

Dataset brief

Comparative proteomic profiling of methicillin-susceptible and resistant *Staphylococcus aureus*

Zhen Xu^{1,2}, Jiazhen Chen³, Kostas Vougas⁴, Ajit Shah⁵, Haroun Shah⁵, Raju Misra⁶, Hermine V Mkrtchyan^{2,7*}

1. Department of Sanitary Toxicology and chemistry, Tianjin Key Laboratory of Environment, Nutrition and Public Health; Center for International Collaborative Research on Environment, Nutrition and Public health, Tianjin Medical University, Tianjin, China
2. School of Biological and Chemical Sciences, Queen Mary University of London, London, UK
3. Department of Infectious Diseases, Huashan Hospital, Fudan University, Shanghai, China
4. Biotechnology Division, Biomedical Research Foundation of the Academy of Athens, Athens, Greece
5. Department of Natural Sciences, Middlesex University, London, UK
6. Natural History Museum, Core Research Laboratories, Molecular Biology, Cromwell Rd, London, UK
7. School of Health, Sport and Biosciences, University of East London, London, UK

*Corresponding author

*Hermine V Mkrtchyan: School of Health, Sport and Biosciences, University of East London, London, UK. E1 4NS; E-mail: h.mkrtchyan@uel.ac.uk

Received: 7 3, 2019; Revised: 12 10, 2019; Accepted: 12 20, 2019

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/pmic.201900221](https://doi.org/10.1002/pmic.201900221).

This article is protected by copyright. All rights reserved.

List of abbreviations: BP: Biological Process; CC: Cellular Component; EMRSA-15: Epidemic methicillin-resistant *Staphylococcus aureus* types 15; FDR: False discovery rate; Fg: human fibrinogen; GO: Gene Ontology; HPLC: high-performance liquid chromatography; iTRAQ: isobaric tag for relative and absolute quantitation; KEGG: Kyoto Encyclopedia of Genes and Genomes; MALDI-TOF MS: Matrix Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry; MF: Molecular Function; MMTS: methanethiosulfonate; MRSA: Methicillin resistant *Staphylococcus aureus*; MSSA: Methicillin susceptible *Staphylococcus aureus*; MS: mass spectrometry; NCBI: National Center for Biotechnology Information; PPI: Protein-Protein Interact Network; RP: Reverse Phase; TCEP: tris-2-carboxymethyl phosphine; TEAB: triethylammonium bicarbonate

Keywords: Comparative proteomics, MRSA, MSSA

Total number of words: 3800 words

Abstract

Purpose: *Staphylococcus aureus* is a highly successful human pathogen responsible for wide range of infections. In this study, we provide insights into the virulence, pathogenicity, and antimicrobial resistance determinants of methicillin susceptible and methicillin resistant *Staphylococcus aureus* (MSSA; MRSA) recovered from non-healthcare environments.

Experiment design: Three environmental MSSA and three environmental MRSA were selected for proteomic profiling using iTRAQ MS/MS. Gene Ontology (GO) Annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway Annotation were applied to interpret the functions of the proteins detected.

Results: 792 proteins were identified in MSSA and MRSA. Comparative analysis of MRSA and MSSA revealed that 8 of out 792 proteins were up-regulated and 156 were down-regulated. Proteins that had differences in abundance were predominantly involved in catalytic and binding activity. Among 164 differently abundant proteins, 29 were involved in pathogenesis, antimicrobial resistance, stress response, mismatch repair and cell wall synthesis. Twenty-two proteins

associated with pathogenicity, including *spa*, *sbi*, *clfA* and *dlt* were up-regulated in MRSA. Moreover, the up-regulated pathogenic protein *entC2* in MSSA was determined to be a super antigen potentially capable of triggering toxic shock syndrome in the host.

Conclusions: Enhanced pathogenicity, antimicrobial resistance and stress response were observed in MRSA compared to MSSA.

Significance: In this study we have unravelled the variation of virulence, pathogenicity, stress response factors and antimicrobial resistance of environmental (non-healthcare) MRSA and MSSA. iTRAQ MS/MS analyses were used to compare differences in protein abundance among representative strains of these *S. aureus* isolates.

S. aureus isolated from different ecological niches has frequently been reported, however, studies reporting proteomic profiling of these isolates have been fragmentary. We believe that our study is of major importance as it reports differences in protein abundance between environmental MRSA and MSSA and demonstrates a significant variation in pathogenicity, antimicrobial resistance and stress response, hence identifying the pathogenic potential of these isolates.

Staphylococcus aureus is a highly successful human pathogen responsible for a wide range of mild to life threatening infections ^[1]. The pathogenicity of *S. aureus* is initiated by adhesion of the organism to the host cells and subsequent secretion of toxins that are detrimental to the cells. Various virulence factors are involved in colonization and invasion of host tissue and evasion of host immune system ^[2]. The abundance of virulence factors can be affected by different conditions, including environmental stress and nutritional variations ^[3].

MRSA infections occur both in the community and healthcare settings, posing greater challenges for public health. More recently, attention has been paid to the dissemination of MRSA in the environment ^[5,6]. Traditionally healthcare-associated MRSA clones, including ST5, ST30 harbouring various virulence factors have been found in the environment ^[7, 8].

Comparative proteomic analysis of hospital-associated MRSA and MSSA have been reported ^[1], however, to date, studies reporting comparative proteomic profiling of environmental MRSA and MSSA are scarce.

Therefore, in an attempt to unravel the variation of virulence, pathogenicity, and stress response factors and antimicrobial resistance of environmental MRSA and MSSA, iTRAQ MS/MS analyses were used to compare differences in protein abundance among representative strains of *S. aureus* isolates.

Three different biological replicates of MSSA and MRSA isolates recovered from non-healthcare environments were included in this study^{[5][7][8]}. Six strains were selected from several hundred strains as part of a larger study on the distribution and antibiotic profiling of *S. aureus*^{[7][8]}. Antibiotic susceptibility of all *S. aureus* isolates was determined using disc diffusion method. In addition, the minimum inhibitory concentrations (MIC) for oxacillin was determined for MRSA isolates and those included in this study had MICs of 2 mg/L. In addition, the presence of *mecA* gene was confirmed by PCR in all MRSA isolates and sequence types were determined to be ST22. The cell lysate was prepared by using a 'glass bead beating' method as described previously^[9]. The samples were frozen and then freeze dried using a Coolsafe (Jencons-VWR, East Grinstead, UK). Protein was digested with trypsin following reduction and alkylation and subsequently labelled with iTRAQ 8 plex. High-pH RP C18 fractionation of the iTRAQ 8plex labeled peptides was performed using a Dionex HPLC system composed of P680 pumps, and a PDA-100 photodiode array detector. Eight 1 ml fractions were collected and dried using a speed vac concentrator and stored at -20° C until the LC-MS analysis. All LC-MS experiments were performed using a Dionex Ultimate 3000 UHPLC system coupled to a high resolution nano-ESI Orbitrap-Elite mass spectrometer (Thermo Scientific). The HCD tandem mass spectra collected from the analysed fractions were processed using Proteome Discoverer® (version 1.4, Thermo Scientific) for peptide and protein identification and relative quantification. Database pattern was Decoy (peptide FDR≤0.01) and the protein ratios were calculated as the median of only unique peptides of the protein. The peptide false discovery rates (FDR) was controlled through the software at 1%. iTRAQ ratios meta-analysis were carried out in R language [R_Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>]. Proteins with a *p*-value<0.05 were considered differences in protein abundance. The FASTA protein sequences of differently abundant proteins were retrieved from UniProtKB database (Release 2016_10), and the sequences were searched against SwissProt database using the NCBI BLAST+ client software (ncbi-blast-2.2.28+-win32.exe) to

annotate the studied sequences. The annotation configuration was as follows: E-value filter of 1e-6, default gradual EC weights, a GO weight of 5, and an annotation cut-off of 75. Moreover, InterProScan10 against EBI databases was used to annotate the proteins that failed to be annotated using Blast2GO (Version 3.3.5). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD014478.

The protein sequences of differently abundant proteins were blasted against the KEGG database (<http://geneontology.org/>) in FASTA format to retrieve their Kos (KEGG Orthologies) and then mapped to pathways in KEGG11. The matching KEGG pathways were extracted. Fisher' exact test (threshold p-value < 0.05) was used to explore GO enrichment on biological process, molecular function, and cellular component, as well as KEGG pathway enrichment. Cluster 3.0 (<http://bonsai.hgc.jp/~mdehoon/software/cluster/software.htm>) and the Java Treeview software (<http://jtreeview.sourceforge.net>) were used to do hierarchical clustering analysis. The data of protein–protein interaction was retrieved from IntAct molecular interaction database (<http://www.ebi.ac.uk/intact/>) by using gene symbols or STRING software (<http://string-db.org/>).

Sequencing of *S. aureus* genomes are now undertaken almost routinely; thus the core and accessory genomes are more clearly defined ^[10]. By contrast the proteomes of strains, particularly those of non-clinical isolates have received less attention. Environmental isolates are known to harbour antibiotic resistance to different extents but the factors that enable selective strains to enter the human ecosystem and give rise to disease outbreaks are still partly conjectural. In a preliminary study, we sampled a vast number of environmental sites and collected nearly 1000 isolates which were subtyped and antibiograms carried out to ascertain any direct correlations ^[11]. While this was not achievable, it was clear that methicillin resistance occurred among environmental strains. In an early study we reported a continuum of strains in the process of transition from sensitive to methicillin resistance based upon rapid scanning of their proteomes using mass spectrometry ^[12]. Here we focused on three representative strains each of MRSA and MSSA that were distinctly sensitive or resistant to methicillin and carried out more detailed analysis of their proteomes to gain insight into how methicillin resistance may affect the global regulation of the cell.

The profiles of MRSA and MSSA at protein level were determined using an LC-MS/MS quantitative method based on iTRAQ. PD (Proteome Discoverer 1.3 (Thermo Electron, San Jose, CA) was used to identify and quantify the proteins, and a total of 792 proteins with an $FDR \leq 0.01$ were identified in both MSSA and MRSA. In comparison with MRSA, the proteins of MSSA that were up regulated two fold or down-regulate 0.5 times were considered to have statistical significance. There were 8 up-regulated and 156 down-regulated proteins.

The GO-annotation is a bioinformatic tool, which provides functional information of gene products and describes functions through the adoption of domain-specific ontologies. Moreover, GO annotation is based on the protein abundance levels ^[13]. In this study, the differences in protein abundance were interpreted as follows: Biological Process (BP), Molecular Function (MF), and Cellular Component (CC). The BP includes metabolic process (n=128), cellular process (n=126), biological regulation (n=23), regulation of biological process (n=21), response to stimulus (n=21), localization (n=19), multi-organism process (n=13), cellular component organization or biogenesis (n=8), detoxification (n=6) developmental process (n=5), signalling (n=3), positive regulation of biological process (n=3), multicellular organismal process (n=2), negative regulation of biological process (n=2), cell killing (n=1), immune system process (n=1), behaviour (n=1), carbohydrate utilization (n=1), and biological adhesion (n=1). The differences in protein abundance were mainly involved in metabolic and cellular process. On the other hand the differences in protein abundance of MF were involved in catalytic activity (n=124), binding (n=96), transporter activity (n=14), antioxidant activity (n=6), structural molecule activity (n=2), molecular carrier activity (n=2), transcription regulator activity (n=2), signal transducer activity (n=1) and molecular transducer activity (n=1). The majority differences in protein abundance of CC occurred in cell (n=97), followed by cell part (n=93), membrane (n=25), membrane part (n=19), macromolecular complex (n=18), extracellular region (n=12), organelle (n=3), membrane-enclosed lumen (n=1), and organelle part (n=1). The un-ignored amount of differences in protein abundance was mainly located on the membrane. The pathogenesis of *S. aureus* is associated with the synthesis of cell wall associated adhesions and the secretion of extracellular toxins with damaging effects on host cells ^[2] (Figure 1). Therefore, the changes in the membrane may reflect variations in toxicity.

Among the 164 proteins that were tested, 29 were involved in the pathogenesis, antimicrobial resistance, stress response, mismatch repair and cell wall synthesis, and 7 proteins were up-regulated in MSSA compared with MRSA, four of which were involved in pathogenesis. These were mainly located at the extracellular region and cell membrane. *KdpC* and *mscL* were located on the plasma membrane and were involved in ion and potassium transfer. In addition, *mraY* that was involved in the cell wall organization was also located on the plasma membrane. There were 156 proteins that were up regulated in MRSA compared with MSSA. Of these, 22 proteins were involved in pathogenesis, antimicrobial resistance and cell wall synthesis. Nine proteins were involved in pathogenesis, including ISAB, SARS, EBH, TPIS, SPA, SBI, ENO, ROT, CLPL. Immunodominant antigen B (*isaB*) gene is considered as core-variable gene and encodes for proteins that bind to the host tissue^[14]. The up-regulation of *isaB* gene suggested increased pathogenicity of MRSA in comparison with MSSA. Cell wall-associated fibronectin binding protein (EBH) alters cell size and complement resistance in *S. aureus*, therefore the up-regulation of EBH in this study suggested increased resistance to methicillin. Meanwhile, EBH enhances the stability of peptidoglycan structure of the cell wall and thus maintains the stability of cell structure^[15]. Triosephosphate isomerase (TPIS) is vital for carbohydrate metabolism of cells. TPIS is associated with pathogenicity. Staphylococcal protein A (SPA) enhances nasal colonization cell adhesion, and facilitates its dissemination^[16]. The up-regulation of *spa* gene product is an enigma finding. SBI is known as staphylococcal immunoglobulin-binding protein, and favour the host evasion^[17]. The *eno* gene was reported in *S. aureus* that causes infections, and is known to be responsible for colonization^[18]. *Rot* is a regulator of virulence genes, and affects the ability of the organism to bind to human fibrinogen (FG)^[19]. CLP proteolytic complexes are known to adapt to stress environment and degrade the mis-folded proteins^[20]. *SarS* is the homolog of *sarA*, which is located upstream of the *spa* gene. SARS is an activator of *spa* expression and up-regulated SARS increases the expression of *spa*^[21]. This is consistence with our findings of increased abundance of SPA. CLFA, DLTC and SSPB proteins are involved in cell shape maintenance. Both MRSA and MSSA had proteins that were involved in pathogenesis, antimicrobial resistance and cell wall synthesis changes. However, proteins that were involved in stress response, mismatch repair and antimicrobial resistance were exclusively up-regulated in MRSA. Generally, MRSA is different from MSSA due to acquisition of an additional *mecA* gene, which encodes penicillin binding protein 2a and is responsible for the ensuing methicillin resistance. Three proteins, including

THIO, BSAA, and SAR are involved in stress response. The staphylococcal accessory regulator (SAR) is known to repress the transcription of collagen adhesin gene, which is independent from *agr* regulation^[22]. MUTS2 and MUTL are involved in mismatch repair, which is known to play a vital role in the adaptation of bacteria into changing to a stress environment. The up-regulation of MUTS and MUTL is necessary to maintain the stability of bacterial cells and avoid hypermutability, which may contribute to the erratic chromosome structure^[23]. Five proteins were involved in antimicrobial resistance, including MSRA, MSRB, ERMA, PLS and TETM. *msrA* and *msrB* are genes that encode macrolide efflux pump, which mediates the macrolide resistance of *S. aureus*^[24]. The up-regulation of MSRA and MSRB dramatically increases the macrolide resistance in MRSA. *ermA* is another antimicrobial resistance gene that confers macrolide resistance in *S. aureus*^[25]. PLS is one of the surface proteins of *S. aureus*, and is believed to be associated with the methicillin resistance^[26]. *tetM* encodes an energy dependent efflux pump of tetracycline, and thus confers tetracycline resistance^[27]. As a result of a number of up-regulated antibiotic resistance genes, infections caused by MRSA have become more difficult to treat with alternative antibiotics.

S. aureus can cause a range of infections, including minor skin infections, life-threatening diseases and food poisoning. It employs several pathways to make the immune system ineffective, including inhibition of neutrophils by immune modulating proteins, accumulation of the positive net charge of cytoplasmic membrane to resist cationic antimicrobial peptides (such as defensin), and inhibiting immune response with the expression of superantigens^[28]. SPA, SBI and CLFA are surface proteins that release cytotoxic toxin, which have leukotoxic activity. Furthermore, it may inhibit the transmission of neutrophils, and thus inhibit the host immune system. Moreover, DLT may inhibit the activity of β -Defensin, and thus increase the resistance to antimicrobial peptides (Figure 2). In this study, SPA, SBI, CLFA and DLT were up-regulated in MRSA suggesting the enhanced pathogenicity of MRSA (Table S2). In this study we also showed that the up-regulated pathogenic protein ENTC2. ENTC2 is a super antigen that potentially can trigger toxic shock syndrome in the host^[29].

We demonstrated difference in protein abundance between environmental MRSA and MSSA strains used in this study. Despite the limited number of representative strains of MRSA and MSSA, we provide insights into

the variations in pathogenicity, antimicrobial resistance and stress response factors. From the standpoint of human healthcare, additional comparative analyses are required to substantiate these findings and ascertain the degree of pathogenicity in various environments and subsequently aid our understanding whether they are transferrable to human pathogens.

Acknowledgement

This work was part of Zhen Xu's PhD study funded by China Scholarship Council. All authors read and approved the final manuscript.

Conflict of interest:

The authors declare they have no conflict of interest.

References

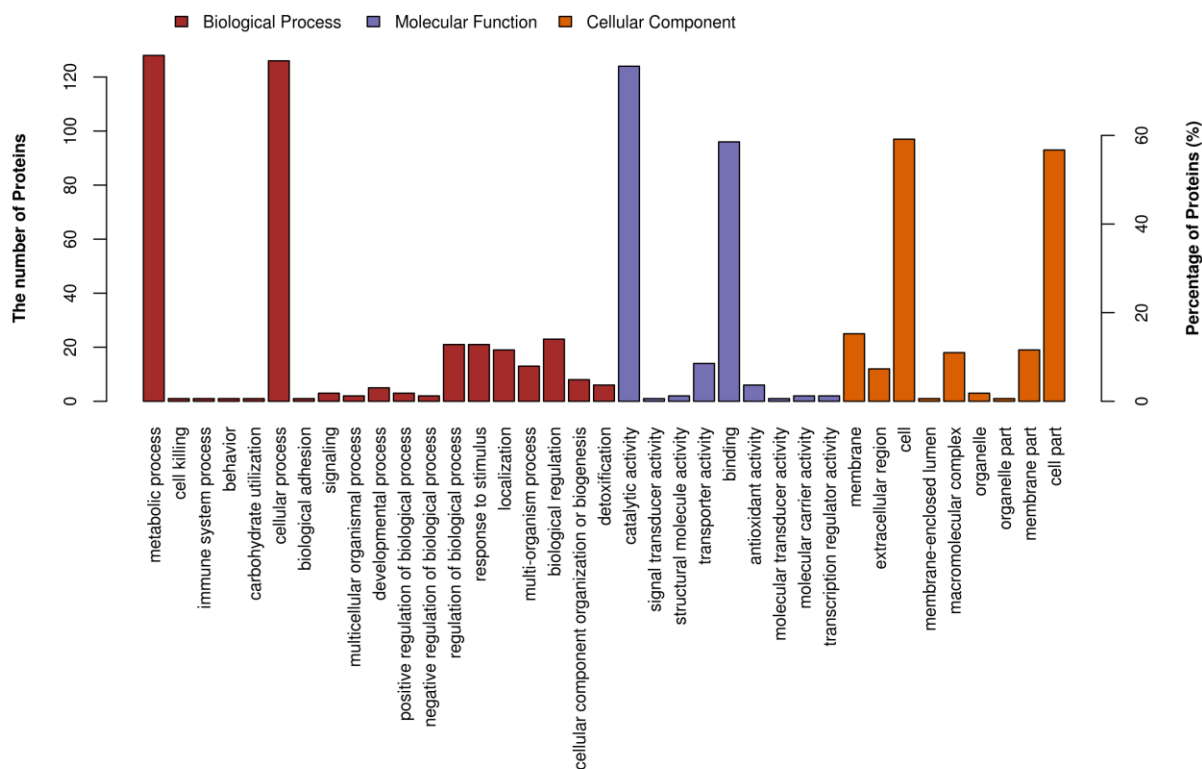
- [1] Atshan, S.S., Shamsudin, M.N., Sekawi, Z., Lung, L.T.T., Barantalab, F., Yun, K.L., Alreshidi, M.A., Abduljaleel, S.A., Hamat, R.A., *Front. Microbiol.* **2015**, 6, 524.
- [2] Enany, S., Yoshida, Y., Yamamoto, T., *World J. Microbiol. Biotechnol.* **2014**, 30, 1269–1283.
- [3] Bernardo, K., Fleer, S., Pakulat, N., Krut, O., Hunger, F., Kronke, M., *Proteomics* **2002**, 2, 740–746.
- [4] Monecke, S., Coombs, G., Shore, A.C., Coleman, D.C., Akpaka, P., Borg, M., Chow, H., Ip, M., Jatzwauk, L., Jonas, D., Kadlec, K., Kearns, A., Laurent, F., O'Brien, F.G., Pearson, J., Ruppelt, A., Schwarz, S., Scicluna, E., Slickers, P., et al., *PLoS One* **2011**, 6, e17936.
- [5] Mkrtchyan, H. V., Xu, Z., Cutler, R.R., *BMC Microbiol.* **2015**, 15, 120.
- [6] Peng, Y., Ou, Q., Lin, D., Xu, P., Li, Y., Ye, X., Zhou, J., Yao, Z., *Sci. Rep.* **2015**, 5, 16087.
- [7] Mkrtchyan, H. V., Russell, C.A., Wang, N., Cutler, R.R., *PLoS One* **2013**, 8, e54223.

- [8] Xu, Z., Shah, H.N., Misra, R., Chen, J., Zhang, W., Liu, Y., Cutler, R.R., Mkrtchyan, H. V, *Antimicrob. Resist. Infect. Control* **2018**, 7, 73.
- [9] Dekio, I., Culak, R., Fang, M., Ball, G., Gharbia, S., Shah, H.N., *Biomed Res. Int.* **2013**, 2013, 151797.
- [10] Koser, C.U., Holden, M.T.G., Ellington, M.J., Cartwright, E.J.P., Brown, N.M., Ogilvy-Stuart, A.L., Yang, H.L., Claire, C., Croucher, N.J., Harris, S.R., *N Engl J Med* **2012**, 366, 2267–2275.
- [11] Zhen Xu, Dissemination, Antibiotic Susceptibility, Proteomic and Genomic Characterization of Antibiotic-Resistant Staphylococci Recovered from General Public Settings. Queen Mary University of London, **2016**.
- [12] Lancashire, L., Schmid H, O., Ball, G., *Bioinformatics* **2005**, 21, 2191–2199.
- [13] Dong, F., Ji, Z.B., Chen, C.X., Wang, G.Z., Wang, J.M., *Int. J. Genomics*, 2013, (2013-9-30) **2013**, 2013, 917342.
- [14] Lindsay, J.A., Moore, C.E., Day, N.P., Peacock, S.J., Witney, A.A., Stabler, R.A., Husain, S.E., Butcher, P.D., Hinds, J., *J. Bacteriol.* **2005**, 188, 669–676.
- [15] Cheng, A.G., Missiakas, D., Schneewind, O., *J. Bacteriol.* **2014**, 196, 971–81.
- [16] Hong, X., Qin, J., Li, T., Dai, Y., Wang, Y., Qian, L., Lei, H., Lu, H., Gao, Q., Yong, L., *Front. Microbiol.* **2016**, 7, 951.
- [17] Burman, J.D., Leung, E., Atkins, K.L., O’Seaghda, M.N., Lango, L., Bernado, P., Bagby DI, Ss., Foster, T.J., Isenman, D.E., *Mol. Immunol.* **2007**, 44, 3982.
- [18] Ghasemian, A., Peerayeh, S.N., Bakhshi, B., Mirzaee, M., *Arch. Clin. Infect. Dis.* **2015**, 10, 8538–8540.
- [19] Xue, T., You, Y., Shang, F., Sun, B., *Med. Microbiol. Immunol.* **2012**, 201, 81–92.
- [20] Michel, A., Agerer, F., Hauck, C.R., Herrmann, M., Ullrich, J., Hacker, J., Ohlsen, K., *J. Bacteriol.* **2006**, 188, 5783–96.

- [21] Cheung, A.L., Schmidt, K., Bateman, B., Manna, A.C., *Infect. Immun.* **2001**, 69, 2448–2455.
- [22] Blevins, J.S., Gillaspay, A.F., Rehtin, T.M., Hurlburt, B.K., Smeltzer, M.S., *Mol. Microbiol.* **2010**, 33, 317–326.
- [23] Prunier, A.L., Leclercq, R., *J. Bacteriol.* **2005**, 187, 3455–3464.
- [24] Wang, Y., Wu, C.-M., Lu, L.-M., Ren, G.-W.N., Cao, X.-Y., Shen, J.-Z., *Vet. Microbiol.* **2008**, 130, 118–25.
- [25] Zmantar, T., Kouidhi, B., Miladi, H., Bakhrouf, A., *BMC Res. Notes* **2011**, 4, 453.
- [26] Gill, S.R., Fouts, D.E., Archer, G.L., Mongodin, E.F., Deboy, R.T., Ravel, J., Paulsen, I.T., Kolonay, J.F., Brinkac, L., Beanan, M., Dodson, R.J., Daugherty, S.C., Madupu, R., Angiuoli, S. V, Durkin, A.S., Haft, D.H., Vamathevan, J., Khouri, H., Utterback, T., et al., *J. Bacteriol.* **2005**, 187, 2426–38.
- [27] Ng, L.K., Martin, I., Alfa, M., Mulvey, M., *Mol. Cell. Probes* **2001**, 15, 209–15.
- [28] Foster, T.J., *Nat. Rev. Microbiol.* **2005**, 3, 948–958.
- [29] Sergeev, N., Volokhov, D., Chizhikov, V., Rasooly, A., *J. Clin. Microbiol.* **2004**, 42, 2134–2143.
- [30] Kanehisa, M. , Goto, S. , Kawashima, S. , Nakaya, A. , & S., S. K. *Nucleic Acids Research* **2000**, volume 28, 27-30.

Figure 1 GO annotation of differences in proteins abundance in Gene Ontology level 2 group

X axis GO terms in three categories: biological process (red), molecular function (purple) and cellular compartment (orange), Y axis on the left: the percentage of annotated proteins in all differences in protein abundance, Y axis on the right: the number of the annotated differently abundant proteins Figure 2 KEGG pathway map



Proteins involved in pathogenicity of *S. aureus* are highlighted red in the KEGG pathway map.

Staphylococcus aureus develop several ways to compromise the efficiency of the immune system, including secreting immune modulating proteins SpA which impede apoptosis, Dlt may mediate increase of the positive net charge of cytoplasmic membrane in order to reduce the sensitivity to cationic antimicrobial peptides (such as defensin), and preventing immune response by expression of superantigens SpA, Sbi and ClfA^[30].

