imported to CLC Main Workbench (Qiagen, CLC Bio) together with data obtained from Biosample database, as metadata, and visualized.

Analysis of Coding Sequences and Putative Promoters

Self-developed Python scripts utilizing nucleotide and protein BLAST from NCBI BLAST+ toolkit (version 2.3.0) and muscle [version 3.8.31 (Edgar, 2004)] were used for creating multiple sequence alignments to analyze coding sequences and putative promoter regions. Annotated coding sequences were extracted from the analyzed genomes together with 200 bp upstream fragments containing the putative promoters. Translations of the coding sequences were clustered with a similarity threshold of 90%. All identical protein sequences within a cluster were designated as a single allele. The multiple sequence alignments of different alleles and promoter sequences within each cluster were scanned for differentiating columns.

Visual Genome Comparison With BRIG

The nine genomes, published here, and the genome of N315 strain as a human reference of ST5 were compared to ED98 genome using BRIG tool (Alikhan et al., 2011). The identity thresholds were set to 98 and 95% and ED98 mobile genetic elements, prophages ϕ Av1, ϕ Av β , the pathogenicity island SaPIAv, and plasmids pAvX, pT181, and pAvY were annotated.

Whole Genome General Comparison of CH21 and ch22 Strains

Continuous similar sequence segments were identified using nucleotide BLAST. Most segments were separated by short

insertions or duplications. Similar sequence segments and their equivalence between strains were visualized using Circos [version 0.69-3 (Krzywinski et al., 2009)].

Proteomic Analysis

Exoproteome Analysis

For comparative analysis of the exoproteomes of the CH21 and ch22 strains, triplicates of cleared culture fluids were handled as described previously (Bonar et al., 2016). To assure no contamination with intracellular proteins, the supernatants were passed through a 0.22- μ m PVDF filter. Then, the proteins were precipitated with an equal volume of 20% (w/v) trichloroacetic acid in acetone and recovered by centrifugation. The pellet was washed with acetone and air dried. The samples were dissolved in lysis buffer (30 mM TrisHCl pH 8.5 containing 7 M urea, 2 M thiourea and 4% CHAPS). The protein samples were labeled with spectrally resolvable fluorescent G-dyes (NH Dye AGNOSTICS GmbH) and subjected to two-dimensional difference gel electrophoresis [2D DIGE; (Alban et al., 2003; Timms and Cramer, 2008; Minden et al., 2009)]. Isoelectrofocusing (IEF) was performed using 17-cm immobilized pH (3–10) gradient strips and Protean IEF Cell (Bio-Rad). Proteins were separated in the second dimension using an 12% acrylamide gel according to the Laemmli method (Laemmli, 1970). The gels were scanned using Typhoon Trio + (GE), and images were analyzed using Image Quant v.7.0 and DeCyder 2D software v.7.2 (GE). Protein spots were considered as differentiating if the standardized average spot volume ratio exceeds 1.5-fold at the 95% confidence level (Student's *t*-test *p*-value <0.05). Subsequently, the gels were silver stained (Shevchenko et al., 1996). The differentiating spots were excised and destained by several subsequent washes in 25% and 50% (v/v) acetonitrile (ACN) in 25 mM ammonium bicarbonate buffer (NH₄HCO₃), pH 8.0 at 37°C. The gel fragments were dehydrated in ACN, dried using a vacuum concentrator, and rehydrated using trypsin solution (10 ng/ μ l in 25 mM NH₄HCO₃, pH 8.0), and digestion was carried out overnight at 37°C. Peptides were extracted by sonication, dehydrated in ACN and dried using a vacuum concentrator. Samples were suspended in 2% (v/v) ACN in water containing 0.05% (v/v) trifluoroacetic acid (TFA) and separated using an UltiMate 3000RS LCnanoSystem (Dionex). Peptides were analyzed on a coupled MicroTOF-QII mass spectrometer (Bruker) equipped with an Apollo Source ESI-nanospray with low-flow nebulizer. Raw data were pre-processed with Data Analysis 4.0 software (Bruker, Germany) into Mascot Generic format. The SwissProt non-redundant protein database taxonomically restricted to *Firmicutes* (gram-positive bacteria) was queried with the obtained peak lists using an in-house Mascot server. Additionally, an in-house prepared database constructed based on the full genome sequences of CH21 and ch22 obtained in this study was used. Only identifications with a score value over 100 were considered relevant for further analysis. If more than a single protein was identified in particular spot, only hits scoring over 50% of the highest scoring protein were considered in further analysis.

Analysis of Intracellular Proteome

For comparative analysis of the intracellular proteomes of strains CH21 and ch22, 3 ml of overnight cultures (OD₆₀₀ ~ 12) was used in triplicate. The cells were collected by centrifugation, washed three times with 10 mM TrisHCl pH 8.0 and suspended in 1 ml of Tri Reagent (Sigma Aldrich), transferred to Lysis Matrix Tubes (MP Biomedicals) and disrupted with Precellys 24 Homogenizer (Bertin Instruments). Lysates were clarified by centrifugation, and the protein phase was isolated according to the Tri Reagent protocol. The obtained protein pellet was suspended in lysis buffer and analyzed by 2D DIGE and MS as described for exoproteins.

Surface Proteome Analysis

Proteolytic “shaving” was used to compare surface proteomes (surfacomers) of strains CH21 and ch22. All samples were analyzed in biological triplicates. Bacteria were cultured overnight in TSB at 37°C with agitation. The fresh medium was inoculated with the overnight culture at 1:100 dilution and cultured until OD₆₀₀ reached 1. The cultures were divided into two samples, experimental and control, based on the method of Solis et al. (2010) to control for contamination by cytoplasmic proteins resulting from unspecific cell lysis. Bacteria were collected by centrifugation, washed with phosphate buffered saline (PBS) and suspended in 30% sucrose in PBS. Experimental samples were incubated with trypsin [1 μ g/ μ l; “Gold, MS Grade” (Promega)] for 30 min at 37°C with gentle agitation, clarified by centrifugation and filtered through a 0.22- μ m PVDF filter. The control sample was handled as above save that trypsin was not added during the initial incubation. The control sample was treated with trypsin (1 μ g/ μ l) only after filtering (conditions as above). All samples were reduced with DTT and alkylated with idoacetamide, digested overnight with trypsin 0.2 μ g/ μ l (Biocentrum) at 37°C, supplemented with 0.5% TFA and 5% ACN (final concentration), cleaned using Pierce C18 Spin columns (Thermo Fisher Scientific) and vacuum dried. Samples were analyzed by MS as described for exoproteins, save score values over 50 were considered relevant.

RESULTS AND DISCUSSION

Overall Genomic Characteristics of Virulent and Non-virulent Strains of *S. Aureus*

In a prior study, we evaluated the virulence in a chicken embryo infection model of a number of poultry originating strains of *Staphylococcus aureus* on a background of their genetic relationships. To determine the genomic basis of phenotypic differences between the selected strains, in this study, we obtained the genomic sequences of five highly virulent strains (CH3, CH5, CH9, CH21, and CH23) and four strains characterized by low virulence (ch22, ch24, pa3, and ph2). Two genomes (CH21 and ch22) were obtained in their complete form, whereas the remaining seven genomes were obtained in the form of ordered contigs. The overall characteristics of the obtained genomes are summarized in **Supplementary Table 1**. Genome size does not correlate with virulent phenotypes since all genomes, save that of strain ch22 (considered in more detail in the next section), have

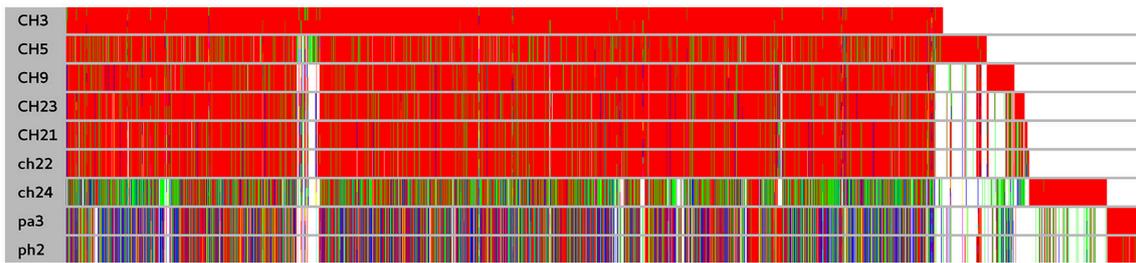


FIGURE 1 | Comparison of protein clusters among nine *S. aureus* strains. Clusters were obtained for annotated coding sequences at threshold of 90% similarity. Within a cluster identical sequences are denoted by the same color. White denotes the absence of an ortholog. Long white gaps at the end of the plot correspond to the absence of different mobile genetic elements such as plasmids. The cluster order is related to the order of coding sequence occurrence in the genome of ED98.

a similar size of ca. 2.8 Mbp. Furthermore, the overall number of genes and coding sequences does not correlate with virulent phenotypes and is comparable among the analyzed strains save for ch22. Significantly, the virulent phenotype correlates with the presence of phage ϕ AvB (46,768 bp; 62 ORFs) and pAvY plasmid (1,442 bp, 1 ORF). All VIR strains carry the above genetic elements, whereas all NVIR strains, with the exception of ch22, are devoid of both of these elements. Apart from the above mobile genetic elements, we additionally identified a number of novel prophages and pathogenicity islands and several known plasmids, neither of which differentiated the VIR and NVIR groups (**Supplementary Tables 3, 4; Supplementary Figures 1, 2**). Another clear distinction differentiating the VIR and NVIR groups becomes apparent when ORFs longer than 165 nucleotides are extracted and clustered according to protein sequence similarities above 90% (**Figure 1**). When the core genome of all nine strains is considered, only approximately one-quarter of genes fall within such defined clusters, which clearly demonstrates significant genetic differences among strains. Interestingly, the percentage of genes clustered with the above criteria among the core genomes of the VIR and NVIR groups is much higher at 72 and 42%, respectively, demonstrating the closer relatedness of strains within both groups than between groups. Such analysis also shows that the VIR group is more homogenous than the NVIR group (**Table 1**). When compared to human-origin strains of sequence type 5 (ST5), the ST5 poultry-origin strains comprise a clear evolutionary subgroup. This phylogenetic relationship based on more in depth analysis of genome-extracted 97 different loci (Nubel et al., 2008) places poultry-origin ST5 strains on a separate branch of the tree regardless to geographical origin. Interestingly, based on this analysis CH21 and ch22 strains are still genetically identical (**Figure 2**). Markedly, the whole genome analysis confirmed their high genetic resemblance by revealing only 85 single nucleotide polymorphisms (SNPs). The maximal SNP number among the whole poultry ST5 subgroup is 355, whereas the distance to the type strain N315 of human origin is higher (from 663 to 823 SNPs). Non-ST5 poultry strains analyzed in this study, all being non-virulent, are clearly more distant from the aforementioned ST5 subgroup by having from 19,384 to 21,127 SNPs. The ST692 strains pa3 and ph2, although

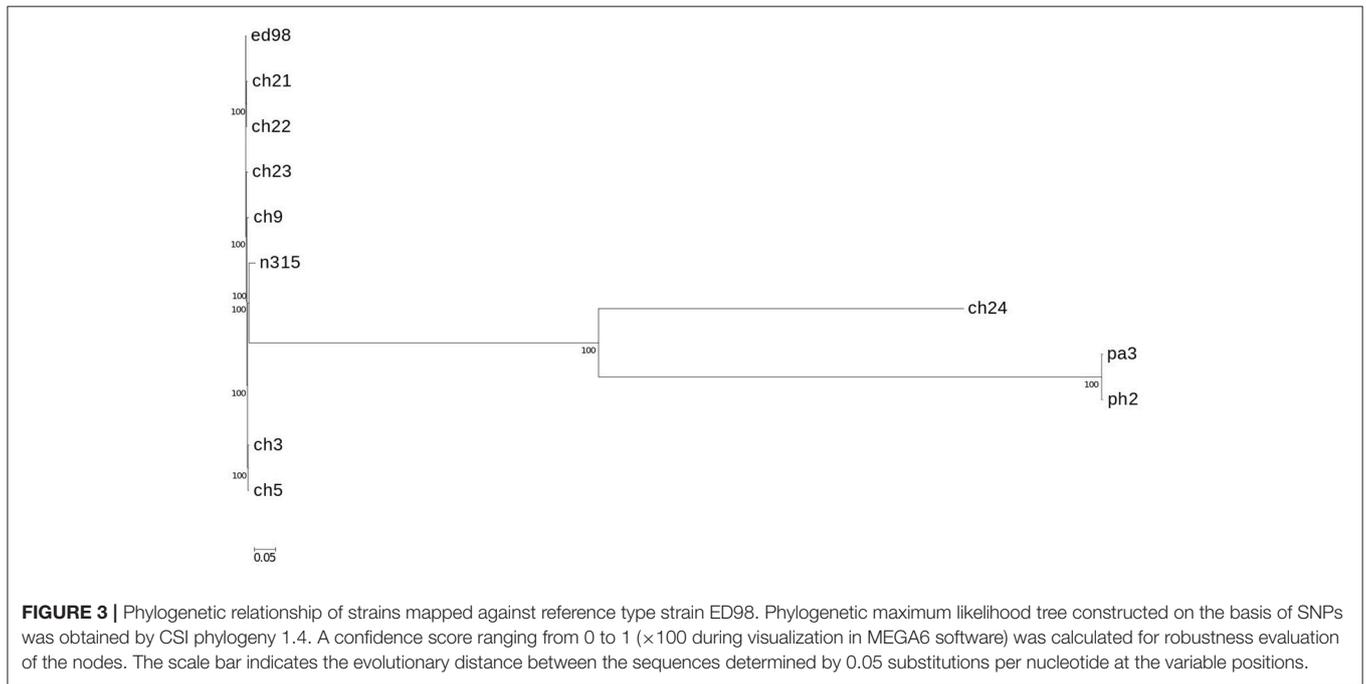
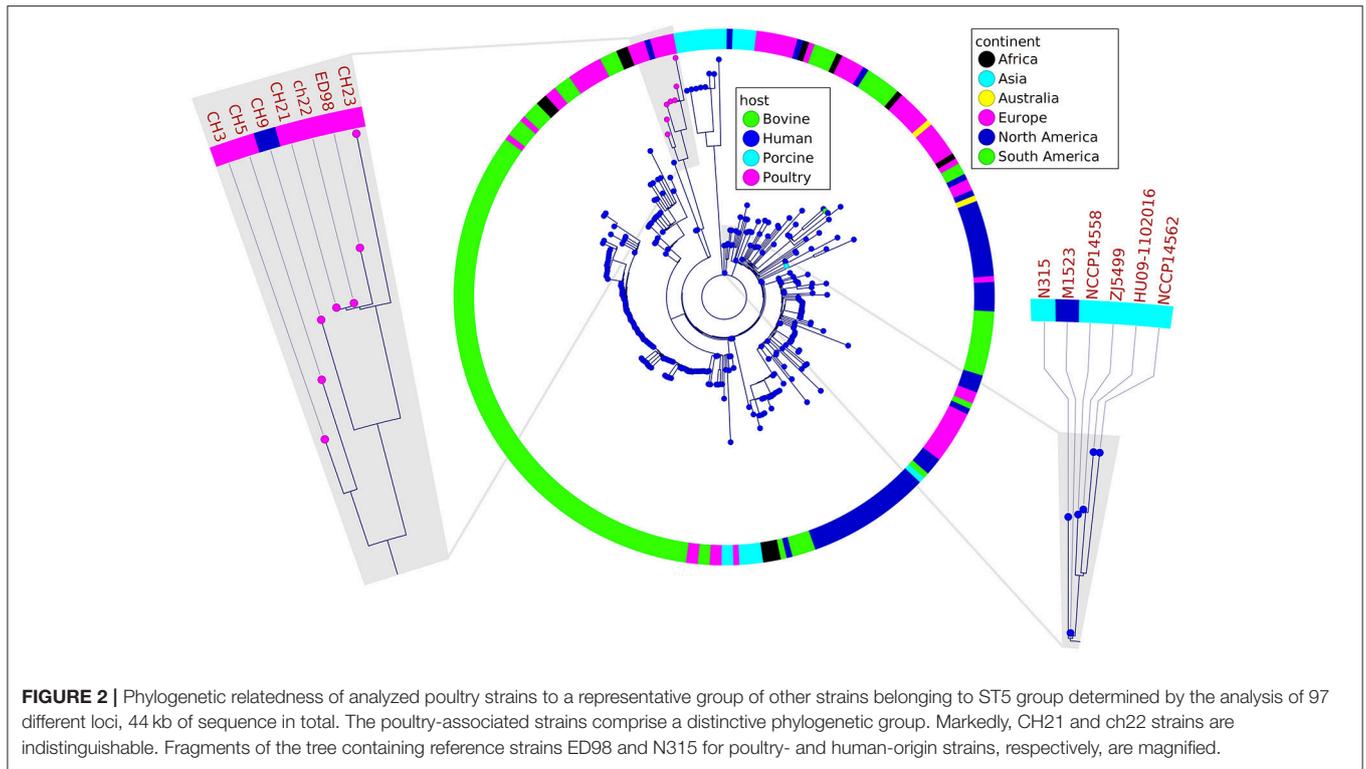
TABLE 1 | Number of clusters gathering open reading frames exhibiting over 90% protein sequence similarity in the studied group of *S. aureus* strains.

	Total number of identified clusters	Number (percent) of identical clusters
Pan-genome	3,307	1,277 (38.62)
Core-genome	2,317	586 (25.29)
VIR strains pan-genome	2,949	2,513 (85.22)
VIR strains core-genome	2,527	1,826 (72.26)
NVIR strains pan-genome	3,009	1,862 (61.88)
NVIR strains core-genome	2,471	1,050 (42.25)

isolated from different poultry hosts, respectively partridge and pheasant, are genetically highly similar to each other by displaying only 165 SNPs. The ST1 strain ch24 is highly distant from the rest of the analyzed strains by having from 19,384 to 20,875 SNPs (**Figure 3; Supplementary Table 12**). All these genetic similarities and differences are also visible when the analyzed genomes are visually compared to ED98 genome as a reference (Lowder et al., 2009). Notably, none of the non-ST5 strains contain mobile genetic element such as prophages ϕ Av1, ϕ Av β and the pathogenicity island SaPIAv, which seem to be characteristic to the ST5 poultry subgroup (**Figure 4**).

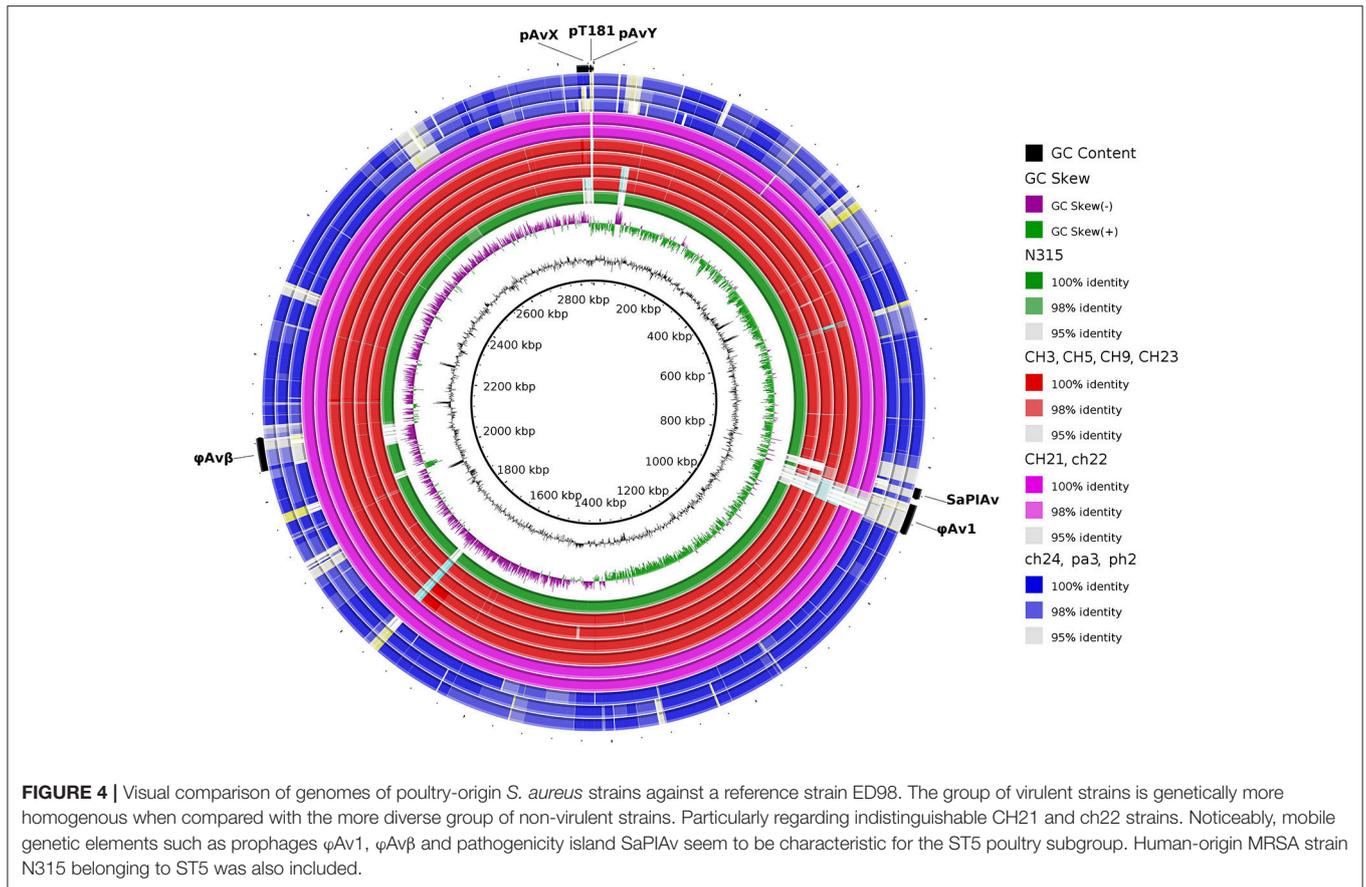
Genetic Differences Between Virulent (CH21) and Non-virulent (ch22) Strains

CH21 and ch22 show opposite phenotypes in a chicken embryo infection model (VIR vs. NVIR, respectively) but are indistinguishable by common typing techniques. We assumed that their different phenotypes may be explained by deep genome sequencing. Such an approach is generally applied to elucidate the genetic bases of the sudden acquisition of antibiotic resistance (Sabat et al., 2015). However, it was likely that the switch between the VIR and NVIR phenotypes in CH21 vs. ch22 was also due to mutations or the loss or acquisition of genetic material, which would be traceable by NGS. Therefore, we obtained the complete genomic sequences of both strains. The



first striking difference between the strains was the total length of their genomes, which was 2.8 and 3.1 Mbp, respectively, for the virulent and non-virulent strains. The additional genetic pool of over 250 kbp within the genome of ch22 arose mainly from three duplications. The longest one (CP017807 1552041..

1722216 and 1777324.. 1947499) encompassed 170 kbp and encoded 230 proteins. Two shorter ones (CP017807 1496929.. 1551606 and 1722212.. 1776889, 54 kbp; and CP017807 1978955.. 1990796 and 2244863.. 2256780, 12 kbp) encoded 70 and 16 proteins, respectively (**Supplementary Tables 5–9**).



Interestingly, the major duplications resided within the core part of the bacterial chromosome, whereas the pool of mobile genetic elements, encompassing plasmids pAvX and pAvY, two prophages and pathogenicity island SaPIAv, remained identical in both strains (Figure 5; Supplementary Table 3). Therefore, the first conclusion is that loss of the virulent phenotype by ch22 was not related to losing a substantial part of its genetic material, which could have been responsible for virulence. On the opposite, pathogenic bacteria were shown to have generally smaller genomes than their close non-virulent relatives (Georgiades and Raoult, 2011). Therefore, one can imagine that the burden of dispensable genetic material could impair the rate of growth and thus the virulence of ch22. However, the growth rate in rich TSB medium as well as in minimal M9-CAA medium were comparable between CH21 and ch22 (Supplementary Figure 6). Moreover, a possible relation between genome size and virulence is not supported by analysis of the genomes of the remaining strains, which have similar sizes within both the VIR and NVIR groups.

A detailed comparative analysis revealed genetic differences that possibly affect the virulence status of CH21 and ch22. The first notable difference between CH21 and ch22 follows the overall distinction characteristic for the analyzed VIR and NVIR strains and encompasses 10 ORFs (Table 2) with established roles in staphylococcal colonization and virulence (Foster et al., 2014).

Among others, a deletion and an internal stop codon within the repeat-containing fragments of *clfA* and *clfB* result in the truncation of protein sequences by 113 and 254 amino acid residues, respectively, in CH22. In the case of *sdrD*, eight sense mutations and a point mutation upstream of the coding sequence are different in the strains. CH21 and ch22 are indistinguishable with popular typing methods, including MLVF (Sabat et al., 2003). Significantly, MLVF is based on virulence-related genes and thus allowed us to clearly distinguish virulent (CH3, CH5, CH9, CH21, and CH23) from non-virulent (ch24, pa2, and ph2) strains (Polakowska et al., 2012). This study shows that CH21 and ch22 indeed contain different alleles of three genes (*clfA*, *clfB*, and *sdrD*) among the seven used for typing; however, the differences were only detectable using deep sequencing. This finding corroborates multiple previous reports on the involvement of these proteins in staphylococcal virulence, where clumping factors and other Sdr proteins facilitate successful colonization-mediating interactions with host proteins, such as fibrinogen, desmoglein-1 and cytokeratin 10 (O'Brien et al., 2002; Walsh et al., 2008; Wertheim et al., 2008; Askarian et al., 2016). As an additional example differentiating CH21 and ch22, the genes encoding pyruvate carboxylase (*pycA*) and exotoxin SeN (*seN*) are pseudogenized in the NVIR ch22 strain (Table 2). Interestingly, *seN* is totally absent in all other NVIR strains. PycA catalyzes the ATP-dependent carboxylation of

TABLE 2 | Differences in genes between VIR and NVIR strains exemplified by differences identified in the VIR CH21 and NVIR ch22 strains.

No.	Cl.*	Product	Acronym	CH21 loci	ch22 loci	Protein lengths		Differences
1.	188	Hypothetical protein (associated with type VII secretion system operon)		BJL64_01345 BJL64_01350	BJL65_01345 -----	166 161	161 --	Merge between two coding sequences of CH21.
2.	438	Serine-aspartate repeat protein D	SdrD	BJL64_02690	BJL65_02680	1,385	1,385	Eight sense mutations, one point mutation upstream of CDS.
3.	670	Clumping factor A	CifA	BJL64_03875	BJL65_03870	875	762	113 aa deletion within the repeat-containing fragment.
4.	891	Hypothetical protein (associated with bacteriocin operon)		BJL64_05095	[BJL65_05090]	654	--	Internal stop codon in ch22.
5.	938	Hypothetical protein (lipoprotein)		BJL64_05335	[BJL65_05325]	69	--	Internal stop codon in ch22.
6.	975	Pyruvate carboxylase	PycA	BJL64_05525	[BJL65_05515]	1,150	--	Internal stop codon in ch22.
7.	1017	Hypothetical protein		BJL64_05735	[BJL65_05725]	55	--	Internal stop codon in ch22.
8.	1702	Exotoxin	SeN	BJL64_09410	[BJL65_10575]	251	--	Internal stop codon in ch22.
9.	1767	Cell division protein	FtsK	BJL64_08675	BJL65_09840	453	453	Five sense mutation and 13 point mutations upstream of CDS.
10.	2529	Clumping factor B	CifB	BJL64_13955	[BJL65_15235]	865	--	Internal stop codon in ch22.

*Cluster's number, for detailed protein sequences and alignment see "<https://mol058.mol.uj.edu.pl/extra/clusters.htm>."

TABLE 3 | Identified differences in genes exclusively found between the VIR CH21 and NVIR ch22 strains.

No.	Cl.*	Product	Acronym	CH21 loci	ch22 loci	Protein lengths		Differences
1.	89	Hypothetical protein		BJL64_00820	BJL65_00820	199	199	Seven sense mutations in a highly variable region.
2.	385	Zinc metalloprotease	FtsH	BJL64_02360	BJL65_02355	710	697	One sense mutation. One SNP resulting in a frameshift at the end of the gene. 36 terminal aa changed to 23 different aa.
3.	475	NAD(P)H-dependent oxidoreductase	WrbA	BJL64_02885	BJL65_02880	178	178	One sense mutation.
4.	1221	Choline transporter	BetT	BJL64_06800	BJL65_06835	548	548	One sense mutation.
5.	1368	Hypothetical protein (lipoprotein)		BJL64_07580	BJL65_07610	230	308	A highly variable region shortened by 78 aa in CH21.
6.	1573	Gamma-aminobutyrate permease	GabP	BJL64_08625	[BJL65_09790]	453	--	Internal stop codon in ch22.
7.	1700	Hypothetical protein		BJL64_09400	BJL65_10565	258	258	One sense mutation.
8.	1766	AAA family ATPase (FtsK operon associated)	VirB4	[BJL64_08695]	BJL65_09860	--	832	Internal stop codon in CH21.
9.	2127	Malonate transporter	YfdV	BJL64_11910	BJL65_13180	302	302	Two sense mutations.
10.	2185	Conjugal transfer protein (FtsK operon associated)	TpcC	BJL64_00290 BJL64_08710	BJL65_00290 BJL65_09875	353 358	353 358	Four sense mutations. A variable 17 aa region containing 15 sense mutations.
11.	2356	Peptide ABC transporter ATP-binding protein	NikE	[BJL64_13030]	BJL65_14305	--	249	Internal stop codon in CH21.
12.	2902	Hypothetical protein		BJL64_11025 -----	BJL65_12290 BJL65_12295	30 --	30 30	Identical protein sequences. Tandem duplication. A copy with one point mutation.

*Cluster's number, for detailed protein sequences and alignment see "<https://mol058.mol.uj.edu.pl/extra/clusters.htm>."

TABLE 4 | Identified differences in putative promoter regions between the VIR CH21 and NVIR ch22 strains.

No.	Downstream coding sequence	Acronym	CH21 loci	ch22 loci	Protein lengths		Differences
1.	Ring-cleaving dioxygenase	GloA	BJL64_01535	BJL65_01530	308	308	One point mutation upstream of CDS.
2.	Hypothetical protein		BJL64_01915	BJL65_01910	227	227	Completely different sequences from 113 bp upstream of CDS.
3.	Serine-aspartate repeat protein D	SdrD	BJL64_02690	BJL65_02680	1,385	1,385	One point mutation upstream of CDS.
4.	Hypothetical protein		BJL64_04510	BJL65_04505	37	37	Completely different sequences from 145 bp upstream of CDS.
5.	Hypothetical protein		BJL64_05100	BJL65_05095	106	106	Completely different sequences from 54 bp upstream of CDS. In ch22, the preceding CDS is pseudogenized.
6.	Fibrinogen-binding protein	FnBP	BJL64_05765	BJL65_05755	116	116	One single nucleotide deletion upstream of CDS.
7.	Hypothetical protein		BJL64_07020	BJL65_07055	55	55	Five point mutations upstream of CDS.
8.	DNA-binding response regulator	OmpR	BJL64_08605	BJL65_09770	234	234	Three point mutations upstream of CDS.
9.	Hypothetical protein		BJL64_08650	BJL65_09815	120	120	Two point mutations upstream of CDS.
10.	Hypothetical protein (FtsK operon associated)		BJL64_08655	BJL65_09820	197	197	Two point mutations upstream of CDS.
11.	Hypothetical protein (FtsK operon associated)		BJL64_08670	BJL65_09835	77	77	19 point mutations and one single nucleotide deletion upstream of CDS.
12.	Cell division protein	FtsK	BJL64_08675	BJL65_09840	453	453	13 point mutations upstream of CDS.
13.	Hypothetical protein (FtsK operon associated)		BJL64_08680	BJL65_09845	110	110	One point mutation upstream of CDS.
14.	Exotoxin	SeO	BJL64_09435	BJL65_10600	254	254	Four one-nucleotide deletions upstream of CDS.
15.	Cysteine protease (staphopain A)	ScpA	BJL64_10140	BJL65_11405	388	388	Completely different sequences starting from the first bp upstream of CDS because of translocation.
16.	Hypothetical protein		BJL64_11025	BJL65_12290 BJL65_12295	30 30	30	Gene duplication. four point mutations in the region up to 100 bp upstream of CDS and substantial differences in the region from 101 to 200 bp upstream of CDS.
17.	Thiaminase II	TenA	BJL64_11050	BJL65_12320	229	229	One point mutation upstream of CDS.
18.	Hypothetical protein		BJL64_11905	BJL65_13175	43	43	One point mutation upstream of CDS.
19.	Hypothetical protein		BJL64_13105	BJL65_14385	140	140	Deletion of 46 bp right upstream of CDS in CH21.

the respective groups of differentiating spots (**Table 5; Figure 6; Supplementary Table 13**).

Although intracellular proteins do not directly interact with the host, they significantly contribute toward maintaining the virulent phenotype by maintaining metabolism in stress conditions. Proteins characterized by more abundant expression in virulent strain CH21 compared to ch22 included diapolycopene oxygenase (CrtP), which is involved in the biosynthesis of staphyloxanthin. Interestingly, however both strains exhibited the same level of pigmentation (**Supplementary Figure 5**). This golden pigment has been linked with staphylococcal virulence as it shields the microbe from oxidation-based clearance, which is the innate host immune response to infection (Clauditz et al., 2006). Interestingly,

it has been shown that strains with gene knock-outs in the staphyloxanthin biosynthesis pathway exhibit attenuated virulence (Liu et al., 2005). Nevertheless, staphyloxanthin overexpression does not directly correlate with virulence. It has been shown that strains with elevated pigmentation (associated with mutations in genes responsible for oxidative phosphorylation and purine biosynthesis) were less virulent in a murine abscess model of infection. Some strains were even unable to exhibit long-term colonization (Lan et al., 2010). Furthermore, supplementation of *S. argenteus* with a gene cluster responsible for staphyloxanthin production led to increased susceptibility to the host defense peptides LL-37 and hNP-1 *in vitro* and reduced virulence in an experimental rabbit endocarditis model (Xiong et al., 2015). Therefore, CrtP

TABLE 5 | List of proteins differentially expressed by VIR CH21 and NVIR ch22 strains as identified by proteomics.

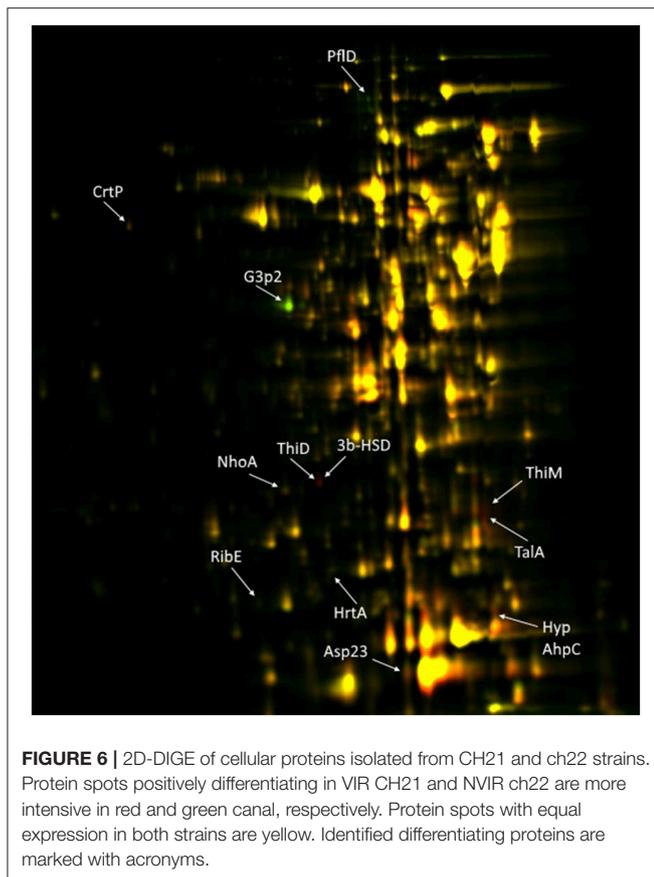
Location	Locustag CH21	Locustag ch22	Protein	Acronym	Cluster of orthologous groups (functional category)	Elevated in/shaved from strain
in	BJL64_13565	BJL65_14845	Diapolycopene oxygenase	CrtP	1233 (Q)	CH21
in	BJL64_11045	BJL65_12315	Bifunctionalhydroxymethylpyrimidine kinase/phosphomethylpyrimidine kinase	ThiD/J	0351 (H)	CH21
in	BJL64_01840	BJL65_01835	NAD(P)-dependent oxidoreductase	3 β -HSD	0702 (R)	CH21
in	BJL64_14305	BJL65_15585	Arylamine N-acetyltransferase	NhoA	2162 (Q)	CH21
in	BJL64_11040	BJL65_12310	Hydroxyethylthiazole kinase	ThiM	2145 (H)	CH21
in	BJL64_09170	BJL65_10335	Translaldolase	TalA	0176 (G)	CH21
in	BJL64_09115	BJL65_10280	Riboflavin synthase subunit alpha	RibE	0307 (H)	CH21
in	BJL64_00690	BJL65_00690	Hypothetical protein	Hyp	n/a	CH21
in	BJL64_01760	BJL65_01755	Peroxiredoxin	AhpC	0450 (V)	CH21
in	BJL64_11525	BJL65_12795	Alkalineshock protein	Asp23	1302 (S)	CH21
in	BJL64_00965	BJL65_00965	Formate C-acetyltransferase	PflD	1882 (C)	ch22
in	BJL64_08575	BJL65_09740	Type I glyceraldehyde-3-phosphate dehydrogenase	G3p2	0057 (G)	ch22
in	BJL64_12445	BJL65_13715	Hemin ABC transporter ATP-binding protein	HrtA	1136 (M)	ch22
out	BJL64_09350	BJL65_10515	DUF4888 domain-containing protein	DUF4888	n/a	CH21
out	BJL64_02695	BJL65_02685	MSCRAMM family adhesion SdrE	SdrE	4932 (S)	ch22
out	BJL64_07905	BJL65_09090	Superoxide dismutase	SodM1	0605 (P)	ch22
out	BJL64_01760	BJL65_01755	Peroxiredoxin	AhpC	0450 (V)	ch22
out	BJL64_03485	BJL65_03480	Glycerol phosphate lipoteichoic acid synthase	LtaS	1368 (M)	ch22
out	BJL64_01445	BJL65_01440	Lipase2	Lip2	1075 (I)	ch22
out	BJL64_11645	BJL65_12915	Toxin		n/a	ch22
out	BJL64_14165	BJL65_15445	Lipase1	Lip1	1075 (I)	ch22
surf	BJL64_02580	BJL65_02570	50S ribosomal protein L7/L12	RL7/12	0222 (J)	CH21
surf	BJL64_12760	BJL65_14035	Immunoglobulin-binding protein Sbi	Sbi	n/a	CH21
surf	BJL64_14035	BJL65_15315	N-acetylmuramoyl-L-alanine amidase	Y2979	1705 (MN)	CH21
surf	BJL64_03805	BJL65_03800	Phosphopyruvate hydratase	Eno	0148 (G)	ch22
surf	BJL64_07465	BJL65_08690	DNA-binding protein	HU	0776 (L)	ch22
surf	BJL64_11780	BJL65_13050	30S ribosomal protein S5	RS5	0098 (J)	ch22
surf	BJL64_13785	BJL65_15065	Fructose bisphosphate aldolase	Alf2	3588 (G)	ch22
surf	BJL64_14045	BJL65_15325	Adhesin (surface protein F)	SasF	n/a	ch22

overexpression in CH21 may not necessarily relate to the virulence of this strain.

Peroxiredoxin, another enzyme protecting these bacteria against oxidative host attack by scavenging hydrogen peroxide, was also upregulated in CH21 compared to ch22. This finding demonstrates the general importance of antioxidative mechanisms in staphylococcal virulence. The level of alkaline shock protein (Asp23), which was linked with cell homeostasis and protection of the cell envelope in non-growing cells (Müller et al., 2014), was also elevated in CH21 compared to ch22. In addition,

enzymes implicated in the synthesis of vitamins [thiamine (hydroxymethylpyrimidine/phosphomethylpyrimidine kinase and hydroxyethylthiazole kinase), and riboflavin (riboflavin synthase subunit alpha)] and an enzyme involved in the metabolism of bile acid [NAD(P)-dependent oxidoreductase] were also elevated in the virulent CH21 strain compared to ch22; however, the significance of this fact in virulence remains elusive.

Proteins characterized by higher expression in non-virulent strain ch22 compared to virulent strain CH21 included type I glyceraldehyde-3-phosphate dehydrogenase 2 (G3P2), an isoenzyme of G3PD-1 involved in glycolysis, and formate



C-acetyltransferase, which is implicated in anaerobic glucose metabolism. Hemin ABC transporter ATP-binding protein (HrtA) was also elevated in ch22 (**Figure 6**). HrtA is responsible for coping with heme stress and was already demonstrated to negatively correlate with staphylococcal virulence (Hammer and Skaar, 2011). The transporter prevents heme-mediated toxicity. It has been demonstrated that an inability to cope with heme stress paradoxically yields a hypervirulent phenotype (Hammer and Skaar, 2011). In *hrtA* mutants, such a phenotype was associated with the loss of membrane integrity and increased secretion of immunomodulatory proteins (Attia et al., 2010). The above findings point to the likely significance of the inability to produce HrtA by CH21 in maintaining the strain virulence.

In conclusion, although the expression of only 13 intracellular proteins is significantly different between the VIR and NVIR strains, the pattern of expression is strikingly consistent. VIR strain CH21 over-expresses proteins involved in coping with oxidative stress, which characterizes the host response against invading pathogens. In contrast, NVIR strain ch22 overproduces enzymes adapting the cell to anaerobic conditions, which characterize its niches of commensal coexistence with the host.

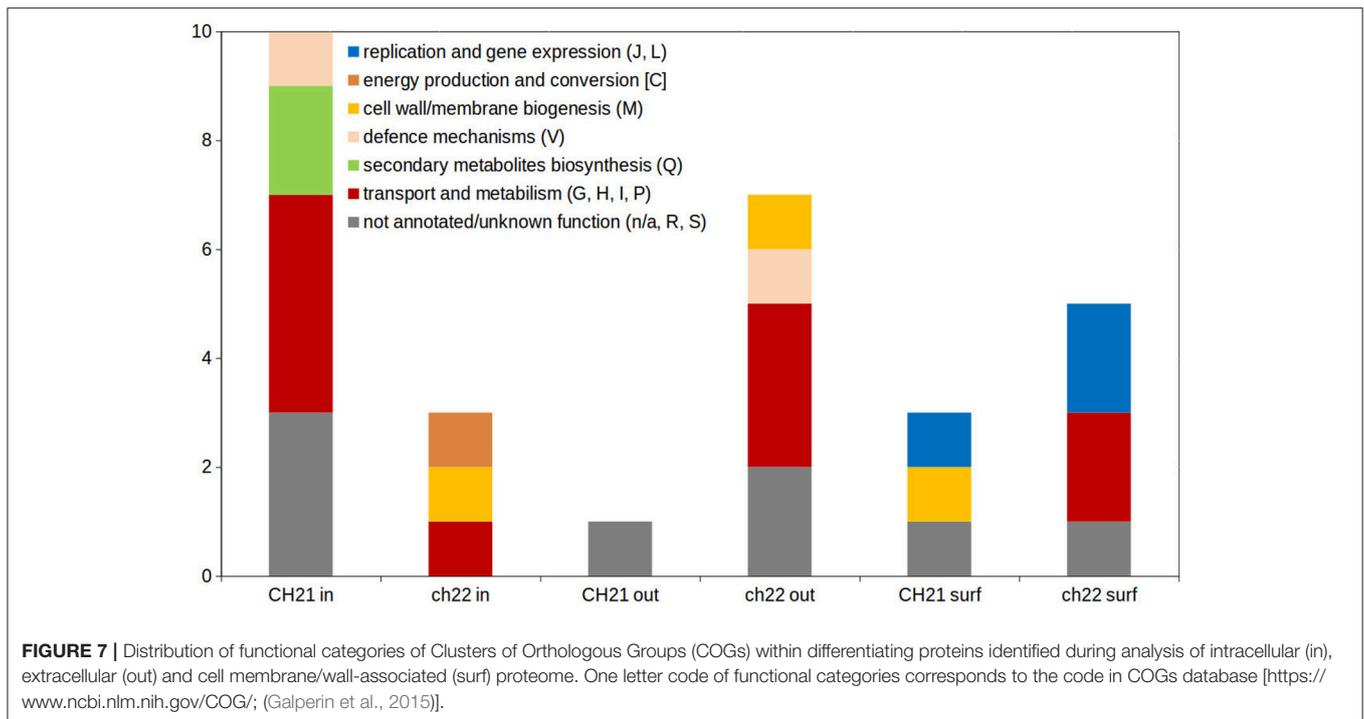
The exoproteomes of strains CH21 and ch22 are dominated by staphopain C, a plasmid-encoded cysteine protease

overproduced by a range of poultry-derived *S. aureus* strains that was previously demonstrated to be a factor unrelated to virulence (Bonar et al., 2016). Only three differentiating spots were identified as overexpressed within the exoproteome of virulent strain CH21 compared to ch22, and all three spots contained DUF4888 domain-containing protein (**Table 5; Supplementary Table 13**). This protein of 193 residues has no known function (Marchler-Bauer et al., 2017). Homologues are primarily found in various species of staphylococci, suggesting a species-specific role, the elucidation of which is of significant interest in the light of this study. Interestingly, the exoproteome of the virulent strain CH21 is not characterized by the overexpression of any known virulence factors compared to nonvirulent strain ch22. Even alpha-hemolysin, which was previously consistently found to be overexpressed in the exoproteomes of virulent strains isolated from poultry, is not overproduced by CH21 compared to ch22 (Bonar et al., 2016).

Within the exoproteome of NVIR strain ch22, seven proteins were found to be overexpressed compared to CH21 (**Table 5; Supplementary Table 13**). The most pronounced differences in expression were characterized by lipase 1 (Lip1, found in 15 differentiating protein spots) and lipoteichoic acid synthase (LtaS, found in 10 differentiating spots). Further, peroxiredoxin (AhpC) was identified in three differentiating spots, whereas superoxide dismutase 1 (SodM1) and lipase 2 (Lip2) were found in two differentiating spots. Additional proteins were identified as single differentiating spots.

The higher complexity of the exoproteome of nonvirulent strain ch22 compared to virulent strain CH21 corroborates the results of our previous study, which documented a general regularity in the more complexed exoproteomes of NVIR strains relative to VIR strains (Bonar et al., 2016). Moreover, Clusters of Orthologous Groups (COGs) analysis confirmed that proteins differentiating both exo- and intracellular proteomes of CH21 and ch22 were assigned to different functional categories (**Figure 7; Supplementary Figure 4**).

Gel proteomics performs poorly in comparing cell membrane / cell wall associated proteomes. Therefore, to compare these subsets of the proteomes of CH21 and ch22, we used the cell “shaving” approach (Solis et al., 2010) coupled to direct, LC-MS based identification. This approach enabled direct comparisons, but unlike semiquantitative gel proteomics, the direct approach is qualitative only and suffers from large variability among pools identified in different biological repeats. As such, only identifications present in two or more samples were considered further (**Table 5; Supplementary Table 13**). Proteins identified exclusively on the surface of virulent strain CH21 include immunoglobulin-binding protein (Sbi) and N-acetylmuramoyl-L-alanine amidase (Y2979; BJL64_14035). Sbi and staphylococcal protein A (SpA) are IgG binders with a demonstrated role in staphylococcal virulence (Gonzalez et al., 2015). While SpA binds Fc γ of IgGs, Sbi has two Ig-binding domains and two domains that bind to complement component C3 (Zhao et al., 2016). A prior study demonstrated the pseudogenization of *spa*, a gene encoding a major *S. aureus* IgG-binding protein, as a specific adaptation to a poultry host since SpA is unable to bind



chicken immunoglobulins (Lowder et al., 2009). It is tempting to speculate that Sbi may replace SpA in poultry strains. Sbi has a broader ability to bind mammalian IgG than SpA, but its ability to bind avian Igs remains to be tested (Atkins et al., 2008). Amidase Y2979 is consistently found in the exoproteomes of different virulent *S. aureus* strains isolated from poultry (Bonar et al., 2016), but its role in staphylococcal virulence remains unknown.

SasF adhesin and DNA-binding protein (HU) are distinguished among proteins identified exclusively on the surface of non-virulent strain ch22. SasF was initially identified as a cell wall attached adhesion protein (Roche et al., 2003), but later, it was demonstrated that it provides resistance to unsaturated free fatty acids, such as linoleic acid (Kenny et al., 2009). It is thus likely that SasF has a protective role, compensating the increased lipase content in the exoproteome of ch22. High lipase would result in high production of free fatty acids, which in turn, could adversely affect the integrity of the bacterial cell membrane (Cadieux et al., 2014). SasF would counteract such adverse effects. HU protein is one of the major nucleoid-associated proteins involved in DNA bending and thus the determination of prokaryotic chromosome structure (Kim et al., 2014a). It also acts as a transcriptional regulator of genes, responding to anaerobiosis, acid stress, high osmolarity and SOS induction (Oberto et al., 2009). Abundant intracellular proteins are often concomitantly present at the surface of these bacteria, and they usually have moonlighting functions. Extracellular DNA is an important component of bacterial biofilms, and DNA-binding proteins other than HU are associated with biofilm formation (Joo and Otto, 2012), suggesting that HU may have a similar function.

Advantageous Influence of “Individualized” Genomics in Proteomic Identifications

The number of studies combining genomic and proteomic approaches is relatively limited, but the complexity of host-pathogen interactions calls for such a holistic approach. By applying the advantage of the growing availability of NGS, this study, for the first time, based proteomic identifications on “individualized” information retrieved from the genomes of the studied strains. More importantly, however, such an approach enabled the correlation of genomic and proteomic data, providing rational explanations for the varied expression of specific proteins differentiating the virulent and nonvirulent strains. Bifunctional hydroxymethylpyrimidine kinase/phosphomethylpyrimidine kinase (BJL64_11045) and hydroxyethylthiazole kinase (BJL64_11040), proteins upregulated in CH21 compared to ch22, are encoded in a single operon (Müller et al., 2009) together with the thiaminase gene (BJL64_11050, TenA). Genomics demonstrates a single nucleotide polymorphism in the putative promoter region of the operon in which a polymorphism may be responsible for differential expression. Even more interestingly, in the case of superoxide dismutase (BJL65_09090) and DNA-binding protein (BJL65_08690; HU), which were upregulated in ch22 compared to CH21, additional copies (BJL65_07935 and B JL65_07500, respectively) of the encoding genes were found in 170 and 54 kbp duplications, respectively, explaining the observed increased expression. Nevertheless, of the 28 proteins differentially expressed in CH21 and ch22, differences in the expression of only the above four proteins could have been explained on genetic basis. This result demonstrates that global expression regulators influence the proteome more significantly

than alterations in particular promoter regions. Superoxide dismutases (SODs) give bacteria a defense mechanism against professional phagocytes. Mn-dependent SOD (SodM1) is specific for *S. aureus* and not found in coagulase-negative staphylococci, a group of staphylococci that is generally less pathogenic than the former species (Wright Valderas et al., 2002). SodM1 is generally a cytosolic protein, however extracellular SodM1 was previously identified and implicated in biofilm formation (Atshan et al., 2015). Although CH21 and ch22 are weak biofilm producers, the latter indeed forms a slightly higher biofilm when tested *in vitro* than the former (**Supplementary Figure 3**). In the above context, it is interesting to note that another primarily intracellular DNA-binding protein (HU), identified in the exoproteome of ch22, may be involved in biofilm formation. Extracellular DNA is an important component of bacterial biofilms, and DNA-binding proteins have been associated with biofilm formation (Joo and Otto, 2012). The importance of such moonlighting proteins is shown by the fact that specific monoclonal antibodies are able to disrupt an established biofilm and restore the antibiotic sensitivity of released bacteria (Estellés et al., 2016). Nonetheless, it remains to be determined whether the overexpression of HU by ch22 truly influences biofilm formation and virulence.

In conclusion, staphylococcal virulence has been investigated using genetic, proteomic, biochemical and molecular approaches (Hecker et al., 2018). It is currently well-established that the virulent phenotype relies on multiple factors that are subject to fine-tuned regulation and act in a highly orchestrated manner (Thomer et al., 2016). Nevertheless, the corresponding interconnections and guiding principles remain elusive. Here, we demonstrated that the level of virulence of wild-type *S. aureus* strains may be significantly influenced by minor differences in their genetic material. Performed proteomics indicate that coping with oxidative stress is crucial for virulent strain CH21, whereas basic metabolism and nutrient acquisition are fine-tuned in non-virulent ch22. However, we are also aware that the study was performed on a limited number of strains and changes in proteomes were assessed using *in vitro* bacterial cultures

which not optimally reflects *in vivo* conditions. Nevertheless, we pointed mutations in wild-type *S. aureus* which have been previously showed to be linked with staphylococcal virulence using recombinant strains (Benton et al., 2004), which strengthen accuracy of our findings. Other results are candidates for further detailed studies.

AUTHOR CONTRIBUTIONS

EB, MB, and BW designed the study. EB, MB, MH, UJ, SK-K, VA, and AS performed the experiments. EB, MB, SK-K, MG, GD, JM, AS, AF, and BW analyzed and interpreted data. EB, MB, GD, and BW wrote the manuscript. All authors revised the manuscript and agreed to be accountable for all aspects of the work herein.

FUNDING

This research was supported by funds granted by the National Science Centre (NCN, Poland) on the basis of decision no. DEC-2012/07/D/NZ2/04282 (to BW). Proteomics studies were carried out with equipment purchased through European Union structural funds, grant POIG.02.01.00-12-167/08 (Malopolska Centre of Biotechnology).

ACKNOWLEDGMENTS

We are thankful to Iwona Wojcik for help in proteomics. The Faculty of Biochemistry, Biophysics and Biotechnology of Jagiellonian University is a partner of the Leading National Research Center (KNOW), which is supported by the Ministry of Science and Higher Education, Warsaw, Poland.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcimb.2018.00313/full#supplementary-material>

REFERENCES

- Aanensen, D. M., Feil, E. J., Holden, M. T., Dordel, J., Yeats, C. A., Fedosejev, A., et al. (2016). Whole-genome sequencing for routine pathogen surveillance in public health: a population snapshot of invasive *Staphylococcus aureus* in Europe. *MBio* 7:e00444-16. doi: 10.1128/mBio.00444-16
- Alban, A., David, S. O., Bjorksten, L., Andersson, C., Sloge, E., Lewis, S., et al. (2003). A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. *Proteomics* 3, 36–44. doi: 10.1002/pmic.200390006
- Alikhan, N. F., Petty, N. K., Ben Zakour, N. L., and Beatson, S. A. (2011). BLAST ring image generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12:402. doi: 10.1186/1471-2164-12-402
- Askarian, F., Ajayi, C., Hanssen, A. M., van Sorge, N. M., Pettersen, I., Diep, D. B., et al. (2016). The interaction between *Staphylococcus aureus* SdrD and desmoglein 1 is important for adhesion to host cells. *Sci. Rep.* 6:22134. doi: 10.1038/srep22134
- Atkins, K. L., Burman, J. D., Chamberlain, E. S., Cooper, J. E., Poutrel, B., Bagby, S., et al. (2008). *S. aureus* IgG-binding proteins SpA and Sbi: host specificity and mechanisms of immune complex formation. *Mol. Immunol.* 45, 1600–1611. doi: 10.1016/j.molimm.2007.10.021
- Atshan, S. S., Shamsudin, M. N., Sekawi, Z., Thian Lung, L. T., Barantalab, F., Liew, Y. K., et al. (2015). Comparative proteomic analysis of extracellular proteins expressed by various clonal types of *Staphylococcus aureus* and during planktonic growth and biofilm development. *Front. Microbiol.* 6:524. doi: 10.3389/fmicb.2015.00524
- Attia, A. S., Benson, M. A., Stauff, D. L., Torres, V. J., and Skaar, E. P. (2010). Membrane damage elicits an immunomodulatory program in *Staphylococcus aureus*. *PLoS Pathog.* 6:e1000802. doi: 10.1371/journal.ppat.1000802
- Benton, B. M., Zhang, J. P., Bond, S., Pope, C., Christian, T., Lee, L., et al. (2004). Large-scale identification of genes required for full virulence of *Staphylococcus aureus* large-scale identification of genes required for full virulence of *Staphylococcus aureus*. *J. Bacteriol.* 186, 8478–8489. doi: 10.1128/JB.186.24.8478-8489.2004
- Bonar, E., Wojcik, I., Jankowska, U., Kedracka-Krok, S., Bukowski, M., Polakowska, K., et al. (2016). Identification of secreted exoproteome fingerprints of highly-virulent and non-virulent *Staphylococcus aureus* strains. *Front. Cell. Infect. Microbiol.* 6:51. doi: 10.3389/fcimb.2016.00051

- Bonar, E., Wójcik, I., and Wladyka, B. (2015). Proteomics in studies of *Staphylococcus aureus* virulence. *Acta Biochim. Pol.* 62, 367–381. doi: 10.18388/abp.2015_1083
- Bouchiat, C., Moreau, K., Devillard, S., Rasigade, J. P., Mosnier, A., Geissmann, T., et al. (2015). *Staphylococcus aureus* infective endocarditis versus bacteremia strains: subtle genetic differences at stake. *Infect. Genet. Evol.* 36, 524–530. doi: 10.1016/j.meegid.2015.08.029
- Bukowski, M., Wladyka, B., and Dubin, G. (2010). Exfoliative toxins of *Staphylococcus aureus*. *Toxins* 2, 1148–1165. doi: 10.3390/toxins2051148
- Cadieux, B., Vijayakumaran, V., Bernards, M. A., McGavin, M. J., and Heinrichs, D. E. (2014). Role of lipase from community-associated methicillin-resistant *Staphylococcus aureus* strain USA300 in hydrolyzing triglycerides into growth-inhibitory free fatty acids. *J. Bacteriol.* 196, 4044–4056. doi: 10.1128/JB.02044-14
- Calderwood, M. S., Desjardins, C. A., Sakoulas, G., Nicol, R., Dubois, A., Delaney, M. L., et al. (2014). Staphylococcal enterotoxin P predicts bacteremia in hospitalized patients colonized with methicillin-resistant *Staphylococcus aureus*. *J. Infect. Dis.* 209, 571–577. doi: 10.1093/infdis/jit501
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. doi: 10.1186/1471-2105-10-421
- Clauditz, A., Resch, A., Wieland, K. P., Peschel, A., and Götz, F. (2006). Staphyloxanthin plays a role in the fitness of *Staphylococcus aureus* and its ability to cope with oxidative stress. *Infect. Immun.* 74, 4950–4953. doi: 10.1128/IAI.00204-06
- Diep, B. A., Stone, G. G., Basuino, L., Graber, C. J., Miller, A., des Etages, S., et al. (2008). The arginine catabolic mobile element and staphylococcal chromosomal cassette *mec* linkage: convergence of virulence and resistance in the usa300 clone of methicillin-resistant *Staphylococcus aureus*. *J. Infect. Dis.* 197, 1523–1530. doi: 10.1086/587907
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Estellés, A., Woischnig, A. K., Liu, K., Stephenson, R., Lomongsod, E., Nguyen, D., et al. (2016). A high-affinity native human antibody disrupts biofilm from *Staphylococcus aureus* bacteria and potentiates antibiotic efficacy in a mouse implant infection model. *Antimicrob. Agents Chemother.* 60, 2292–2301. doi: 10.1128/AAC.02588-15
- 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
- Galperin, M. Y., Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2015). Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res.* 43, D261–D269. doi: 10.1093/nar/gku1223
- Gatermann, S., John, J., and Marre, R. (1989). *Staphylococcus saprophyticus* urease: characterization and contribution to uropathogenicity in unobstructed urinary tract infection of rats. *Infect. Immun.* 57, 110–116.
- Gatermann, S., and Marre, R. (1989). Cloning and expression of *Staphylococcus saprophyticus* urease gene sequences in *Staphylococcus carnosus* and contribution of the enzyme to virulence. *Infect. Immun.* 57, 2998–3002.
- Georgiades, K., and Raoult, D. (2011). Genomes of the most dangerous epidemic bacteria have a virulence repertoire characterized by fewer genes but more toxin-antitoxin modules. *PLoS ONE* 6:e17962. doi: 10.1371/journal.pone.0017962
- Gill, S. R., McIntyre, L. M., Nelson, C. L., Remortel, B., Rude, T., Reller, L. B., et al. (2011). Potential associations between severity of infection and the presence of virulence-associated genes in clinical strains of *Staphylococcus aureus*. *PLoS ONE* 6:e18673. doi: 10.1371/journal.pone.0018673
- Gonzalez, C. D., Ledo, C., Gaiá, C., Garófalo, A., and Gómez, M. I. (2015). The Sbi protein contributes to *Staphylococcus aureus* inflammatory response during systemic infection. *PLoS ONE* 10:e0131879. doi: 10.1371/journal.pone.0131879
- Hammer, N. D., and Skaar, E. P. (2011). Molecular mechanisms of *Staphylococcus aureus* iron acquisition. *Annu. Rev. Microbiol.* 65, 129–147. doi: 10.1146/annurev-micro-090110-102851
- Hecker, M., Mäder, U., and Völker, U. (2018). From the genome sequence via the proteome to cell physiology – pathoproteomics and pathophysiology of *Staphylococcus aureus*. *Int. J. Med. Microbiol.* 308, 545–557. doi: 10.1016/j.ijmm.2018.01.002
- Hiron, A., Posteraro, B., Carrière, M., Remy, L., Delporte, C., La Sorda, M., et al. (2010). A nickel ABC-transporter of *Staphylococcus aureus* is involved in urinary tract infection. *Mol. Microbiol.* 77, 1246–1260. doi: 10.1111/j.1365-2958.2010.07287.x
- Jolley, K. A., and Maiden, M. C. (2010). BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* 11:595. doi: 10.1186/1471-2105-11-595
- Joo, H. S., and Otto, M. (2012). Molecular basis of *in vivo* biofilm formation by bacterial pathogens. *Chem. Biol.* 19, 1503–1513. doi: 10.1016/j.chembiol.2012.10.022
- Kaas, R. S., Leekitcharoenphon, P., Aarestrup, F. M., and Lund, O. (2014). Solving the problem of comparing whole bacterial genomes across different sequencing platforms. *PLoS ONE* 9:e104984. doi: 10.1371/journal.pone.0104984
- Kenny, J. G., Ward, D., Josefsson, E., Jonsson, I. M., Hinds, J., Rees, H. H., et al. (2009). The *Staphylococcus aureus* response to unsaturated long chain free fatty acids: survival mechanisms and virulence implications. *PLoS ONE* 4:e4344. doi: 10.1371/journal.pone.0004344
- Kim, D. H., Im, H., Jee, J. G., Jang, S. B., Yoon, H. J., Kwon, A. R., et al. (2014a). β -arm flexibility of HU from *Staphylococcus aureus* dictates the DNA-binding and recognition mechanism. *Acta Crystallogr. D Biol. Crystallogr.* 70, 3273–3289. doi: 10.1107/S1399004714023931
- Kim, H. K., Missiakas, D., and Schneewind, O. (2014b). Mouse models for infectious diseases caused by *Staphylococcus aureus*. *J. Immunol. Methods* 410, 88–99. doi: 10.1016/j.jim.2014.04.007
- Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685. doi: 10.1038/227680a0
- Lan, L., Cheng, A., Dunman, P. M., Missiakas, D., and He, C. (2010). Golden pigment production and virulence gene expression are affected by metabolisms in *Staphylococcus aureus*. *J. Bacteriol.* 192, 3068–3077. doi: 10.1128/JB.00928-09
- Liu, G. Y., Essex, A., Buchanan, J. T., Datta, V., Hoffman, H. M., Bastian, J. F., et al. (2005). *Staphylococcus aureus* golden pigment impairs neutrophil killing and promotes virulence through its antioxidant activity. *J. Exp. Med.* 202, 209–215. doi: 10.1084/jem.20050846
- Lowder, B. V., Guinane, C. M., Ben Zakour, N. L., Weinert, L. A., Conway-Morris, A., Cartwright, R. A., et al. (2009). Recent human-to-poultry host jump, adaptation, and pandemic spread of *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19545–19550. doi: 10.1073/pnas.0909285106
- Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., et al. (2017). CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.* 45, D200–D203. doi: 10.1093/nar/gkw1129
- Minden, J. S., Dowd, S. R., Meyer, H. E., and Stühler, K. (2009). Difference gel electrophoresis. *Electrophoresis* 30, S156–S161. doi: 10.1002/elps.200900098
- Müller, I. B., Bergmann, B., Groves, M. R., Couto, I., Amaral, L., Begley, T. P., et al. (2009). The vitamin B1 metabolism of *Staphylococcus aureus* is controlled at enzymatic and transcriptional levels. *PLoS ONE* 4:e7656. doi: 10.1371/journal.pone.0007656
- Müller, M., Reiß, S., Schlüter, R., Mäder, U., Beyer, A., Reiß, W., et al. (2014). Deletion of membrane-associated Asp23 leads to upregulation of cell wall stress genes in *Staphylococcus aureus*. *Mol. Microbiol.* 93, 1259–1268. doi: 10.1111/mmi.12733
- Nubel, U., Roumagnac, P., Feldkamp, M., Song, J.-H., Ko, K. S., Huang, Y.-C., et al. (2008). Frequent emergence and limited geographic dispersal of methicillin-resistant *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14130–14135. doi: 10.1073/pnas.0804178105
- Oberto, J., Nabti, S., Jooste, V., Mignot, H., and Rouviere-Yaniv, J. (2009). The HU regulon is composed of genes responding to anaerobiosis, acid stress, high osmolarity and SOS induction. *PLoS ONE* 4:e4367. doi: 10.1371/journal.pone.0004367
- O'Brien, M. M., Walsh, E. J., Massey, R. C., Peacock, S. J., and Foster, T. J. (2002). *Staphylococcus aureus* clumping factor B (ClfB) promotes adherence to human type I cytokeratin 10: implications for nasal colonization. *Cell. Microbiol.* 4, 759–770. doi: 10.1046/j.1462-5822.2002.00231.x
- Polakowska, K., Lis, M. W., Helbin, W. M., Dubin, G., Dubin, A., Niedziolka, J. W., et al. (2012). The virulence of *Staphylococcus aureus* correlates with strain genotype in a chicken embryo model but not a nematode model. *Microbes Infect.* 14, 1352–1362. doi: 10.1016/j.micinf.2012.09.006

- Rasigade, J. P., and Vandenesch, F. (2014). *Staphylococcus aureus*: a pathogen with still unresolved issues. *Infect. Genet. Evol.* 21, 510–514. doi: 10.1016/j.meegid.2013.08.018
- Recker, M., Laabei, M., Toleman, M. S., Reuter, S., Saunderson, R. B., Blane, B., et al. (2017). Clonal differences in *Staphylococcus aureus* bacteraemia-associated mortality. *Nat. Microbiol.* 2, 1381–1388. doi: 10.1038/s41564-017-0001-x
- Roche, F. M., Massey, R., Peacock, S. J., Day, N. P., Visai, L., Speziale, P., et al. (2003). Characterization of novel LPXTG-containing proteins of *Staphylococcus aureus* identified from genome sequences. *Microbiology* 149, 643–654. doi: 10.1099/mic.0.25996-0
- Sabat, A., Krzyszton-Russjan, J., Strzalka, W., Filipek, R., Kosowska, K., Hryniewicz, W., et al. (2003). New method for typing *Staphylococcus aureus* strains: multiple-locus variable-number tandem repeat analysis of polymorphism and genetic relationships of clinical isolates. *J. Clin. Microbiol.* 41, 1801–1804. doi: 10.1128/JCM.41.4.1801-1804.2003
- Sabat, A. J., Pournaras, S., Akkerboom, V., Tsakris, A., Grundmann, H., and Friedrich, A. W. (2015). Whole-genome analysis of an oxacillin-susceptible CC80 *mecA*-positive *Staphylococcus aureus* clinical isolate: insights into the mechanisms of cryptic methicillin resistance. *J. Antimicrob. Chemother.* 70, 2956–2964. doi: 10.1093/jac/dkv210
- Sauer, U., and Eikmanns, B. J. (2005). The PEP-pyruvate-oxaloacetate node as the switch point for carbon flux distribution in bacteria. *FEMS Microbiol. Rev.* 29, 765–794. doi: 10.1016/j.femsre.2004.11.002
- Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996). Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal. Chem.* 68, 850–858. doi: 10.1021/ac950914h
- Sievers, F., and Higgins, D. G. (2014). Clustal omega, accurate alignment of very large numbers of sequences. *Methods Mol. Biol.* 1076, 105–116. doi: 10.1007/978-1-62703-646-7_6
- Solis, N., Larsen, M. R., and Cordwell, S. J. (2010). Improved accuracy of cell surface shaving proteomics in *Staphylococcus aureus* using a false-positive control. *Proteomics* 10, 2037–2049. doi: 10.1002/pmic.200900564
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 9, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Tamura, K., Stecher, G., Peterson, D., Filipowski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Tavaré, S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. *Am. Math. Soc. Lect. Math. Life Sci.* 17, 59–86.
- Thomer, L., Schneewind, O., and Missiakas, D. (2016). Pathogenesis of *Staphylococcus aureus* bloodstream infections. *Annu. Rev. Pathol. Mech. Dis.* 11, 343–364. doi: 10.1146/annurev-pathol-012615-044351
- Thurlow, L. R., Joshi, G. S., Clark, J. R., Spontak, J. S., Neely, C. J., Maile, R., et al. (2013). Functional modularity of the arginine catabolic mobile element contributes to the success of USA300 methicillin-resistant *Staphylococcus aureus*. *Cell Host Microbe* 13, 100–107. doi: 10.1016/j.chom.2012.11.012
- Timms, J. F., and Cramer, R. (2008). Difference gel electrophoresis. *Proteomics* 8, 4886–4897. doi: 10.1002/pmic.200800298
- Walsh, E. J., Miajlovic, H., Gorkun, O. V., and Foster, T. J. (2008). Identification of the *Staphylococcus aureus* MSCRAMM clumping factor B (ClfB) binding site in the alphaC-domain of human fibrinogen. *Microbiology* 154, 550–558. doi: 10.1099/mic.0.2007/010868-0
- Wertheim, H. F., Walsh, E., Choudhury, R., Melles, D. C., Boelens, H. A., Miajlovic, H., et al. (2008). Key role for clumping factor B in *Staphylococcus aureus* nasal colonization of humans. *PLoS Med.* 5:e17. doi: 10.1371/journal.pmed.0050017
- Wright Valderas, M., Gatson, J. W., Wreyford, N., and Hart, M. E. (2002). The superoxide dismutase gene *sodM* is unique to *Staphylococcus aureus*: absence of *sodM* in coagulase-negative staphylococci. *J. Bacteriol.* 184, 2465–2472. doi: 10.1128/JB.184.9.2465-2472.2002
- Xiang, S., and Tong, L. (2008). Crystal structures of human and *Staphylococcus aureus* pyruvate carboxylase and molecular insights into the carboxyltransfer reaction. *Nat. Struct. Mol. Biol.* 15, 295–302. doi: 10.1038/nsmb.1393
- Xiong, Y. Q., Yang, S. J., Tong, S. Y. C., Alvarez, D. N., and Mishra, N. N. (2015). The role of staphylococcal carotenogenesis in resistance to host defense peptides and *in vivo* virulence in experimental endocarditis model. *Pathog. Dis.* 73:ftv056. doi: 10.1093/femspd/ftv056
- Zhao, F., Chong, A. S., and Montgomery, C. P. (2016). Importance of B lymphocytes and the IgG-binding protein Sbi in *Staphylococcus aureus* skin infection. *Pathogens* 5:12. doi: 10.3390/pathogens5010012

Conflict of Interest Statement: 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.

Copyright © 2018 Bonar, Bukowski, Hydzik, Jankowska, Kedracka-Krok, Groborz, Dubin, Akkerboom, Miedzobrodzki, Sabat, Friedrich and Wladyka. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.