Insights into Alpha-Hemolysin (Hla) Evolution and Expression among *Staphylococcus aureus* Clones with Hospital and Community Origin

Ana Tavares¹, Jesper B. Nielsen², Kit Boye², Susanne Rohde², Ana C. Paulo³, Henrik Westh²,⁴, Kristian Schønning²,⁴, Hermínia de Lencastre¹,⁵, Maria Miragaia¹,⁶*

¹Laboratory of Molecular Genetics, Instituto de Tecnologia Química e Biológica (ITQB), Oeiras, Portugal, ²Dept. of Clinical Microbiology 445, Copenhagen University Hospital, Hvidovre, Denmark, ³Molecular Microbiology of Human Pathogens, ITQB, Oeiras, Portugal, ⁴Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, ⁵Laboratory of Microbiology and Infectious Diseases, The Rockefeller University, New York, New York, United States of America, ⁶Laboratory of Bacterial Evolution and Molecular Epidemiology, ITQB, Oeiras, Portugal

**Abstract**

**Background:** Alpha-hemolysin (Hla) is a major virulence factor in the pathogenesis of *Staphylococcus aureus* infection, being active against a wide range of host cells. Although *hla* is ubiquitous in *S. aureus*, its genetic diversity and variation in expression in different genetic backgrounds is not known. We evaluated nucleotide sequence variation and gene expression profiles of *hla* among representatives of hospital (HA) and community-associated (CA) *S. aureus* clones.

**Methods:** 51 methicillin-resistant *S. aureus* and 22 methicillin-susceptible *S. aureus* were characterized by PFGE, spa typing, MLST and SCCmec typing. The internal regions of *hla* and the *hla* promoter were sequenced and gene expression was assessed by RT-PCR.

**Results:** Alpha-hemolysin encoding- and promoter sequences were diverse, with 12 and 23 different alleles, respectively. Based on phylogenetic analysis, we suggest that *hla* may have evolved together with the *S. aureus* genetic background, except for ST22, ST121, ST59 and ST93. Conversely, the promoter region showed lack of co-evolution with the genetic backgrounds. Four non-synonymous amino acid changes were identified close to important regions of *hla* activity. Amino acid changes in the RNAIII binding site were not associated to *hla* expression. Although expression rates of *hla* were in general strain-specific, we observed CA clones showed significantly higher *hla* expression (p = 0.003) when compared with HA clones.

**Conclusion:** We propose that the *hla* gene has evolved together with the genetic background. Overall, CA genetic backgrounds showed higher levels of *hla* expression than HA, and a high strain-to-strain variation of gene expression was detected in closely related strains.

**Citation:** Tavares A, Nielsen JB, Boye K, Rohde S, Paulo AC, et al. (2014) Insights into Alpha-Hemolysin (Hla) Evolution and Expression among *Staphylococcus aureus* Clones with Hospital and Community Origin. PLoS ONE 9(7): e98634. doi:10.1371/journal.pone.0098634

**Editor:** J. Ross Fitzgerald, University of Edinburgh, United Kingdom

**Received** February 7, 2014; **Accepted** May 6, 2014; **Published** July 17, 2014

**Copyright:** © 2014 Tavares et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was funded by project Ref. P-99911 from Fundação Calouste Gulbenkian (http://www.gulbenkian.pt/Institucional/pt/Homepage) and Project PTDC/BIA-MIC/3195/2012 from Fundação para a Ciência e Tecnologia (http://www.fct.pt/) awarded to Hdi; Project PTDC/BIA-EVF/117507/2010 from Fundação para a Ciência e Tecnologia awarded to MMe; and through grant Ref. Pest-OE/EQB/LAO004/2011 from Fundação para a Ciência e Tecnologia (FCT), Portugal. A. Tavares was supported by grant SFRH/BD/44220/2008 from FCT. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** Co-author Herminia de Lencastre is a PLOS ONE Editorial Board member. This does not alter the authors’ adherence to PLOS ONE Editorial policies and criteria.

* Email: miragaia@itqb.unl.pt

**Introduction**

*Staphylococcus aureus* is a human opportunistic pathogen responsible for a wide range of infections that can vary in its clinical presentation and severity. Methicillin-resistant *S. aureus* (MRSA) emerged in 1960 in the United Kingdom and has been a major problem in hospitals (HA-MRSA) worldwide during the last 40 years; however since the late 1990s, MRSA has been emerging as a leading cause of severe infection also in the community, in individuals without recent health-care contact (CA-MRSA) [1,2].

CA-MRSA present distinct genetic backgrounds from their hospital counterparts, are more susceptible to antibiotics other than beta-lactams, carry the smallest staphylococcal cassette chromosome *mec* types (SCC*mec* IV or V), and have higher virulence capacity [1,2,3]. The underlying reasons behind the enhanced virulence of CA-MRSA appear to be multiple including a different capacity to overcome host cell response [4], different distribution of mobile genetic elements carrying virulence determinants [5], allelic variation in virulence determinants located in the core genome and in mobile genetic elements [6], and different levels of expression and protein production of virulence determinants (alpha-hemolysin, collagen adhesin, staphylokinase, coagu-
lase, lipase, enterotoxins C3 and Q, V8 protease and cysteine protease) [7,8,9].

The alpha-hemolysin or α-toxin (Hla), is one of the major virulence determinants implicated in the pathogenesis of S. aureus, associated to severe skin and soft tissue infections (SSTI), necrotizing pneumonia and even sepsis [10]. Hla is the most prominent S. aureus cytotoxin that can act against a wide range of host cells including erythrocytes, epithelial cells, endothelial cells, T cells, monocytes and macrophages [10,11,12]. The gene encoding Hla is located in the core genome and is expressed as a water-soluble monomer (33.2 kD) that assembles to form a membrane-bound heptameric β-barrel pore (232.4 kD) on susceptible cells leading to cell death and lysis [11]. The overall structure is mushroom-like, divided into three domains: 1) Cap domain: largely hydrophobic, defining the entry of the pore; 2) Rim domain: underside of the Cap, in close proximity to membrane bilayer; 3) Stem domain: part of the transmembrane channel, forming the membrane-perforating β-barrel pore (Figure 1) [10,11]. Hla expression is mainly controlled by the global toxin accessory gene regulator (agr), via the regulatory effector molecule RNAIII [13]. While agr provides the first and most important mechanism of up-regulation of hla, expression can also be modulated by other regulators, such as SarR, SarZ, ArlS [14,15,16] (up-regulators) and Rot, SarT [17] (down-regulators).

Although polymorphisms in the hla promoter region have been described [18], the range of genetic diversity and evolution of this tox in has never been assessed in a large representative S. aureus collection. Furthermore, although differences in hla expression have been described between community- and hospital-associated MRSA, these studies have been performed with a limited number of CA-MRSA epidemic clones [9], or almost exclusively with representatives of the USA300 clone [19,20,21]. To better understand the evolutionary history of hla and its importance as a virulence factor for CA-MRSA, in this study we compared the hla nucleotide sequence and expression among the major epidemic and minor CA and HA clones, including both MRSA and MSSA strains.

Materials and Methods

Ethics Statement

Isolates were obtained from routine diagnostic and were analyzed anonymously and only the isolates, not humans, were studied. All data was collected according to the European Parliament and Council decision for the epidemiological surveillance and control of communicable disease in the European community. Ethical approval and informed consent were for that reason not required.

Bacterial collection

A total of 73 S. aureus, including 51 MRSA and 22 MSSA were analyzed in this study. Strains were collected in 13 different countries (Belgium, Bulgaria, Czech Republic, Denmark, Greece, Netherlands, Portugal, Romania, Spain, Sweden, United Kingdom, USA and Brazil), between 1961 and 2009 from both community (n = 46) and hospital (n = 27). The strains comprised a total of 52 spa types and 23 sequence types (STs) (see Table S1). Strains were defined as belonging to CA or HA clones if they contained the same or related genetic backgrounds as the reference CA-MRSA and HA-MRSA epidemic control strains, based on ST, spa type and SCCmec (in case of MRSA).

Media and bacterial growth conditions

Before RT-PCR analysis, strains were grown overnight at 37°C on tryptic soy agar plates (TSA). Bacterial growth experiments were performed by growing bacteria in tryptic Soy Broth (TSB) at 37°C with shake and measuring OD (600 nm) each hour in the follow up automatic incubator Bioscreen C (Oy Growth Curves AB, Helsinki, Finland). Plates of 100-well honeycomb (Oy Growth Curves AB, Helsinki, Finland) were filled with 300 µl/well of

Figure 1. HLA protein structure. A) wildtype (highlighted the non-synonymous mutations Gln87, Glu208, Thr239 and Ser243) and B) truncated protein due to a stop codon at Gln87. Structure generated by the program PyMOL v.1.6.

doi:10.1371/journal.pone.0098634.g001
overnight culture diluted to OD_{600} = 0.05 in TSB growth medium. Three individual growth experiments (SetC, SetD and SetE) were performed for each strain and named accordingly e.g. HLZ6C, HLZ6D and HLZ6E (see Figure S2.I to III).

Nucleotide sequence of hla and promoter region

Chromosomal DNA was extracted from overnight cultures, using the boiling method (100°C for 10 min followed by centrifuged at 13,000 g for 3 min). Two sets of primers were designed to span the most polymorphic regions within the hla gene and hla promoter (considered as the region located -600 bp from hla starting codon), after alignment of sequences available on NCBI for S. aureus. One set of primers (Forward: hla-F_CGAAAGGTACCATTTGCTGTT; Reverse: hla-R_GCAATCGATTATTATATCTTTC) amplified an internal fragment of the hla gene (nt 1170419–1170982, CP000730.1) and the other set (Forward: hlaPro-F_CACTATATTTAAATACATAC; Reverse: hlaPro-R_GTGTTTACTGAGCTGAC) amplified an internal fragment of the hla promoter region (nt 1171289–1171773, CP000730.1) (Figure S3). PCR products were sequenced (Macrogen Europe, Amsterdam, The Netherlands) and sequences were analyzed using SeqMan (DNAStar, Lasergene v9, Madison, WI, USA). To each unique hla promoter (P) and gene sequence (hla) - allelotype - a single Arabic number was attributed (e.g. P1, P2; hla1, hla2). Gene and promoter sequences were deposited in GenBank (accession numbers KM019554–KM019606; KM019607–KM019674).

Phylogenetic analysis

Phylogenetic relatedness was analyzed using the MEGA5 v5.05 software (http://www.megasoftware.net/) for gene, promoter region and concatenated sequences obtained from 1) gene with promoter region and 2) seven MLST alleles from the 23 representative STs within the collection. Phylogenetic trees were constructed using the Neighbor-Joining clustering method, and 1000 bootstrap replicates, which assigns confidence values for the groupings in the tree.

Moreover, nucleotide diversity (ND) between the two clusters was calculated based on the estimation of the average evolutionary divergence over sequence pairs within the two groups, where the number of base substitutions per site from averaging over all sequence pairs within each group are compared using the maximum composite likelihood model [22].

Detection of recombination

Alignments from the hla gene, hla promoter and internal fragments of each of the seven MLST gene were screened for the occurrence of putative recombination events using Recombination Detection Program version 4 (RDP4) (http://web.chio.uct.ac.za/) with the default settings (with highest acceptable probability value of 0.05). Identification of recombinant sequences recombination breakpoints and major parent was determined using simultaneously nine recombination detection methods (RDP, BOOTSCAN, GENECONV, MAXCHI, CHIMAERA, SISCAN, PhyPro, LARD and 3SEQ). The “minor parent” is considered a sequence closely related to that from which sequences in the proposed recombinant region may have been derived (the presumed donor). The “major parent” was considered as a sequence closely related to that from which the greater part of the recombinant’s sequence may have been derived.

RT-PCR analysis

Culture growth was stopped at late exponential phase, when alpha-toxin is described to have maximal activity [23], corresponding to the time-points 1) 3 hours 30 min in one group (65 strains) and 2) 4 hours 30 min in another (9 strains). Total RNA was extracted from three biological replicates. Cells were mechanically disrupted with FastPrep-24 Instrument (MP Biomedicals, Solon, OH, USA) and RNA was protected using RNA Protect (Qiagen, Valencia, USA). RNA was extracted automatically using the QiAsymphony platforms (Qiagen, Valencia, USA) with QiAsymphony RNA kit (Qiagen, Valencia, USA).

The RT-PCR assay was performed on a 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA) using the following primers and TaqMan probes: Hla RT-F: TAATGAAATCCTGTTCGCTAATGTC; HlaRT-R: CACCTGTTTTTACTGTTAGTTTCTTCC; Hla RT Probe: 6FAM-AAACCGGTACTAGCAGATAT-MGBNFQ. The RT-PCR reaction was performed using the EZ RT-PCR Core Reagents (Applied Biosystems, Foster City, USA), in which RNA is reverse transcribed and amplified in a single reaction. The following PCR protocol was used: 50°C for 2 min, 60°C for 30 min, 95°C for 5 min, followed by 42 cycles of 95°C for 20 sec and 62°C for 1 min. The 16S gene was used as internal or reference control. The primers used for 16S RNA amplification were those previously described [24].

RT-PCR data analysis

The relative hla gene expression was calculated based on the C_t (RT-PCR output) of the gene of interest (C_t hla) as compared to the C_t of the internal control (C_t 16S) as follows: Delta C_t = C_t hla - C_t 16S. The lower the Delta C_t the higher is the amount of hla mRNA and the more the gene is expressed. The reproducibility of the assay was evaluated by the calculation of the arithmetic mean of the relative expression of the three biological replicates (Mean Delta C_t = Average [Delta C_t1; Delta C_t2; Delta C_t3]). The reproducibility of RT-PCR reaction was evaluated by the calculation of the standard deviation (STDEV) of Delta C_t, obtained for each biological replica (Delta C_t1; Delta C_t2; Delta C_t3). Values were considered valid when at least two C_t readings exist with STDEV<2.

Protein structure visualization (pyMOL)

The protein structure was modeled using PyMOL v.1.6 (http://www.pymol.org/) if a nucleotide mutation gave rise to a stop codon.

Statistical analysis

The statistical analysis was performed using the Graphpad Prism 6 (http://www.graphpad.com/scientific-software/prism/) with the two-tailed Student’s t-test to determine whether the differences of mean expression rates (MSSA versus MSSA; HA backgrounds versus CA backgrounds) were statistically significant (p≤0.05).

Regression tree analysis was used to explore which variables could be related with the hla expression [25]. Trees explain the variation of a single response variable (in this study the hla mRNA expression) by repeatedly splitting the data into more homogeneous groups, using combinations of explanatory variables (in our case, the ST, spa type, MRSA, MSSA and the type of SCCmec).

Results

Analysis of polymorphisms in the hla gene and hla promoter

The sequence analysis of the internal region of hla and the hla promoter region among the 73 strains identified a total of 12 hla and 23 promoter region different sequences (allelotypes) (Table 1). We obtained no amplification products for hla and hla promoter
Table 1. Summary of molecular characterization, sequence variation and relative expression rates of S. aureus strains collection.

| No | Isolate ID | SCCmec | spa type | MLST | Branch | Promotor Allotype | Gene Allotype | Nonsynonymous Mutation | Hla Expression (Mean Delta Ct)^2 | Stddev Delta Ct^3 | Expression (High/Low) |
|----|------------|--------|----------|------|--------|-------------------|---------------|------------------------|-----------------------------------|-----------------|---------------------|
| 1  | HLZ6       | II     | t002     | ST5  | L      | P4               | hla1          | D208E                  | 8.69                             | 2^*             | Low                 |
| 2  | BK2464     | II     | t002     | ST5  | L      | nt               | hla1          | D208E                  | 5.37                             | 1               | High                |
| 3  | HBR73      | II     | t067     | ST5  | L      | P5               | hla1          | D208E                  | 8.75                             | 1               | Low                 |
| 4  | C013       | VI     | t002     | ST5  | L      | P3               | hla1          | D208E                  | 6.84                             | 1               | Low                 |
| 5  | HDES26     | VI     | t062     | ST5  | L      | P3               | hla1          | D208E                  | 8.01                             | 1               | Low                 |
| 6  | HDE288     | VI     | t311     | ST5  | L      | P3               | hla1          | D208E                  | 6.67                             | 1               | Low                 |
| 7  | HSA29      | –      | t002     | ST5  | L      | P3               | hla1          | D208E                  | Not Valid                        | Not Valid        | Not Valid           |
| 8  | HDE461     | IV     | t022     | ST22 | H      | P10              | hla1          | S239T; T243S           | 6.60                             | 1               | Low                 |
| 9  | HAR22      | IV     | t022     | ST22 | H      | P11              | hla1          | S239T; T243S           | 6.43                             | 1               | Low                 |
| 10 | HSMB280    | IV     | t032     | ST22 | H      | P10              | hla1          | S239T; T243S           | 4.71                             | 1               | High                |
| 11 | LBM12      | IV     | t747     | ST1806 | H | nt          | hla1          | S239T; T243S           | 9.28                             | 1               | Low                 |
| 12 | HSMB184    | –      | t5951    | ST1806 | H | P10      | hla1          | S239T; T243S           | 6.74                             | 1               | Low                 |
| 13 | HPH2       | II     | t018     | ST36 | L      | P7               | hla8          | D208E; S239T; stop codon | 8.02                             | 2^*             | Low                 |
| 14 | HAR24      | II     | t018     | ST36 | H      | nt               | hla8          | D208E; S239T; stop codon | 9.62                             | 2^*             | Low                 |
| 15 | DEN4415    | II     | t021     | ST36 | H      | P7               | hla8          | D208E; S239T; stop codon | 8.95                             | 2^*             | Low                 |
| 16 | C563       | IV     | t015     | ST45 | H      | nt               | hla10         | S239T                  | 7.02                             | 1               | Low                 |
| 17 | C036       | V      | t015     | ST45 | H      | nt               | hla10         | S239T                  | 6.24                             | 0               | Low                 |
| 18 | HAR38      | IV     | t004     | ST45 | H      | P7               | hla10         | S239T                  | 10.38                            | 1               | Low                 |
| 19 | HDFX77     | III    | t037     | ST239 | L | P1      | hla4          | –                     | 8.74                             | 2^*             | Low                 |
| 20 | HUC343     | IIIA   | t037     | ST239 | L | P1      | hla4          | –                     | 8.27                             | 0               | Low                 |
| 21 | HU25       | IIIA   | t138     | ST239 | L | P1      | hla4          | –                     | 8.17                             | 1               | Low                 |
| 22 | BK1953     | IA     | t051     | ST247 | L | P1      | hla4          | –                     | 7.71                             | 1               | Low                 |
| 23 | HPV107     | IA     | t051     | ST247 | L | P1      | hla4          | –                     | 7.56                             | 0               | Low                 |
| 24 | HSI419     | IA     | t725     | ST247 | L | P1      | hla4          | –                     | 8.23                             | 1               | Low                 |
| 25 | E2125      | I      | t051     | ST247 | L | P1      | hla4          | –                     | 7.29                             | 0               | Low                 |
| 26 | COL        | I      | t008     | ST250 | L | P2      | hla4          | –                     | 8.15                             | 1               | Low                 |
| 27 | HFX74      | IV     | t008     | ST8  | L      | P1      | hla4          | –                     | 8.01                             | 1               | Low                 |
| 28 | USA300     | IV     | t008     | ST8  | L      | P1      | hla4          | –                     | 6.46                             | 1               | Low                 |
| 29 | USA400     | IV     | t127     | ST1  | L      | P17     | hla2          | –                     | 6.01                             | 2^*             | Low                 |

Hla Evolution and Expression in S. aureus from Community and Hospital
| No | Isolate ID | SCCmec | spotype | MultiType | Branch | Allotype | Gene | Allotype | Nonsynonymous Mutation | Hla Expression (Mean Delta Ct) | StdDevDelta Ct | Expression (High/Low) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 36 | LBM05 | V | 127 | ST1 | L | P18 | hla2 | | | | | |
| 37 | CS77 | V | 121 | ST59 | L | P20 | hla5 | | | | | |
| 38 | CS38 | V | 107 | ST59 | L | P19 | hla5 | | | | | |
| 39 | C434 | V | 119 | ST79 | L | P19 | hla5 | | | | | |
| 40 | C118 | V | 101 | ST79 | L | P21 | hla7 | | | | | |
| 41 | C491 | V | 101 | ST79 | L | P12 | hla7 | | | | | |
| 42 | LBM45 | N | 101 | ST398 | H | P7 | hla11 | | | | | |
| 43 | C492 | N | 101 | ST398 | H | P13 | hla11 | | | | | |
| 44 | C496 | V | 118 | ST398 | H | P11 | hla11 | | | | | |
| 45 | LBM40 | | | ST398 | H | P11 | | | | | | |
| 46 | C385 | V | 101 | ST398 | H | P11 | | | | | | |
| 47 | C135 | N | 101 | ST398 | H | P11 | | | | | | |
| 48 | HUC56 | V | 204 | ST398 | H | P12 | | | | | | |
| 49 | C301 | N | 194 | ST398 | H | P13 | | | | | | |
| 50 | USA700 | N | 194 | ST398 | H | P14 | | | | | | |
| 51 | CUS3 | N | 194 | ST398 | H | P14 | | | | | | |
| 52 | SAMS1024 | N | 194 | ST398 | H | P14 | | | | | | |
| 53 | HUC594 | N | 194 | ST398 | H | P14 | | | | | | |
| 54 | HFA28 | N | 194 | ST398 | H | P14 | | | | | | |
| 55 | C230 | N | 194 | ST398 | H | P14 | | | | | | |
| 56 | C141 | N | 103 | ST398 | H | P14 | | | | | | |
| 57 | HBA34 | N | 103 | ST398 | H | P14 | | | | | | |
| 58 | HBA35 | N | 103 | ST398 | H | P14 | | | | | | |
| 59 | C231 | N | 103 | ST398 | H | P14 | | | | | | |

Hla Evolution and Expression in S. aureus from Community and Hospital
region in one and 13 strains, respectively, which probably result from misparing of the primers used.

From the 12 hla (hla1–12), we observed that only a single hla-allotype was found among representatives of a specific ST, except for ST22 (hla12; hla13) and ST30 (hla6; hla9) where two different alleles were identified. On the other hand, the most frequent alleles, hla1 (33.3%, n = 24) and hla4 (20.8%, n = 15), were identified in more than one ST.

Regarding the nucleotide changes identified in the hla, some correspond to non-synonymous mutations (E208, T239 and S243) and, in one particular case, to a stop codon (Table 1 and 2). The substitutions observed did not correspond to any difference in the charge or polarity of the amino acid (aa). However, changes in molecular weight were observed: i) changes from aa D208 to aa E208 (D208E) and from aa S239 to T239 (S239T) gave rise to a higher molecular weight aa; and ii) change from aa T243 to S243 (T243S) resulted in a lower molecular weight aa; of note all changes occurred in the Rim domain of the protein. In a particular case, the aa change gave rise to a stop codon located in the CAP domain, in strains of ST36. Protein structure modeling showed that a protein of about one third of its real size is produced, truncated at the Gln87 (Figure 1, A and B). The truncation is in the outside part of the domain, suggesting that this will affect the capacity of the Hla to form cell wall pores, and ultimately to induce hemolysis.

A high number of sequence variations were identified in the hla promoter region, (n = 23) (P1–23) (Table 1 and 2). Although we found that some STs were associated to a specific promoter allotype, and some promoters were identified in a single ST, we also identified cases where single STs were associated to different promoters (8 out of 23) and examples in which a single promoter allotype was associated to different STs (5 out of 23). This is the case of the most frequent promoter (P1) that was found in about one third of the strains analyzed (25.4%, n = 16), including several different STs.

A particular highly polymorphic region corresponding to nt 222 to 224 from the start codon, was found in the majority (16 out of 23) of the promoter allotypes (exceptions P1, P6, P13, P14, P15, P18 and P23). These polymorphisms are located in the vicinity of RNAIII binding site [26]; however, we could not find a direct correlation between a particular nucleotide sequence and a specific expression pattern (high or low expression). For example, the sequence TTT, observed in two strains belonging to ST398 that have a high level expression, was also observed in strains with low expression belonging to other genetic backgrounds (ST8, ST239, ST247, ST250, ST36, ST45 and ST22).

**Alpha-hemolysin evolutionary history**

In order to better understand the evolution of hla gene within the S. aureus population, we constructed phylogenetic trees from the hla and hla promoter sequences, separately or concatenated (Figure 2, A) and compared it with the tree constructed from the concatenated sequences of the seven housekeeping genes used in MLST, including all the STs represented in the strain collection described here (Figure 2, B).

The phylogenetic tree constructed for the hla gene showed two distinct major clusters with different evolutionary clocks that differed in their nucleotide diversity (ND, see Materials and Methods): cluster L with lower diversity (ND = 0.005), and cluster H with higher diversity (ND = 0.019). Cluster L included more than 70% of the strains (71.2%, n = 52), and five sub-clusters; Cluster H contained about 29% of the strains (28.8%, n = 21), and comprised four minor sub-clusters including hla8–hla12 alleles,
| Strains data distribution based on promoter allotypes. |
|------------------------------------------------------|
| **Promotor allotype** | **Gene allotype** | **Non Synonymous Mutation** | **Isolates Molecular Characterization** | **Expression Category** |
| CA backgrounds | ST398 | P13 | hla11 | – | ST398-IV, t011 | High expression |
| | | P12 | | | ST398, t034 | High expression |
| | | NT | | | ST398-VII, t108 | High expression |
| | | P12 | | | ST398-IV, t011 | High expression* |
| ST25 | P6 | hla1 | D208E | | ST25, t258 | High expression |
| | | | | | ST25, t081 | High expression* |
| | | | | | ST25, t209 | High expression |
| ST9 | P22 | hla1 | D208E | | ST9, t100 | High expression* |
| ST93 | P21 | hla7 | – | | ST93-IV, t202 | High expression |
| | | NT | | | ST93-IV, t1819 | High expression |
| ST121 | P2 | hla6 | – | | ST121-IV, t159 | High expression |
| | | P1 | | | ST121, t145 | High expression |
| | | NT | | | ST121-IV, t308 | High expression |
| | | P1 | | | ST121-IV, t1284 | Low expression* |
| ST72 | P14 | hla1 | D208E | | ST72-IV, t1148 | High expression |
| | | P14 | | | ST72, t3682 | High expression |
| | | P14 | | | ST1810-IV, t1346 | High expression |
| | | P14 | | | ST72, t126 | High expression* |
| | | P15 | | | ST72-IV, t791 | Low expression |
| | | P14 | | | ST72-IV, t1148 | Low expression |
| | | P14 | | | ST72, t1148 | Low expression |
| ST80 | P16 | hla1 | D208E | | ST80-IV, t131 | High expression |
| | | | | | ST80-IV, t044 | High expression |
| | | | | | ST80-IV, t044 | Low expression |
| ST30 | P7 | hla9 | D208E; S239T | | ST30, t1342 | High expression |
| | | NT | | | ST30-IV, t019 | High expression |
| | | P7 | | | ST30-IV, t019 | Low expression |
| | | P9 | | | ST30, t318 | Low expression |
| | | NT | | | ST30-IV, t019 | Low expression |
| | | P8 | hla8 | D208E; S239T; stop codon | | ST30, t012 | Low expression |
| ST15 | P2 | hla1 | D208E | | ST15, t084 | High expression |
| | | | | | ST15, t1346 | Low expression* |
| | | | | | ST15, t1509 | Low expression |
| ST59 | P20 | hla5 | – | | ST59-IV, t1216 | High expression |
| | | P19 | | | ST59-IV, t1437 | Low expression |
| | | P19 | | | ST59-IV, t1437 | Low expression |
| ST1 | P17 | hla2 | – | | ST1-IV, t1381 | Low expression |
| | | P17 | | | ST1-IV, t127 | Low expression* |
| | | P18 | | | ST1, t127 | Low expression |
| ST8 | P1 | hla4 | – | | ST8-IV, t008 | Low expression |
| | | | | | ST8-IV, t008 | Low expression* |
| | | | | | ST8, t008 | Low expression |
| | | | | | ST612-IV, t1257 | Not valid** |
| | | | | | ST8, t024 | Low expression |
| ST97 | P23 | hla3 | – | | ST97, t267 | Low expression |
| HA backgrounds | ST22 | P10 | hla13 | S239T; T243S | ST22-IV, t032 | High expression |
| | | P10 | hla12 | | ST22-IV, t022 | Low expression |
| | | P11 | | | ST22-IV, t022 | Low expression |
which were found in strains of ST30, ST36, ST45, ST398 and ST22.

As opposed to the phylogenetic tree constructed from hla gene, the one constructed from the promoter region did not show two distinct evolutionary branches (Figure S1). Moreover, dissimilar subgroup clustering was noticed in the tree constructed from the promoter gene sequence. For example, ST45, ST30 and ST36 backgrounds were clustered together in the promoter sequence-based tree whereas in the hla sequence-based tree ST45 was placed separately from ST30 and ST36 cluster (branch H). The same type of observations can be drawn for most of STs. Overall the promoter region showed to be more diverse than the hla gene sequence among the different backgrounds.

On the other hand, when we compared the phylogenetic tree constructed with the hla gene with that constructed from MLST concatenated genes, the same type of division into two distinct main clusters was observed (Figure 2). Moreover, the majority of STs were equally distributed between the two clusters in the two trees. The only exceptions were ST22, ST121, ST59 and ST93 that in the two trees have exchanged their positions from one cluster to the other (Figure 2, B-blue arrows).

Detection of recombination in hla gene, hla promoter and MLST genes

To understand if recombination could explain the incongruence found between the trees constructed from hla and MLST concatenated genes, we screened the hla gene, hla promoter and each MLST gene for recombination events using the RDP4 software.

The SiScan and 3Seq methods detected one recombination event in the hla gene. This event corresponded to a fragment ending in positions 385–410 of the hla alignment, however the beginning breakpoint was not possible to determine. In the collection analyzed this event was detected in five isolates belonging to ST22 or related STs (HSMB280, HDE461, HAR22 and LBM12 (TLV ST22) and HSMB184 (TLV ST22)) and four isolates of ST398 (LBM54, LBM40, C496, C482_ST398). The ST30 HFF204 strain was identified as the minor parent (97.8% identity with ST22 strains and 99.3% identity with ST398 strains) and ST121 strain HUC587 was identified as the major parent (with 100% identity to ST398 strains and 93.5–95.2% identity with ST22 strains) of the recombining fragment. A trace signal of recombination of this same event was also identified among ST45 isolates; however this signal was not statistically significant. Interestingly all the recombination events were detected in strains belonging to the high genetic diversity cluster in the tree constructed from hla gene. In the hla promoter region no recombination events were detected.

We have performed the same type of analysis using the internal sequences of each of the seven housekeeping used in MLST scheme, including the alleles present in all STs identified in this
study, however no recombination events were detected in any of the
genes.
Altogether the data gathered suggest that for the majority of
strains hla gene evolved together with the genetic background. The
different clustering of ST22 and ST121 strains, in the trees
constructed from MLST concatenated genes and hla gene, may
derive from recombination events occurring in the hla gene.
Similarly these type of events might explain the genetic diversity
observed in cluster H in the hla tree in strains belonging to ST22,
ST398, ST45, ST30 and ST36 (H cluster of hla tree).

Expression of alpha-hemolysin
The expression of alpha-hemolysin in the 73 strains was assessed
by RT-PCR, in three biological replicates. Fifteen of the 73 strains
(20.5%) were excluded from the final analysis, either because a
single valid determination for Delta Ct (N = 2) was obtained or
because Cq obtained from the different biological replicates were
not reproducible (N = 13).

The analysis of the regression tree split the response variable
into two distinct groups, according to the spa type of the strains.
There was a group of strains with mean Delta Cq ≤5.73, that
was classified as a high expression group and a second group
with a mean Delta Cq >5.73 classified as a low expression group
(Table 1, Table 2 and Figure 3). Overall the regression tree
explained 60% of the variance in the data. This is mostly because
there were strains expressing a low or high mean Delta Cq that
were classified in the same spa type; those were the cases of spa
types t002, t019, t044 and t437.

Furthermore, we explored in each of the spa types what other
explanatory variables (ST, MRSA, MSSA and type of SCCmec)
could differentiate the inclusion of some strains in the low or high
expression group, but we found no associations with the variables
we measured in the study.
We observed that the hla expression level varied within strains
of the same ST (Figure 3; Table 1 and 2). In fact, in some cases the
same ST comprised strains with both high and low levels of
expression (ST5, ST15, ST22, ST30, ST59, ST72 and ST80).
Moreover, we found that the expression rates did not differ
significantly (P = 0.665) between MRSA and MSSA strains.
However, we did find a correlation between the
expression (ST5, ST15, ST22, ST30, ST59, ST72 and ST80).

Discussion
Although Hla is one of the most important S. aureus virulence
factors [10], to the best of our knowledge, this is the first study in
which the variation in hla nucleotide sequence and gene expression
was assessed in such a large and representative
collection.
We found that the nucleotide sequence of hla was highly
diverse. The high degree of diversity found within hla is in
accordance to results obtained for other exotoxins, which are
generally highly polymorphic [27]. Four non-synonymous substitu-
tions (Q87 stop codon, D208E, S239T and T243S) were
identified, that are located in two structural protein domains which
are essential for Hla oligomerization and pore formation (Rim and
Cap) [11,28,29]. The impact of these amino acid (aa) changes on
hla activity is uncertain. If by one hand, the aa changes described
implicate differences in the molecular weight of the aa, that can
have influence in the three dimensional structure stability
and activity of the protein; on the other hand these aa changes did
did not match any of the aa previously described to be essential for Hla
pore formation.

Furthermore, Walker and Bayley showed that multiple mutations
in this same region (residues spanning Hla235–250) did not
alter Hla activity in terms of binding, oligomerization or lysis.
Thus, it would not be expected that S239T or T243S had
significant biological impact in terms of toxin function. The
unique mutation with an identified role in Hla function is the stop
codon found in the ST36 and ST30 strains that was previously described
by DeLeo and co-authors [30] to hinder toxin production and to
originate a less virulent strain in a murine infection model. The
total absence of non-synonymous substitutions identified in our
study in the activity of the protein would have to be tested by the
construction of site directed mutagenesis mutants and by
performing binding, oligomerization, hemolysis and in vitro
models assays.
The construction of phylogenetic trees from the hla defined
the existence of two clusters with different levels of genetic diversity
suggesting that hla is evolving at different rates in different genetic
backgrounds. Interestingly, the most diverse cluster included the
clonal types which are presently more disseminated or that
emerged recently (like ST398). This might be related to the fact
that these clones still need to evolve to evade the human immune
system and not enough time as elapsed for the most adapted allele
to have been selected [31]. On the other hand the recombination
events detected in the hla gene in this study were all in strains
belonging to the high genetic diversity cluster, suggesting that this
mechanism might have been important in the most recent hla
evolution and diversification.

Interestingly, the phylogenetic tree constructed from the hla
gene was similar to that constructed from MLST genes, in the
sense that both trees distributed the different STs similarly in two
main clusters. This observation suggests that hla gene has evolved
together with the S. aureus genetic background. A similar type of correlation with the genetic background was previously described for adhesins, either located in the core genome (clfA, clfB, fnbA, map, sdrC, and spa) or accessory genome (ebpS, fnbB, sdrD, and sdrE) [32]. Although this was the case for the great majority of STs, we observed that four STs (ST22, ST121, ST59, ST93) were located in different clusters in the hla and MLST trees. Our results suggest that recombination occurring at the hla level, might explain the different clustering of strains belonging to ST22 and ST121. No recombination events were, however, detected in

Figure 3. HA and CA strains relative expression distribution. Mean of expression rates from three biological replicates. Dashed line corresponding to the mean Ct value 5.73 results from the regression tree analysis which split strains in two distinct groups, at spa type level: a) high expression group - corresponding to strains with Mean Delta Ct ≤ 5.73 and b) low expression group- corresponding to strains with Mean Delta Ct > 5.73). Highlighted in red are the high expressing strains.
doi:10.1371/journal.pone.0098634.g003

Figure 4. Distribution of the relative hla expression. Mean of relative expression of three independent readings. Expression comparison between a) MRSA and MSSA and b) HA and CA backgrounds using the Two-tailed Student’s t-test. Statistically significance (p ≤ 0.05). (**).
doi:10.1371/journal.pone.0098634.g004
MLST genes or hla sequences of strains belonging to ST39 and ST93, suggesting that their displacement in the two trees could derive from different phenomena, like random mutation.

It was previously suggested that CA-MRSA expressed more hla than HA-MRSA [9]. Results from our study allowed us to extend this conclusion to virtually all epidemic CA, but also in two particular cases of HA genetic backgrounds. The CA strains belonging to ST398, ST25, ST121 and ST39 showed uniformly high relative expression rates and strains belonging to ST36, ST45, ST239, ST247 and ST250 showed uniformly low expression rates. To understand if in fact these patterns of expression are characteristic of these clones, more strains within each clone should be studied for hla expression. Nevertheless, we could not correlate the hla expression rate with any particular polymorphism within the promoter or any aa substitution in the hla gene. The results suggest that hla regulation is probably a result of combination of factors which are redundant, rather than associated to a single genetic event. In fact, it has been demonstrated by several authors that alpha-hemolysin is part of a complex regulatory network, that includes the main two-component systems (TCS) – Agr – that in turn is controlled by a diverse pool of regulatory networks that coordinately interact in response to external stimuli and cell signals, namely others TCS (SaeRS, ArlRS and SrrAB), alternative sigma factors (σ^A), and transcription factors (e.g. SarS, SarT, Rot, SarA, SarZ) [33,34].

We showed that hla evolved together with the genetic background. Moreover, the most epidemic CA-MRSA genetic backgrounds express more hla than the most epidemic HA-MRSA genetic backgrounds. However, the finding of frequent strain-to-strain variation in the expression level of hla within strains of the same clonal types suggests that hla polymorphisms cannot be used as genetic markers of virulence and investigators should remain cautious when inferring conclusions for the entire MRSA population from studies performed with a limited number of strains.

References

1. Deurenberg RH, Stobberingh EE (2009) The molecular evolution of hospital- and community-associated methicillin-resistant *Staphylococcus aureus*. Curr Mol Med 9: 100–115.
2. David MZ, Daum RS (2010) Community-associated methicillin-resistant *Staphylococcus aureus*: epidemiology and clinical consequences of an emerging epidemic. Clin Microbiol Rev 23: 616–687.
3. Otto M (2013) Community-associated MRSA: What makes them special? Int J Med Microbiol.
4. Kobayashi SD, Vojich J, Burlak C, DeLeo FR (2005) Neutrophils in the innate immune response. Arch Immunol Ther Exp (Warsz) 53: 505–517.
5. Baba T, Takeuchi F, Kuroda M, Yuzawa H, Aoki K, et al. (2002) Genome and regulation of gene expression by ArlRS, a two-component signal transduction system in Staphylococcus aureus. EMBO J 21: 1819–1827.
6. Diep BA, Gill SR, Chang RF, Phan TH, Chen JH, et al. (2006) Complete genome sequence of USA300, an epidemic clone of community-acquired methicillin-resistant *Staphylococcus aureus*. Lancet 367: 731–739.
7. Burlak C, Hammer CH, Robinson MA, Whitney AR, McGavin MJ, et al. (2007) Global analysis of community-associated methicillin-resistant *Staphylococcus aureus* exoproteins reveals molecules produced in vitro and during infection. Cell Microbiol 9: 1172–1190.
8. Loughman JA, Fritz SA, Storch GA, Hunstad DA (2009) Virulence gene expression in human community-acquired *Staphylococcus aureus* infection. J Infect Dis 199: 294–301.
9. Li M, Cheung GY, Hu J, Wang D, Joo HS, et al. (2010) Comparative analysis of virulence and toxin expression of global community-associated methicillin-resistant *Staphylococcus aureus* strains. J Infect Dis 202: 1686–1676.
10. Bermejo RJ, Bubek Wardenburg J (2013) *Staphylococcus aureus* alpha-Toxin: Nearly a Century of Intrigue. Toxins (Basel) 5: 1140–1166.
11. Song L, Hobaugh MR, Shustak C, Cheley S, Bayley H, et al. (1996) Structure of staphylococcal alpha-hemolysin, a hemopore transmembrane pore. Science 274: 1859–1866.
12. Valeva A, Palmer M, Bhakdi S (1997) Staphylococcal alpha-toxin: formation of the hemopore is partially cooperative and proceeds through multiple intermediate stages. Biochemistry 36: 13298–13304.
13. Novick RP, Ross HE, Projan SJ, Kornblum J, Kreiswirth B, et al. (1993) Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. EMBO J 12: 3967–3975.
14. Ballal A, Ray B, Manna AC (2009) sarZ, a sarC family gene, is transcriptionally activated by MgaC and is involved in the regulation of genes encoding exoproteins in *Staphylococcus aureus*. J Bacteriol 191: 1656–1663.
15. Liang X, Yu C, Sun J, Liu H, Landwehr C, et al. (2006) Inactivation of a two-component signal transduction system, SaeRS, eliminates adherence and attenuates virulence of *Staphylococcus aureus*. Infect Immun 74: 4655–4665.
16. Liang X, Zheng L, Landwehr C, Lumsford D, Holmes D, et al. (2005) Global regulation of gene expression by ArlRS, a two-component signal transduction regulatory system of *Staphylococcus aureus*. J Bacteriol 187: 5486–5492.
17. Schmidt KA, Manna AC, Gill S, Cheung AL (2001) SarT, a repressor of alpha-hemolysin in *Staphylococcus aureus*, is required for hla but not sarZ. J Bacteriol 183: 4749–4758.
18. Liang X, Hall JW, Yang J, Yan M, Doll K, et al. (2011) Identification of single nucleotide polymorphisms associated with hyperproduction of alpha-toxin in *Staphylococcus aureus*. PLoS One 6: e16428.
19. Bubek Wardenburg J, Bae T, Otto M, DeLeo FR, Schneewind O (2007) Poring over pores: alpha-hemolysin and Panton-Valentine leukocidin in *Staphylococcus aureus* pneumonia. Nat Med 13: 1405–1406.
20. Bubek Wardenburg J, Patel RJ, Schneewind O (2007) Surface proteins and exotoxins are required for the pathogenesis of *Staphylococcus aureus* pneumonia. Infect Immun 75: 1040–1044.
21. Inoashima I, Inoashima N, Wilke GA, Powers ME, Frank KM, et al. (2011) A *Staphylococcus aureus* pore-forming toxin subverts the activity of ADAM10 to cause lethal infection in mice. Nat Med 17: 1310–1314.
22. Tamura K, Nii M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci U S A 101: 11030–11035.
23. Vandenesch F, Kornblum J, Novick RP (1991) A temporal signal, independent of age, is required for hla but not spa transcription in *Staphylococcus aureus*. J Bacteriol 173: 6313–6320.
24. Zielsinska AK, Beeken KE, Joo HS, Mrak LN, Griffin LM, et al. (2011) Defining the strain-dependent impact of the Staphylococcal accessory regulator.
