Ajayi et al. BMC Microbiology (2018) 18:34
https://doi.org/10.1186/s12866-018-1179-7

BMC Microbiology

RESEARCH Open Access

Genetic variability in the *sdrD* gene in *Staphylococcus aureus* from healthy nasal carriers

Clement Ajayi*, Espen Åberg, Fatemeh Askarian, Johanna U. E. Sollid, Mona Johannessen and Anne-Merethe Hanssen*

Abstract

Background: *Staphylococcus aureus* cell wall anchored Serine Aspartate repeat containing protein D (SdrD) is a member of the microbial surface component recognising adhesive matrix molecules (MSCRAMMs). It is involved in the bacterial adhesion and virulence. However the extent of genetic variation in *S. aureus sdrD* gene within isolates from healthy carriers are not known. The aim of this study was to evaluate allelic variation of the *sdrD* gene among *S. aureus* from healthy nasal carriers.

Results: The *sdrD* A region from 48 *S. aureus* isolates from healthy carriers were analysed and classified into seven variants. Variations in the *sdrD* A region were concentrated in the N2 and N3 subdomains. Sequence analysis of the entire *sdrD* gene of representative isolates revealed variations in the SD repeat and the EF motifs of the B repeat. In silico structural modelling indicates that there are no differences in the SdrD structure of the 7 variants. Variable amino acid residues mapped onto the 3D structure revealed that the variations are surface located, exist within the groove between the N2-N3 subdomains and distributed mainly on the N3 subdomain. Comparison of adhesion to keratinocytes in an in vitro cell adhesion assay, using NCTC 8325–4Δ*sdrD* strains expressing the various *sdrD* gene variants, indicated a significant difference between only two complements while others showed no major difference in their adhesion.

Conclusions: This study provides evidence of sequence variations across the different domains of SdrD from *S. aureus* isolated from healthy nasal carriers. Proper understanding of these variations is necessary in the study of *S. aureus* pathogenesis.

Keywords: *S. aureus*, Sdr, Healthy carrier, Virulence factors, Allelic variants

Background

Staphylococcus aureus is an opportunistic human microbe, often found in the anterior nares of one out of four healthy adult individuals [1, 2]. It is responsible for a wide range of human diseases including skin infections such as folliculitis, impetigo and more severe diseases such as endocarditis, and septicaemia [3]. Despite antibiotics use and improved health care, *S. aureus* still remains one of the major causes of hospital-related infections. [4]

Critical in *S. aureus* pathogenesis is its adherence to host cells and/or components present in the host's extracellular matrix [5]. *S. aureus* expresses an array of virulence factors such as the Microbial Surface Component Recognising Adhesive Matrix Molecules (MSCRAMMs), which facilitate its successful adherence [6]. These proteins include surface proteins such as clumping factor (Clf) A, ClfB, serine-aspartate repeat containing protein C (SdrC), SdrD and SdrE. They share similar structural organisation consisting of an N-terminal secretory signal peptide, followed by an A domain, B repeat, and R region containing serine-aspartate repeats. The C- terminal consists of an LPXTG cell wall-anchoring motif, hydrophobic membrane-spanning region and a charged cytoplasmic tail [6] (Fig. 1).

* Correspondence: clement.ajayi@uit.no; anne-merethe.hanssen@uit.no
Research group of Host-Microbe Interactions, Department of Medical Biology, Faculty of Health Sciences, UiT-The Arctic University of Norway, 9037 Tromsø, Norway

© The Author(s). 2018 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

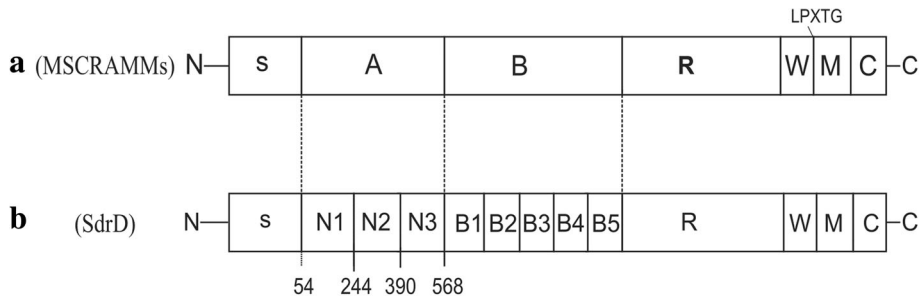


Fig. 1 Schematic representation of MSCRAMMS and the SdrD of *S. aureus* NCTC8325 (**a**) MSCRAMMs have a common structural domain organisation consisting of a signal sequence (S) followed by a ligand binding domain (A), B repeat region (B), Serine-aspartic acid repeat region (R), wall-spanning region (W), LPXTG motif; transmembrane region (M); cytoplasmic domain (C). (**b**) Representation of full length SdrD protein from *S. aureus* NCTC8325. Details of the SdrD A subdomains N1 (54–243), N2 (244–390) and N3 (319–568) are indicated. The B repeat region consists of five repeats (B1-B5)

These proteins interact with host molecules such as fibrinogen [7], desmoglein 1 (Dsg 1) [8] and β-neurexin [9]. Overall, they contribute to *S. aureus* virulence, immune evasion and survival within the host [6].

S. aureus SdrD is encoded within the *sdr* locus in a tandem arrangement with *sdrC* and/or *sdrE* genes [10]. A previous study indicated that while *sdrC* was almost always present in all the *S. aureus* strains studied, *sdrD* had a prevalence of 60% [11]. The different regions within the SdrD protein structure (Fig. 1) contribute to its overall function. The signal peptide directs its translocation to specific sites within the peptidoglycan of the bacterial cell wall [12] The A region is considered to be the ligand binding region and is further divided into three subdomains N1, N2 and N3 [6, 13]. The N2 and N3 subdomains bind the ligand via a "dock-lock-latch" mechanism [13]. In this process, the ligand docks within the groove between these domains, the N3 subdomain locks the ligand in place and provides an additional structure that latches onto the N2 subdomain, further locking the ligand in place. The SdrD B repeat is made up of B1-B5 subdomains of 110–113 amino acid residues, which contains an EF motif that interacts with calcium ions [14]. It is proposed that the B repeat functions as a spacer influencing the distance between the ligand binding A region and the bacterial cell surface [15].

Several studies have delineated the function and expression profile of the SdrD protein, further emphasising its importance in *S. aureus* virulence. It is expressed during nasal colonisation [16] and increase bacterial adherence to desquamated nasal cells [17]. Moreover, SdrD promotes the adhesion of *S. aureus* to keratinocytes via its interaction with Dsg1 [8]. *sdrD* expression is increased in the presence of blood [18] and it was recently shown that the protein also promotes bacterial survival and virulence during systemic infection [19]. Moreover, a correlation between the presence of the *sdrD* gene and bone infections has been observed [11, 20].

Genetic variation within genes encoding virulence factors may influence bacterial pathogenesis. Polymorphisms in the A domain of FnBP A and B in *S. aureus* have been reported [21, 22] and specific single amino acid polymorphisms in FnBP A is associated with infection of cardiovascular devices [23]. Xue et al. [24] observed polymorphisms in the *sdrD* gene from clinical or sub-clinical bovine mastitis associated *S. aureus* isolates. Furthermore, genetic variations within *S. aureus* *sdrD* sequences obtained from GenBank have also been reported [25]. Sources of the *S. aureus* isolates used in the aforementioned studies were either from animal or diseased humans. However, genetic variability in the *sdrD* gene in *S. aureus* isolates from a healthy human population has not been reported. The aim of this study was to determine sequence variability in the *sdrD* gene among *S. aureus* isolated from anterior nares of healthy adult carriers. Furthermore, we aimed to further characterise these variations within the structural domains of the SdrD proteins and evaluate the possible effects of these variations on bacterial adhesion.

Results

Prevalence of genes encoding *sdrC, sdrD* and *sdrE* in *S. aureus*

The prevalence of the three *sdr* genes in 554 *S. aureus* isolates is shown in Table 1. The *sdrC* gene was present in 99% of the isolates investigated while *sdrD* and *sdrE* genes were present in 29% and 84% of the isolates, respectively. Furthermore, the combination of the genes varied across the isolates. Among the 554 isolates, four isolates were negative for the three genes, 146 isolates carried all three, 328 isolates carried two and 76 isolates carried only one. None of the isolates carried the *sdrD* gene as a single gene.

The first 51 *sdrD* positive isolates were selected for further analysis. The selected isolates were distributed into eight *spa* clonal complexes (*spa* CC) (CC2, CC5, CC24, CC65, CC78, CC84, CC267, CC153) (Table 2). In

Table 1 Prevalence of *sdr* genes within *S. aureus* isolates from healthy carriers ($n = 554$)

Gene	*sdrC*	*sdrD*	*sdrE*	Number of isolates (n)
Gene combination	+	+	+	146
	+	+	–	11
	+	–	+	316
	–	+	+	1
	+	–	–	74
	–	+	–	0
	–	–	+	2
	–	–	–	4
Total Number of positive isolates n (%)	547 (99%)	158 (29%)	465 (84%)	554

addition, there were two singletons that could not be clustered. The number of isolates within each *spa* CC varied but the majority of the isolates were found in CC84.

Sequence alignment and phylogenetic analyses of the *sdrD* A-domain

To investigate the sequence diversity of the *sdrD* gene, a region of approximately 1500 bp covering the *sdrD* A region was sequenced from the 51 *sdrD* positive *S. aureus* isolates. Three of the isolates did not yield good sequences and therefore they were excluded from further analysis. Nucleotide sequences and the predicted amino acid sequences from the remaining 48 *S. aureus* isolates were aligned with the corresponding sequences from six *S. aureus* genomic sequences (NCTC8325, N3I5, MW2, MSSA476, Newman and HO 5096 0412).

The alignment revealed a considerable diversity across the sequences with variations in form of nucleotide transitions, transversions, deletions and insertions (Additional file 1). However, similarities were also found. The *sdrD* A region in isolates 12, 14 and 15 were identical to that in N315, while the *sdrD* A-region in isolates 13 and 36 showed 100% identity with *sdrD* in NCTC8325.

Figure 2 shows a maximum likelihood tree based on the deduced amino acid sequences of the *sdrD* A region of the 48 queried and 6 reference strains. The phylogenetic tree split into two major groups, indicated as group 1 and 2. Group 1 further separated into two sub-groups, indicated as A and B, these further divided into smaller clusters. Based on the tree topology, isolates 43 and 49 showed the highest divergence for group 1. The reference sequences included in our analysis were distributed across all clusters. Furthermore, the strains also clustered according to their *spa* clonal complexes. The phylogenetic analysis based on the *sdrD* nucleotide sequences resulted in a similar pattern of clustering and divergence (data not shown).

Table 2 Characterization of the *sdrD* positive *S. aureus* strains isolated from anterior nares of healthy human analyzed in this study

spa-Clonal Complexes (CC)	Number of isolates	Strain ID no	*spa* type*
CC2	7	15,19	t002
		14	t045
		26	t5423
		51	t548
		17	t581
		12	t601
CC5	4	46,16	t005
		3	t223
		44	t309
CC24	7	20,25	t008
		35,36	t024
		13	t1476
		4	t305
CC65	1	43	t209
CC78	10	42	t078
		10,22	t1102
		6	t167
		39	t280
		7	t349
		27	t5257
		2	t5449
		5	t759
		30	t814
CC84	16	1,11,23,28,29,32, 33,34,37,41,46,48,50	t084
		38	t094
		31	t2219
		18	t5232
CC153	1	40	t153
CC267	1	24	t4173
Singleton	2	9	t3160
		49	t186
N/A	2	21	t258
		45	N/A

* Results from *spa* typing are from a previous study [36]

Based on the observed diversity in the A domain of the *sdrD* gene, we classified the isolates into different variants (Fig. 2). For each *sdrD* variant, a representative isolate was randomly selected. *S. aureus* NCTC8325–4 was selected as the representative strain of the sdrD variant 7. Amino acid identity within this region between the seven *sdrD* variants was between 85% and

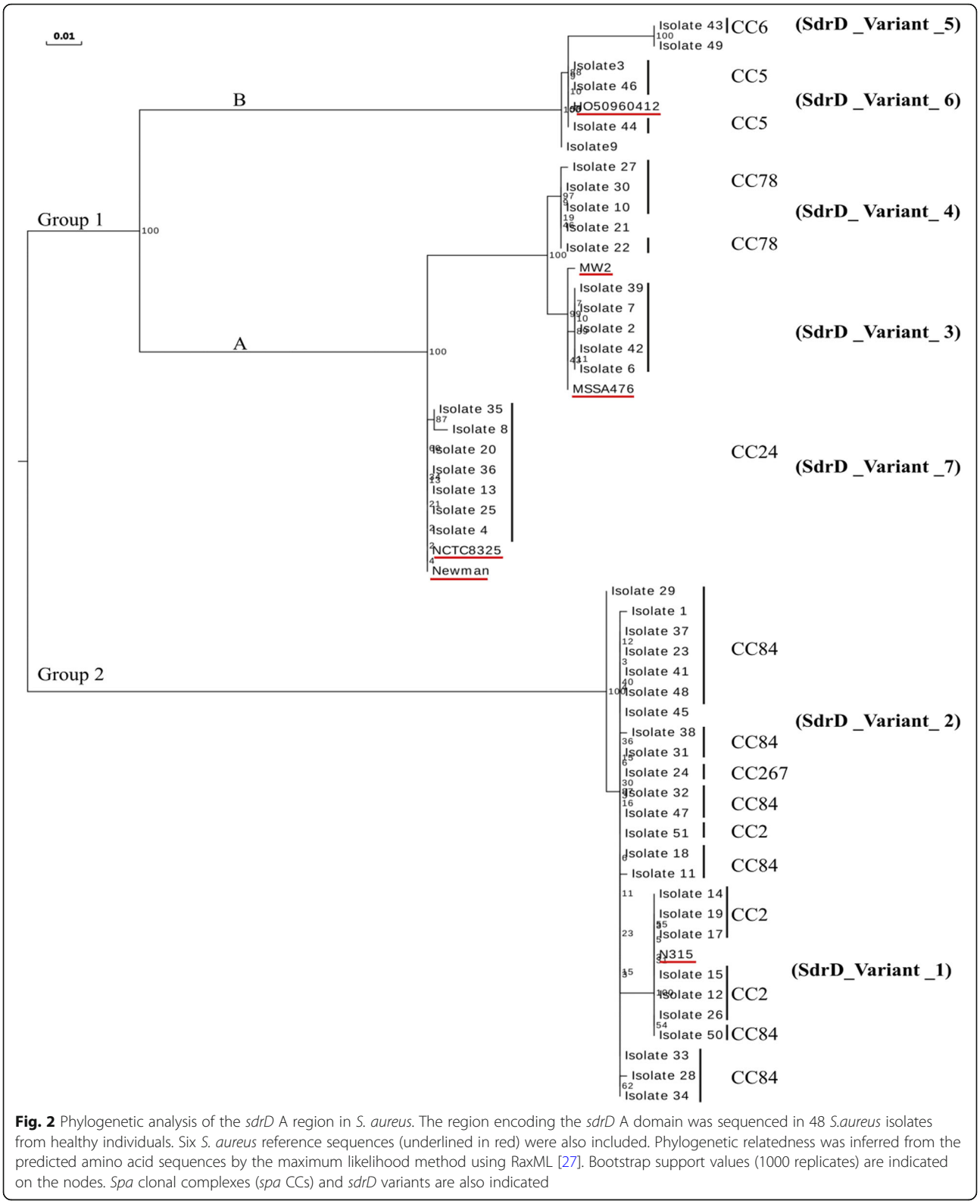


Fig. 2 Phylogenetic analysis of the *sdrD* A region in *S. aureus*. The region encoding the *sdrD* A domain was sequenced in 48 *S.aureus* isolates from healthy individuals. Six *S. aureus* reference sequences (underlined in red) were also included. Phylogenetic relatedness was inferred from the predicted amino acid sequences by the maximum likelihood method using RaxML [27]. Bootstrap support values (1000 replicates) are indicated on the nodes. *Spa* clonal complexes (*spa* CCs) and *sdrD* variants are also indicated

99% (Additional file 4: Table S1). These results indicate allelic sequence diversity in the A region of the *sdrD* gene within these *S. aureus* isolates.

Sequence variation within the entire *sdrD* gene

Whole genome sequencing was performed on the representative isolate of each *sdrD* variant, including *S. aureus*

NCTC8325–4 as it had been chosen as representative strain of variant 7. Sequences of the *sdrD* gene were extracted from all the seven and compared. A pairwise comparison of the deduced amino acid sequences showed identity ranging from 93.1% - 99.3%, where *sdrD* variants 3, 4 and 7 were most similar to each other. The size of the SdrD polypeptides ranged from 1315 to 1399 amino acids (Table 3). We found amino acid variations in the consensus sequence of the EF motifs responsible for calcium binding [14]. For instance, in s*drD* variant 6, there were changes in amino acids residues at position 585 (Val to Iso), 591 (Val to Lys), and 594 (Gly to Lys). There were also extensive insertions or deletions within the R-domain, consisting of serine aspartate repeats. Finally, the N2-N3 domain of the A-region varied extensively, which are in agreement with results shown in Fig. 2. In contrast, the N-terminal signal peptide, YSIRK/GS motif, the signal peptide cleavage sites [12, 26], the B repeats and the residues of the hydrophobic tail were highly conserved among the seven SdrD variants (Additional file 2).

Phylogenetic analysis of the deduced amino acid sequences for the entire *sdrD* gene based on the maximum likelihood method also indicated *sdrD* sequence diversity in the representative isolates (Fig. 3). The phylogenetic tree divided into three major clades. These results further affirm the genetic variability within the *sdrD* gene.

Multilocus sequence type (MLST) analysis of *sdrD* variants

We determined the MLST of the representative *S. aureus* isolate for each *sdrD* variant and performed a phylogenetic analysis based on this. Figure 4 shows a maximum likelihood tree based on the concatenated sequences of the seven housekeeping genes for these isolates. The isolates were distributed into seven different sequence types (ST) and the phylogeny tree split into two groups (Fig. 4). The variants were distributed randomly within the groups. The MLST phylogeny had a different topology compared to the *sdrD* phylogeny shown in Fig. 3. For instance, in the MLST tree, *S. aureus sdrD* variants 2 (ST15) and 5 (ST109) formed a clade, but clustered into different groups in the *sdrD* phylogeny. In addition, *sdrD* variants 6 and 5 were more closely related in the *sdrD* phylogeny but clustered differently in the MLST tree.

3D structural modelling and characterization of SdrD variants

Based on the suggested crystal structure of the N2-N3-B1 subdomains of *S. aureus* SdrD [13], 3D models of N2-N3-B1 subdomains for the seven SdrD variants were generated using the Swiss model server. Comparisons of the generated 3D structure indicated a very high structural similarity with RMSD root mean square deviation (RMSD) values ranging from 0.008–0.059 Å (data not shown). Most of the amino acid variations were located on the N3 subdomain (Fig. 5a). In addition, some of the varied amino acid residues were located within and around the groove between the N2-N3 subdomains (Fig. 5b). Furthermore, the majority of the amino acid variations were surface associated (Fig. 5c).

Adhesion of SdrD variants to keratinocytes

SdrD have previously been shown to promote *S. aureus* adhesion to human keratinocytes [8], and we next wanted to evaluate whether the allelic variation in SdrD could influence this. In order to have identical genetic background, the *sdrD* gene variants were cloned into pMG36e. However, we encountered technical challenges in cloning the *sdrD* gene for *sdrD* variant 1, therefore we could not investigate the effect of this variant on adhesion to keratinocytes. The cloned plasmid was used to complement NCTC8325*ΔsdrD* and all had similar growth rate (Additional file 3a). The plasmids were confirmed to contain *sdrD* by direct sequencing. Complemented NCTC8325*ΔsdrD* that were confirmed to express SdrD (Additional file 3b) were used for adhesion to mammalian cells. However, we could not detect SdrD expression in variants 3 and 4, though direct sequencing and PCR on plasmid isolated from the complements, indicated the presence of the *sdrD* sequence (results not shown). Therefore, we continued our assay with *sdrD* variants 2, 5,6 and 7.

Table 3 Percentage Amino Acid Identity using Amino acid for the entire *sdrD* gene for the seven representative *sdrD* variants

	(aa)	SdrD Variant 1	SdrD Variant 2	SdrD Variant 3	SdrD Variant 4	SdrD Variant 5	SdrD Variant 6	SdrD Variant 7
SdrD Variant 1	1365	100	99.3	93.9	93.7	94.5	93.1	94.5
SdrD Variant 2	1399	99.3	100	94.4	93.9	94.8	93.5	94.8
SdrD Variant 3	1329	93.9	94.4	100	98.7	95.3	94.5	98.5
SdrD Variant 4	1365	93.7	93.9	98.7	100	94.7	93.9	98.2
SdrD Variant 5	1315	94.5	94.8	95.3	94.7	100	97.6	95.6
SdrD Variant 6	1375	93.1	93.5	94.5	93.9	97.6	100	94.9
SdrD Variant 7	1380	94.5	94.8	98.5	98.2	95.6	94.9	100.0

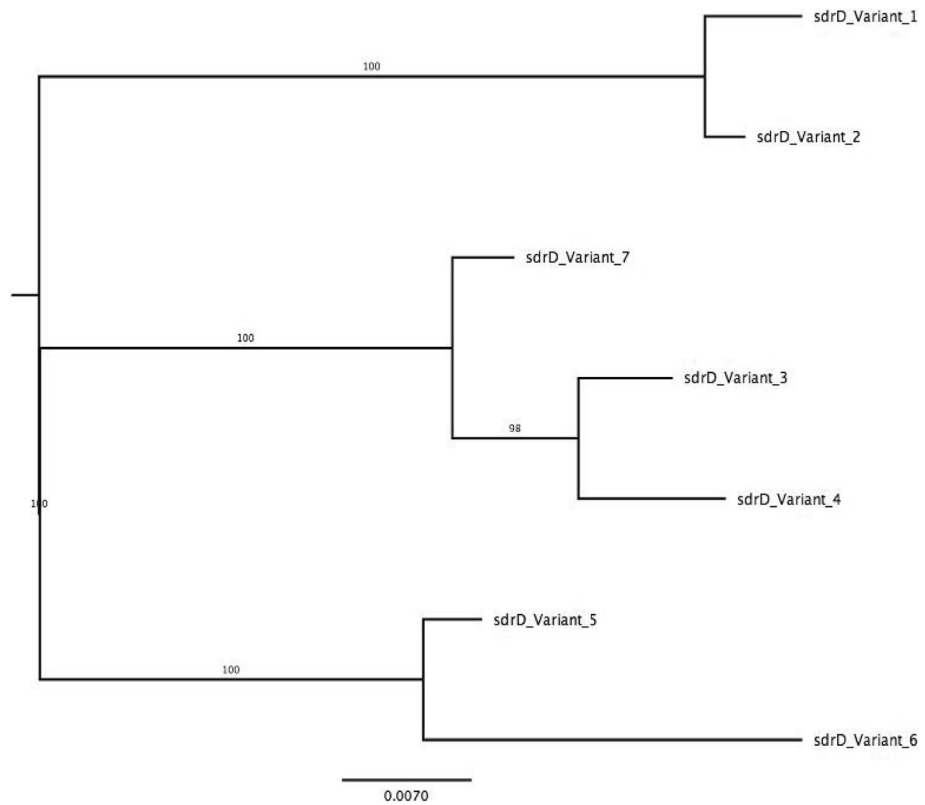


Fig. 3 Phylogenetic relationship between the *sdrD* variants. Phylogenetic tree based on the deduced amino acid sequences for the entire *sdrD* gene from seven representative *S. aureus* strains. The tree was generated by the maximum likelihood method using RaxML [27]. The bootstrap support values (1000 replicates) are indicated on the branches. Branch length is proportional to the number of substitutions as indicated by the scale bar

S. aureus ΔsdrD complemented with *sdrD* variants were co-incubated with HaCaT cells for 90 mins. After washing, the adhered bacteria were harvested and quantified via serial diluting and subsequent plating on blood agar. All the tested complemented *S. aureus* strains showed significant adhesion to HaCaT cells compared to the control strain lacking *sdrD* (Fig. 6). In addition, we also observed a significant difference in the adhesion of variant 5 compared to variant 7.

Discussion

S. aureus is an opportunistic microbe which frequently colonizes humans using its array of virulence factors. We investigated the prevalence of the *sdr* genes among *S. aureus* isolates obtained from anterior nares of a healthy adult population. Most importantly, we report genetic variability in the *sdrD* gene in *S. aureus* isolates from healthy adults.

Sequence diversity within virulence genes in *S. aureus* have been reported; genes such as *coa* [25], *spa* [25], *agr* [27], and *fnb* [21] were reported to be polymorphic in *S. aureus* isolates. Moreover, variation in sequences encoding surface proteins expressed by *S. aureus*, are thought to be essential in its survival to environmental changes [28]. The genetic variability within the *sdrD* gene sequences varied across its defined structural domains. While the N2-N3 sub domain of the A-domain, B1 subdomain of its B-repeats and SD repeats were highly diverse, the other regions showed a high degree of conservation. The findings are in agreement with previous studies, which report variations in the *sdrD* gene in *S. aureus* isolates from livestock [24] and sequences obtained from the GenBank database [25].

Analysis of the entire *sdrD* gene from whole genome sequencing of seven selected representative isolates also indicated regions of insertion and deletion within the R-domain, a region made of tandem repeats of serine- aspartate (SD) dipeptide [10]. The SD residues of *S. aureus* become glycosylated resulting in resistance to proteolytic cleavage in blood [29] and in pathogenesis of *S. aureus* bloodstream infections [30]. It is suggested that glycosylation intensity is dependent on the number of SD repeats available [29]. Thus, insertion of additional residues within SD repeats might have an influence on *S. aureus* pathogenicity. However, glycosylation of the repeats also primes the antibodies of the immune system

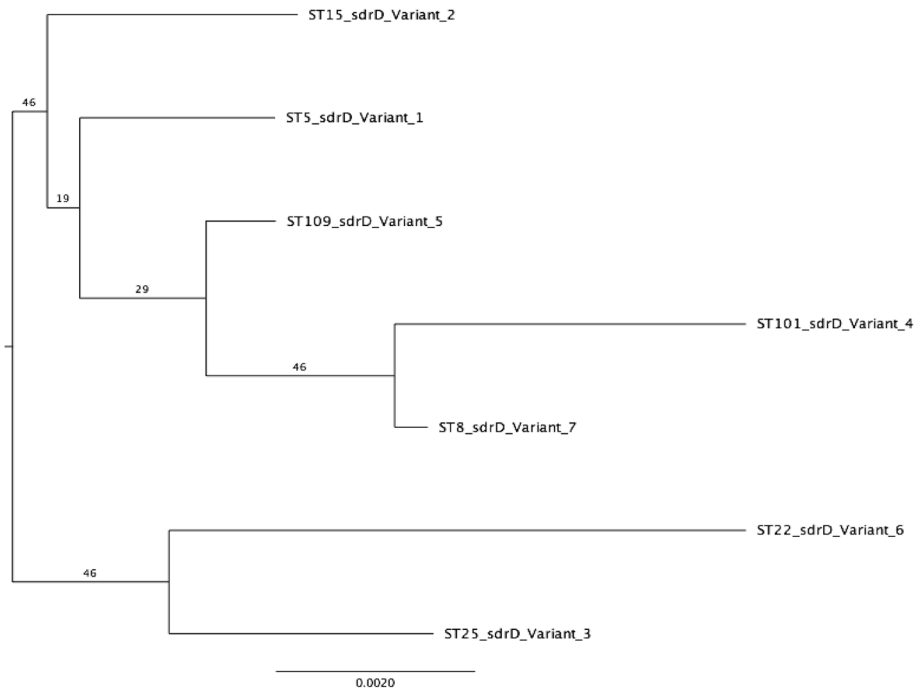


Fig. 4 Maximum likelihood tree based on the concatenated MLST alleles. Sequences types based on the allelic profiles of the MLST alleles in seven selected *S. aureus* strains are indicated in the tree. Bootstrap support values (1000 replicates) are indicated next to the branches. Branches length is proportional to the number of substitutions as indicated by the scale bar. The *sdrD* variants are also indicated

in recognising these proteins [29]. Therefore, it is reasonable that SD residue deletions might be a mechanism employed by *S. aureus* to avoid detection by its host immune surveillance. A previous study has demonstrated that the SD-repeats are hot spot regions with great instability and genetic variation, which can be the result of replication slippage [31]. The genetic variability within this region might suggest the adaptation of *S. aureus* to its environment without losing the functionality of other regions within its surface proteins.

In SdrD, the YSIRK/GS motifs located at the N-terminal signal peptide direct the secretion of SdrD precursor proteins and ensure the ring-like distribution across the *S. aureus* cell wall [12, 32]. In addition, effective anchoring of the SdrD protein by the sortase A to the cell wall is ensured by the LPXTG motif [32]. These consensus sequences were highly conserved among the variants in our study, suggesting a selective pressure that drives the sequence diversity against the functional protein and not the precursor protein. Overall, the B repeat was considerably conserved among the *sdrD* variants. However, there was genetic variability within the sequences of some of the EF motifs. The conserved EF motifs of the SdrD B repeat are involved in calcium binding [14]. Calcium binding to the B repeat has been shown to be important in maintaining its structure [15] while projecting the ligand binding A domain from the bacterial surface [6, 13].

SdrD contributes in *S. aureus* adhesion to human keratinocytes [8] and to desquamated nasal epithelial cells [17]. In concordance with this, we observed that complemented *sdrD* mutants expressing the different SdrD variants exhibited significant adherence to HaCaT cells compared to the isogenic mutant harboring the empty plasmid. The adhesion of the SdrD variants was comparable to each other, except from a significant difference between variants 5 and 7. This could be explained by high diversity in the N2-N3 of the A-region and the B1 subdomain of SdrD. Analysis of these subdomains in the *sdrD* variants showed that there are sequence variations in the groove between the N2 and N3 subdomains and the variations are mostly associated with the N3 subdomain. The groove between the N2-N3 sub domains of the A-domain together with the B1 subdomain is involved in ligand interaction in a dock, latch and lock mechanism [6, 13]. It is possible that the amino acid changes within these regions have resulted in the varied adhesion abilities of the *sdrD* variants to HaCaT cells.

It is tempting to speculate what could account for the sequence variations within the different commensal *S. aureus* isolates. Our MLST phylogeny analysis revealed that *S. aureus* isolates belonging to different clades encoded considerable similar *sdrD* genes. Horizontal gene transfer (HGT), specifically homologous recombination, is suggested to underline the transfer of genes across different *S.*

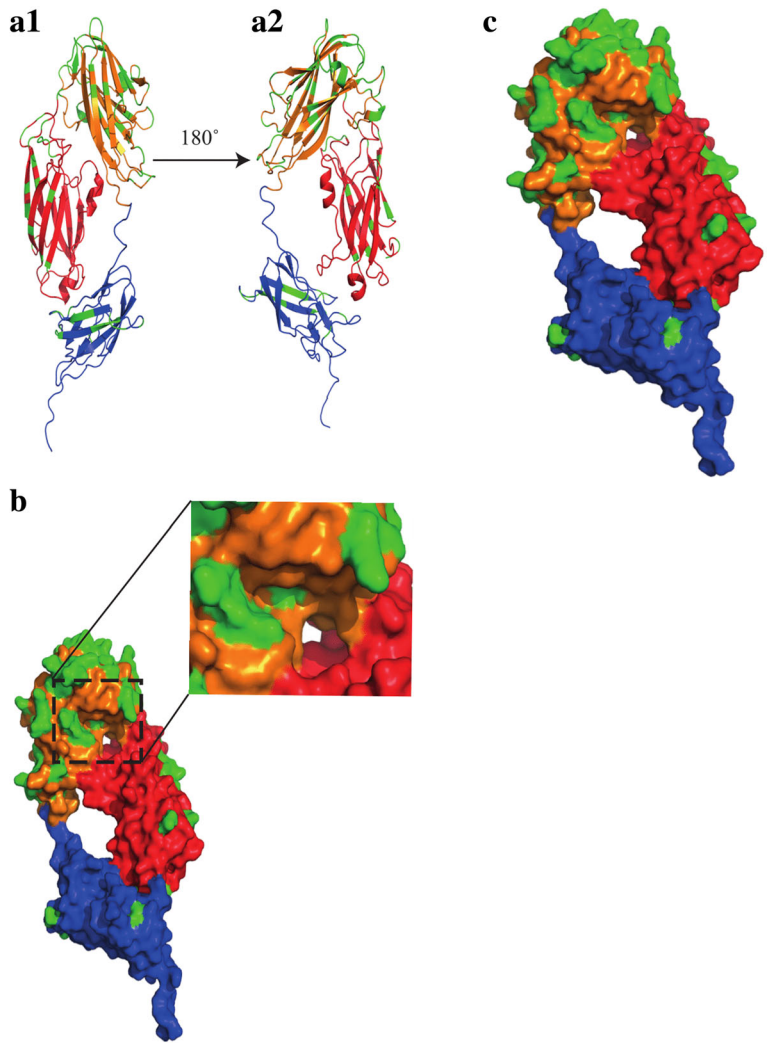


Fig. 5 Structural characterization of the *sdrD* variants. 3D models were generated based on the N2-N3 and B1 subdomains of SdrD [15] in seven representative *S. aureus* strains. The subdomains are shown as N2 (red), N3 (orange) and B1 (Blue). Green colour coding indicates the variable amino acid residues within the seven *sdrD* variants (**a**) AI is ribbon diagram of N2-N3-B1 while A2 shows the diagram when reoriented 180 degrees. (**b**) Space filling diagram of N2-N3-B1. The grove between N2 and N3 is enlarged. (**c**) Space filling diagram of N2-N3-B1. The model shows the location of varied amino acids residues on the surface

aureus isolates [33]. *S. aureus* virulence genes are found on Mobile genetic elements (MGE) and the potential of transfer of these elements between *S. aureus* strains are high [34]. In addition, there have been reports of transfer of non- MGE sequences between *S. aureus* strains [34]. Furthermore, polymorphism in sequences of the genes encoding FnBP A and B have been reported to arise through HGT [35].

Conclusion

In summary, our findings reveal the genetic variability within the *sdrD* gene in *S. aureus* isolates from healthy individuals. Most importantly, our study provides an insight into how the sequences of various domains within the *S. aureus* SdrD protein are conserved. It also indicates the effect the variation might have on bacterial adhesion. However, how this might affect other reported functions of SdrD remains to be investigated.

Materials and methods

Bacterial strains

A total of 4000 nasal culture swabs were taken from the anterior nares of healthy adult participants recruited between 2007 and 2008 in the 6th Tromsø Study. The nasal culture swabs were taken for the purpose of research during the Tromsø Staph and Skin Study, as part of 6th Tromsø Study. Of these, 1113 isolates were confirmed as *S. aureus* and were *spa*-typed previously [36].

From the 1113 *S. aureus* isolates, we selected 554 consecutive isolates and screened for the presence of *sdrC*,

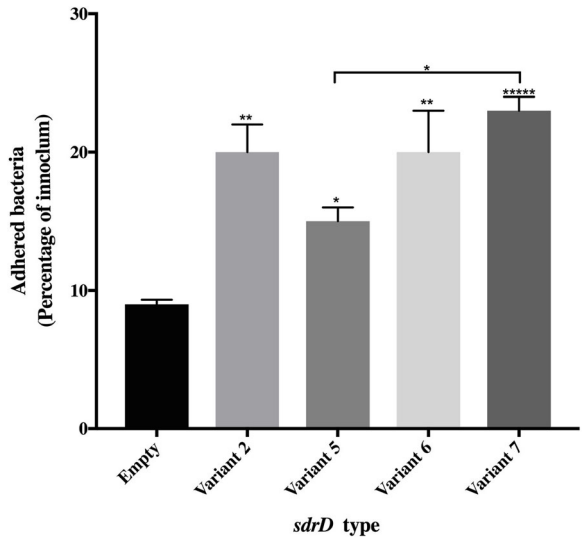


Fig. 6 Adhesion of complemented *S. aureus* mutants to HaCaT cells. Adherence of *S. aureus* isogenic mutant NCTC8325–4Δ*sdrD* complemented with different pMG36e-SdrD variants or pMG36e (empty vector). Adherence was calculated as a percentage of the mean colony-forming unit (cfu/ml) of attached bacteria to the mean cfu/ml of the inoculum. The data represent mean of 3 independent experiments, each with 3 technical repeats. Statistical analysis was performed using one-way ANOVA with Tukey's multiple comparison test (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$)

sdrD and *sdrE* genes by PCR using primers listed in Table 4. The first 51 *sdrD* positive isolates were selected for further analysis, independent of *spa*-types or distribution of *sdr* genes. In addition *S. aureus* NCTC8325–4 and NCTC8325–4Δ*sdrD* [8], was also used in the study.

DNA extraction for PCR

Genomic DNA was extracted from *S aureus* using a boiling method as previously described [37], and stored at − 20 °C until use.

PCR amplification and sequencing

The *sdrC*, *sdrD* or *sdrE* genes in the selected 554 *S. aureus* isolates were amplified by PCR using ReddyMix (Thermo Scientific, USA) according to the manufacturer's instruction. Primers used are listed in Table 4. The cycling conditions for these were: 94 °C for 5 mins; 30 cycles of 94 °C for 10 s, 53 °C for 30 s and 72 °C for 15 s; and 72 °C for 10 mins and thereafter kept at 4 °C. The reference strain *S. aureus* N315 (NC_002745.2) was used as positive control for the amplification of these genes while *Enterococcus faecium* BM4105-RF was used as negative control.

PCR amplification of the *sdrD* A region was also carried out with ReddyMix (Thermo Scientific, USA) according to the manufacturer's instruction. Due to limitation of our DNA sequencer to 1000 bp, amplification of approximately 1500 bp encoding the A domain was performed in two parts using the primer combinations SdrD A-F + SdrD-Rwd and SdrD-Fwd + SdrD-A-R (Table 4) resulting in PCR products of 765 bp and 840 bp, respectively. The PCR products were purified using ExoSAP-IT (Thermo Fisher Scientific, USA) according to the manufacturer's instruction and used as template for direct sequencing using SdrD-A-F and SdrD- A-R primers (Table 4). This was performed using Big Dye Terminator 3.1 cycle sequencing kit (Applied Biosystem, USA) according to the manufacturer's instructions.

Sequence analysis and phylogenetic tree construction

Sequences encoding the *sdrD* A region obtained by PCR described above and *sdrD* A region sequences of six published *S. aureus* genomes obtained from the NCBI GenBank (MW2 (NC_003923.1), Newman (NC_009641.1), N315 (NC_002745.2), NCTC8325 (NC_007795.1), MSSA476 (BX571857.1) and HO 5096 0412 (NC_017763.1) were analysed further. Sequence alignment of the DNA sequences and its predicted amino sequences was performed in Bio Edit sequence alignment editor (version 7.2.5) [38] using the Cluster W algorithm program [39]. Subsequent multiple alignments were also performed using Fast Fourier Transform (MAFFT) (Version 7) [40]. Pairwise sequence alignment of predicted amino acid sequences to calculate amino acid identity was performed using the ExPASy [41].

Phylogenetic trees of the *sdrD* A domain nucleotide sequences and the deduced amino acid sequences were generated using Randomized Axelerated Maximum Likelihood (RaxML) [42]. Bootstrap values were calculated using 1000 replicates to generate confidence in the tree topology. The tree generated was viewed and mid point rooted using Fig tree (version 1.4.3) [43] and Evolview v2 [44].

Whole genome sequencing of *S. aureus*

Genomic DNA from seven representative *S. aureus* isolates was extracted using the GenElute Bacterial Genomic DNA kit (Sigma-Aldrich, US) according to the manufacturers instruction. Samples were submitted to the Norwegian Sequencing Centre, Oslo (www.sequencing.uio.no) for Illumina HiSeq paired-end sequencing. We assessed the quality of the reads using FastQC [45]. The sequence reads were assembled using SPAdes (version 3.6.1) [46] with the –careful option. All genome assemblies were evaluated using the QUAST tool [47]. Protein coding sequences (CDSs) were predicted using MetaGeneMark [48]. The *sdr*-gene sequences were manually isolated using ARTEMIS [49].

Multilocus sequence typing (MLST)

MLST analysis was carried out for the representative *S. aureus* isolates. Nucleotide sequences of the following

Table 4 List of Primers used in this study

Primers	Sequence (5′ - 3′)*	Reference
Prevalence of *sdr* gene		
SdrC prevF	AAAAGGCATGATACCAAATCGA	Sabat et al. [41]
SdrC prev R	AATTCTCCATTCGTATGTTCTG	Sabat et al. [41]
SdrD prev F	AGTGGGAACAGCATCAATTTTA	Sabat et al. [41]
SdrD prev R	GTGGTAGATTGTACACTTTCT	Sabat et al. [41]
SdrE prev F	AGAAAGTATACTGTAGGAACTG	Sabat et al. [41]
SdrE prev R	GATGGTTTTGTAGTTACATCGT	Sabat et al. [41]
Sequencing of SdrD A region		
sdrD A- F	GGAACCAAGAAGCAAAGGCTG	Xue et al. [6]
sdrD A-R	CTTCTTGACCAGCTCCGCCAC	Xue et al. [6]
sdrD Fwd	AGTTGATGACAAAGTTAAATCAGGT	This study
sdrD Rwd	TAATATCTTCCGGATTCAATCCA	This study
Cloning of *sdrD*		
SacI-amh 100 Fwd	AAAAGAGCTCTGAATTAGGAGTAATCTAATGCT	This study
SacI-amh 101 Fwd	AAAAGAGCTCATTTGATGAATTAGGAGTAAT	This study
SphI- amh200R	TTCAGCATGCCGCCTCATATAAGTTTTATTCCGT	This study
SphI- amh201R	ATCAGCATGCAAATTTTGAAATAAAAAACCAGGCC	This study

*restriction sites underlined

house keeping genes: *arcC* (carbamate kinase), *aroE* (shikimate dehydrogenase), *glpF* (glycerol kinase), *gmk* (guanylate kinase), *pta* (phosphate acetyltransferase), *tpi* (triosephosphate isomerase), *yqiL* (acetyl coenzyme A acetyletransferase) [50] were extracted from the genomic sequences of the seven representative *S. aureus* isolates and concatenated. Phylogenetic analysis on these sequences was performed using RaxML [42]. Sequence types for the isolates were obtained from the *S. aureus* MLST database based on the allelic profiles for each representative isolate.

Structural modelling of SdrD

Amino acid sequences of seven representative *sdrD* variants were used in generating a 3D model for their SdrD structure. 3D structures were generated using the Swiss model server [51] based on the resolved crystal structure of the *S. aureus* SdrD N2-N3-B1 subdomains [13]. Further editing and visualization of structures were performed using PyMOL software [52] (Schrodinger, Inc. USA).

sdrD cloning and complementation of *S. aureus* Δ*sdrD*

The open reading frame encoding the entire *sdrD* gene was amplified from the genomic DNA of the seven representative *S. aureus* strains using primers listed in Table 4. The seven amplified PCR products were ligated into the Gram-positive shuttle plasmid pMG36e [53] with erythromycin as the selectable marker (kind gift from Dr. Diep Dzung, Norwegian University of life Science, Norway) creating six pMG36e-*sdrD* plasmids. The constructed plasmids pMG36e-*sdrD* and empty pMG36e were propagated in *E. coli* DH5, under selective pressure (LB supplemented with 400 μg/ml erythromycin).

To overcome the restriction barrier, the plasmids were transformed into and purified from *E. coli* DC10b [54]. The plasmids were electroporated (2.1 kV voltage, 600-Ω resistance, 10 μF capacitance) into *S. aureus* NCTC8325Δ*sdrD* [8] and transformants were selected at 30 °C on TSA supplemented with 8 μg/ml erythromycin. To confirm successful transformation of the isogenic *S. aureus* NCTC8325–4Δ*sdrD,* plasmids were isolated from the recipient strains and confirmed by direct sequencing. Expression of SdrD by the complemented mutant was confirmed by immunoblotting as previously described [8].

Bacterial growth measurement

Complemented *S. aureus* mutants were grown overnight at 30 °C in TSB containing 8 μg/ml erythromycin. Next day, the bacteria were diluted 1:150 into fresh pre-warmed TSB with 8 μg/ml erythromycin. Thereafter, 275 μl of the diluted culture was transferred to wells of Nunclon delta surface 96 well plate (Thermo Scientific, US) in replicates of five. A control containing only the growth medium was included in order to adjust for background fluorescence. The plate was maintained at continuous orbital shaking in a Synergy H1 Reader (BioTek, US). Growth was assayed by measuring optical density at 600 nm at every 15 mins for 24 h.

Mammalian cell line

HaCaT cells, an immortalized human keratinocyte cell line [55], were purchased from CLS Cell Lines Service Germany (item number: 300493). Cells were maintained in Dulbecco's modified Eagle's medium (DMEM) (Sigma Aldrich, Germany), supplemented with 10% (*v*/v) fetal bovine serum (FBS) (Invitrogen Life Technologies, USA), penicillin (100 units/ml), and streptomycin 100 μg/ml (Sigma Aldrich, Germany) and at 37 °C in humidified 5% CO_2. They were routinely sub cultured at 70%–80% confluency.

Bacterial adhesion to HaCaT cells

HaCaT cells were seeded into a 24 well plate at a concentration of 1.5 – 2×10^5 cells/ml in DMEM supplemented with 10% FBS. *S. aureus* NCTC8325Δ*sdrD* complemented with the pMG36e-*sdrD* variants were grown overnight at 30 °C and 230 rpm in TSB with 8 μg/ml erythromycin. Next day, the overnight cultures were diluted into fresh medium and grown to OD_{600nm} of 0.7–0.8. The bacterial cultures were pelleted and washed with PBS. Subsequently, the bacteria were resuspended in DMEM supplemented with 10% FBS and added to the HaCaT cells at multiplicity of infection (MOI) of 100, and incubated at 37 °C for 90 mins. Thereafter, the HaCaT cells were treated, washed and the number of adhered bacteria enumerated as described earlier [8].

Statistical analysis

Statistical analysis was carried out on the pooled data using GraphPad prism 7 (GraphPad Software Inc. USA). Statistical significance was quantified by one-way anova test with Tukey's multiple comparison test. Data with $P < 0.05$ were considered statistically significant.

Additional files

Additional file 1: Multiple Sequence Alignment of nucleotide sequences of SdrD A region for 48 queried and 6 reference strains. Consensus residues within the residues are indicated as conserved (*) and semi-conserved (.). Deletion and insertions are indicated as (–). (PDF 3240 kb)

Additional file 2: Multiple Sequence Alignment of deduced amino acid sequence of *sdrD* gene of the representative sdrD variants. The different domains within the SdrD structure are indicated. The YSIRK/GS and signal cleavage peptide within the signal sequences are indicated in red boxes. The predicted EF motifs within the B repeat and the LPXTG motif is indicated in red boxes also. Alignment was performed with MAFFT and edited using Boxshade and Adobe Photoshop. Consensus residues within the residues are indicated as conserved (*) and semi-conserved (.). Deletion and insertions are indicated as (–). (PDF 10515 kb)

Additional file 3: Growth curve measurement and SdrD Expression **(A)**: Growth curve measurement for the *S. aureus* isogenic mutant NCTC8325–4Δ*sdrD* complemented with either different pMG36e-SdrD variants (variant) or pMG36e (empty vector). Growth was measured as absorbance at OD =600 nm over 24 h. **(B)**. Immunoblot for the detection of SdrD expression by the NCTC8325–4 (WT) NCTC8325–4Δ*sdrD* (Δ*sdrD*), NCTC8325–4Δ*sdrD* complemented with either different pMG36e-SdrD variants (indicated by the number) or pMG36e (Empty). The molecular weight (Kda) is also indicated (PNG 234 kb)

Additional file 4: Table S1 Percentage amino acid identity using the amino acid sequences for the SdrD A domain for the seven representative *sdrD* variants. (PNG 274 kb)

Abbreviations

Dsg1: Desmoglein 1; FnBP: Fibronectin binding protein; HGT: Horizontal gene transfer; MGE: Mobile genetic element; MLST: Multilocus sequence typing; MSCRAMMs: Microbial surface component recognising adhesive matrix molecules; Sdr: Serine-aspartate-repeat protein

Acknowledgements

We thank Dr. Elisabet Josefsson and Dr. Diep Dzung for providing SdrD A-region antibody and pMG36e construct, respectively.

Funding

This work was supported by grants from The Northern Norway Regional Health Authority (Helse Nord RHF) projects SFP877–09, Miljøstøtte MIL963–10 (2010–2012) and Infection Biology and Antimicrobials (IBA) research school (NFR project number 249062). The funding bodies had no role in the design of the study, collection and interpretation of data and in writing the manuscript. The publication charges for this article have been funded by a grant from the publication fund of UiT The Arctic University of Norway.

Availability of data and materials

Fasta files of the sequences encoding the sdrD variants have been deposited in the NCBI's Genbank with accession numbers: sdrD_Variant_7 (MF377380), sdrD_Variant_1 (MF377381), sdrD_Variant_2 (MF377382), sdrD_Variant_3 (MF377383), sdrD_Variant_4 (MF377384), sdrD_Variant_5 (MF377385), sdrD_Variant_6 (MF377386).

Authors' contributions

C.A. performed the experiments and drafted the manuscript. J.U.E, M.J. and A.M.H have taken part in study design and experiments. C.A, M.J, J.U.E and A.M.H analysed and interpreted the data. E. A performed the WGS analysis. F.A provided technical advice regarding the experiment design and manuscript. All authors reviewed and approved the final manuscript.

Ethics approval and consent to participate

The Tromsø Staph and Skin Study and the 6th Tromsø Study was approved by the Regional Committee of Medical Research Ethics, North Norway (Ref. 200,605,174–12/IAY/400), and the Data Inspectorate of Norway (Ref. 07/00886–2/CAO). Written informed consent was obtained from all participants prior to participation.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 5 July 2017 Accepted: 8 April 2018
Published online: 16 April 2018

References

1. Kluytmans J, van Belkum A, Verbrugh H. Nasal carriage of Staphylococcus aureus: epidemiology, underlying mechanisms, and associated risks. Clin Microbiol Rev. 1997;10(3):505–20.
2. van Belkum A. Novel technology to study co-evolution of humans and Staphylococcus aureus: consequences for interpreting the biology of colonisation and infection. Adv Exp Med Biol. 2011;697:273–88.
3. Lowy FD. Staphylococcus aureus Infections. N Engl J Med. 1998;339(8):520–32.
4. Klein E, Smith DL, Laxminarayan R. Hospitalizations and deaths caused by methicillin-resistant Staphylococcus aureus, United States, 1999–2005. Emerg Infect Dis. 2007;13(12):1840–6.

5. Johannessen M, Sollid JE, Hanssen AM. Host- and microbe determinants that may influence the success of S. Aureus colonization. FrontCellInfectMicrobiol 2012; 2(2235–2988):56–56.
6. Foster TJ, Geoghegan JA, Ganesh VK, Höök M. Adhesion, invasion and evasion: the many functions of the surface proteins of Staphylococcus aureus. Nat Rev Microbiol. 2014;12(1):49–62.
7. Keane FM, Loughman A, Valtulina V, Brennan M, Speziale P, Foster TJ. Fibrinogen and elastin bind to the same region within the a domain of fibronectin binding protein a, an MSCRAMM of Staphylococcus aureus. Mol Microbiol. 2007;63(3):711–23.
8. Askarian F, Ajayi C, Hanssen A-M, van Sorge NM, Pettersen I, Diep DB, Sollid JUE, Johannessen M. The interaction between Staphylococcus aureus SdrD and desmoglein 1 is important for adhesion to host cells. Sci Rep 2016; 6: 22134–22134.
9. Barbu EM, Ganesh VK, Gurusiddappa S, Mackenzie RC, Foster TJ, Sudhof TC, Höök M. Beta-Neurexin is a ligand for the Staphylococcus aureus MSCRAMM SdrC. PLoS Pathog 2010; 6(1):e1000726-e1000726.
10. Josefsson E, McCrea KW, Eidhin DN, O'Connell D, Cox J, Hook M, Foster TJ. Three new members of the serine-aspartate repeat protein multigene family of Staphylococcus aureus. Microbiology. 1998;144(12):3387–95.
11. Sabat A, Melles DC, Martirosian G, Grundmann H, Van Belkum A, Hryniewicz W. Distribution of the serine-aspartate repeat protein-encoding sdr genes among nasal-carriage and invasive Staphylococcus aureus strains. J Clin Microbiol. 2006;44(3):1135–8.
12. DeDent A, Bae T, Missiakas DM, Schneewind O. Signal peptides direct surface proteins to two distinct envelope locations of Staphylococcus aureus. EMBO J. 2008;27(20):2656–68.
13. Wang X, Ge J, Liu B, Hu Y, Yang M. Structures of SdrD from Staphylococcus aureus reveal the molecular mechanism of how the cell surface receptors recognize their ligands. Protein and Cell. 2013;4(4):277–85.
14. Josefsson E, O'Connell D, Foster TJ, Durussel I, Cox JA. The binding of calcium to the B-repeat segment of SdrD, a cell surface protein of Staphylococcus aureus. J Biol Chem. 1998;273(47):31145–52.
15. Roman AY, Devred F, Lobatchov VM, Makarov AA, Peyrot V, Kubatiev AA, Tsvetkov PO. Sequential binding of calcium ions to the B-repeat domain of SdrD from Staphylococcus aureus. Can J Microbiol. 2016;62(2):123–9.
16. Jenkins A, An Diep B, Mai TT, Vo NH, Warrener P, Suzich J, Kendall Stover C, Sellman BR. Differential expression and roles of Staphylococcus aureus virulence determinants during colonization and disease. MBio. 2015;6(1)
17. Corrigan RM, Miajlovic H, Foster TJ. Surface proteins that promote adherence of Staphylococcus aureus to human desquamated nasal epithelial cells. BMC Microbiol 2009; 9(22):22–22.
18. Sitkiewicz I, Babiak I, Hryniewicz W. Characterization of transcription within sdr region of Staphylococcus aureus. Antonie Van Leeuwenhoek. 2011;99(2):409–16.
19. Askarian F, Uchiyama S, Valderrama JA, Ajayi C, Sollid JUE, van Sorge NM, Nizet V, van Strijp JAG, Johannessen M. Serine-aspartate-repeat protein D Increases *Staphylococcus aureus* Virulence and Survival in Blood. Infection and immunity. 2016;IAI.00559–00516.
20. Trad S, Allignet J, Frangeul L, Davi M, Vergassola M, Couve E, Morvan A, Kechrid A, Buchrieser C, Glaser P, et al. DNA macroarray for identification and typing of Staphylococcus aureus isolates. J Clin Microbiol. 2004;42(5):2054–64.
21. Peacock SJ, Day NPJ, Thomas MG, Berendt AR, Foster TJ. Clinical isolates of Staphylococcus aureus exhibit diversity in fnb genes and adhesion to human fibronectin. J Infect. 2000;41(1):23–31.
22. Murai M, Moriyama H, Hata E, Takeuchi F, Amemura-Maekawa J. Variation and association of fibronectin-binding protein genes fnbA and fnbB in Staphylococcus aureus Japanese isolates. Microbiol Immunol. 2016;60(5):312–25.
23. Lower SK, Lamlertthon S, Casillas-Ituarte NN, Lins RD, Yongsunthon R, Taylor ES, DiBartola AC, Edmonson C, McIntyre LM, Reller LB, et al. Polymorphisms in fibronectin binding protein a of Staphylococcus aureus are associated with infection of cardiovascular devices. Proc Natl Acad Sci. 2011;108(45):18372–7.
24. Xue H, Lu H, Zhao X, Kluytmans J, vanBelkum A, Verbrugh H, Lowy FD, Baptiste KE, Williams K, Willams NJ et al. Sequence diversities of serine-aspartate repeat genes among Staphylococcus aureus isolates from different hosts presumably by horizontal gene transfer. PLoS One 2011; 6(5):e20332-e20332.
25. McCarthy AJ, Lindsay JA. Genetic variation in Staphylococcus aureus surface and immune evasion genes is lineage associated: implications for vaccine design and host-pathogen interactions. BMC Microbiol 2010; 10(1471–2180):173–173.
26. Bae T, Schneewind O. The YSIRK-G/S motif of staphylococcal protein a and its role in efficiency of signal peptide processing. J Bacteriol. 2003; 185(9):2910–9.
27. Dufour P, Jarraud S, Vandenesch F, Greenland T, Novick RP, Bes M, Etienne J, Lina G. High genetic variability of the agr locus in Staphylococcus species. J Bacteriol. 2002;184(4):1180–6.
28. Messina JA, Thaden JT, Sharma-Kuinkel BK, Fowler VG, Lindsay JA, Lindsay JA, Ruzin A, Ross HF, Kurepina N, Novick RP et al. Impact of bacterial and human genetic variation on Staphylococcus aureus infections. PLoS Pathog 2016;12(1):e1005330-e1005330.
29. Hazenbos WLW, Kajihara KK, Vandlen R, Morisaki JH, Lehar SM, Kwakkenbos MJ, Beaumont T, Bakker AQ, Phung Q, Swem LR et al. Novel staphylococcal glycosyltransferases SdgA and SdgB mediate immunogenicity and protection of virulence-associated cell wall proteins. PLoS Pathog 2013; 9(10):e1003653-e1003653.
30. Thomer L, Becker S, Emolo C, Quach A, Kim HK, Rauch S, Anderson M, Leblanc JF, Schneewind O, Faull KF, et al. N-acetylglucosaminylation of serine-aspartate repeat proteins promotes Staphylococcus aureus bloodstream infection. J Biol Chem. 2014;289(6):3478–86.
31. Cheng J, Xue H, Zhao X. Variation of serine-aspartate repeats in membrane proteins possibly contributes to staphylococcal microevolution. PLoS One 2012; 7(4):e34756-e34756.
32. Schneewind O, Mihaylova-Petkov D, Model P. Cell wall sorting signals in surface proteins of gram-positive bacteria. EMBO J. 1993;12(12):4803–11.
33. Lindsay JA, Holden MTG. Staphylococcus aureus: Superbug, super genome? Trends Microbio. 2004;12(8):378–85.
34. Everitt RG, Didelot X, Batty EM, Miller RR, Knox K, Young BC, Bowden R, Auton A, Votintseva A, Larner-Svensson H, et al. Mobile elements drive recombination hotspots in the core genome of Staphylococcus aureus. Nat Commun. 2014;5:2334–5.
35. Loughman A, Sweeney T, Keane FM, Pietrocola G, Speziale P, Foster TJ. Sequence diversity in the a domain of Staphylococcus aureus fibronectin-binding protein a. BMC Microbiol 2008; 8:74–74.
36. Sangvik M, Olsen RS, Olsen K, Simonsen GS, Furberg A-S, Sollid JUE. Age- and gender-associated Staphylococcus aureus spa types found among nasal carriers in a general population: the Tromso staph and skin study. J Clin Microbiol. 2011;49(12):4213–8.
37. Hanssen A-M, Kjeldsen G, Sollid JUE. Local variants of staphylococcal cassette chromosome mec in sporadic methicillin-resistant Staphylococcus aureus and methicillin-resistant coagulase-negative staphylococci: evidence of horizontal gene transfer? Antimicrob Agents Chemother. 2004;48(1):285–96.
38. Hall T. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nuclei Acids Symposium Series. 1999;41:95–8.
39. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007;23(21):2947–8.
40. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013; 30(4):772–80.
41. SIM Alignment Tool - Protein Sequences. https://web.expasy.org/sim/. Accessed 18 Feb 2017.
42. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics (Oxford, England). 2014;30(9):1312–3.
43. FigTree: http://tree.bio.ed.ac.uk/software/figtree/. Accessed 5 Mar 2017.
44. He Z, Zhang H, Gao S, Lercher MJ, Chen W-H, Hu S, et al. Evolview v2: an online visualization and management tool for customized and annotated phylogenetic trees. Nucleic Acids Res. 2016;44(W1):W236–41.
45. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput SequenceData. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 20 June 2017.
46. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19(5):455–77.
47. Gurevich A, Saveliev V, Vyahhi N, Tesler GQUAST. Quality assessment tool for genome assemblies. Bioinformatics (Oxford, England). 2013;29(8):1072–5.

48. Zhu W, Lomsadze A, Borodovsky M. Ab initio gene identification in metagenomic sequences. Nucleic Acids Res 2010; 38(12):e132-e132.
49. Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics. 2012;28(4):464–9.
50. Enright MC, Day NP, Davies CE, Peacock SJ, Spratt BG. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of Staphylococcus aureus. J Clin Microbiol. 2000;38(3):1008–15.
51. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. 2014;42(W1)
52. Schrodinger LLC. The PyMOL molecular graphics system. Version. 2015;1:8.
53. van de Guchte M, van der Vossen JM, Kok J, Venema G. Construction of a lactococcal expression vector: expression of hen egg white lysozyme in Lactococcus lactis subsp. lactis. Appl Environ Microbiol. 1989;55(1):224–8.
54. Monk IR, Shah IM, Xu M, Tan M-W, Foster TJ. Transforming the untransformable: application of direct transformation to manipulate genetically Staphylococcus aureus and Staphylococcus epidermidis. MBio. 2012;3(2):e00277–11.
55. Boukamp P, Petrussevska RT, Breitkreutz D, Hornung J, Markham A, Fusenig NE. Normal keratinization in a spontaneously immortalized aneuploid human keratinocyte cell line. J Cell Biol. 1988;106(3):761–71.

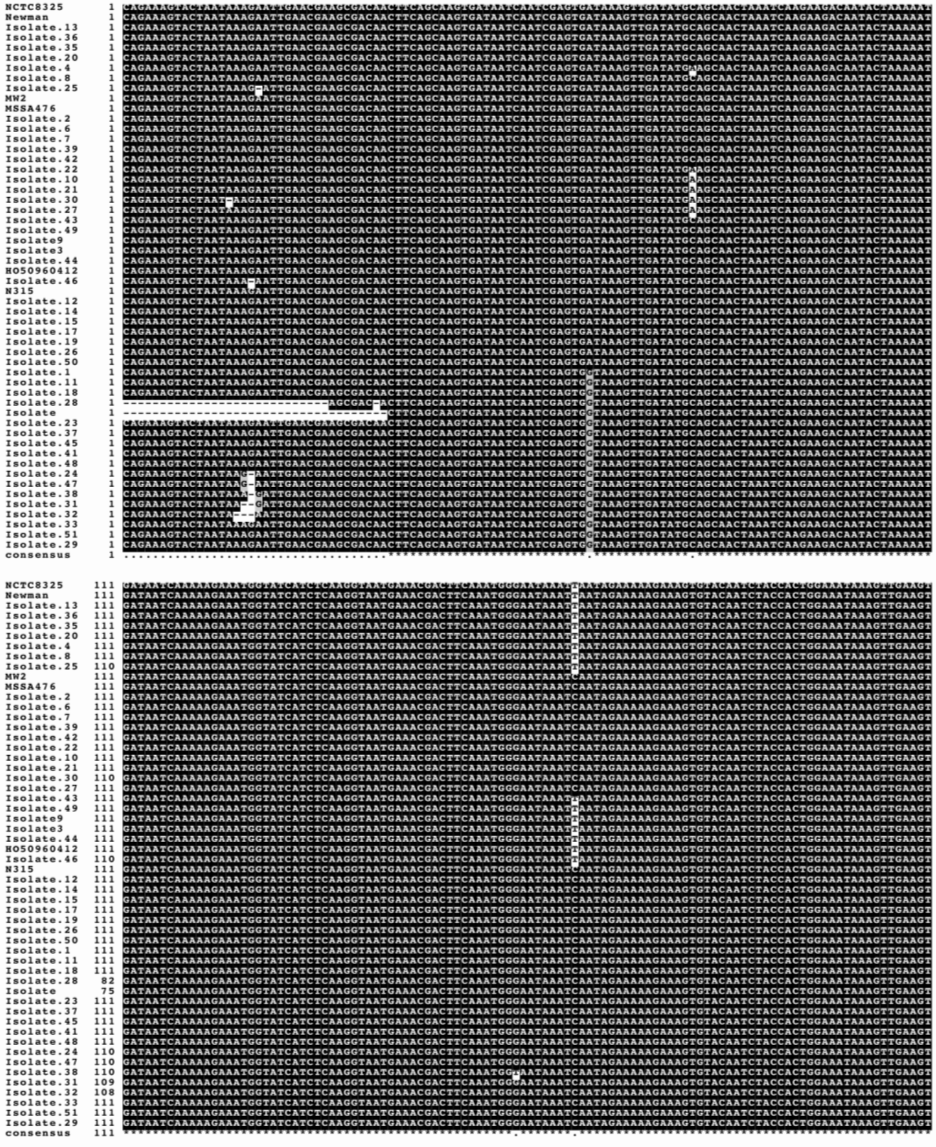
Ready to submit your research? Choose BMC and benefit from:

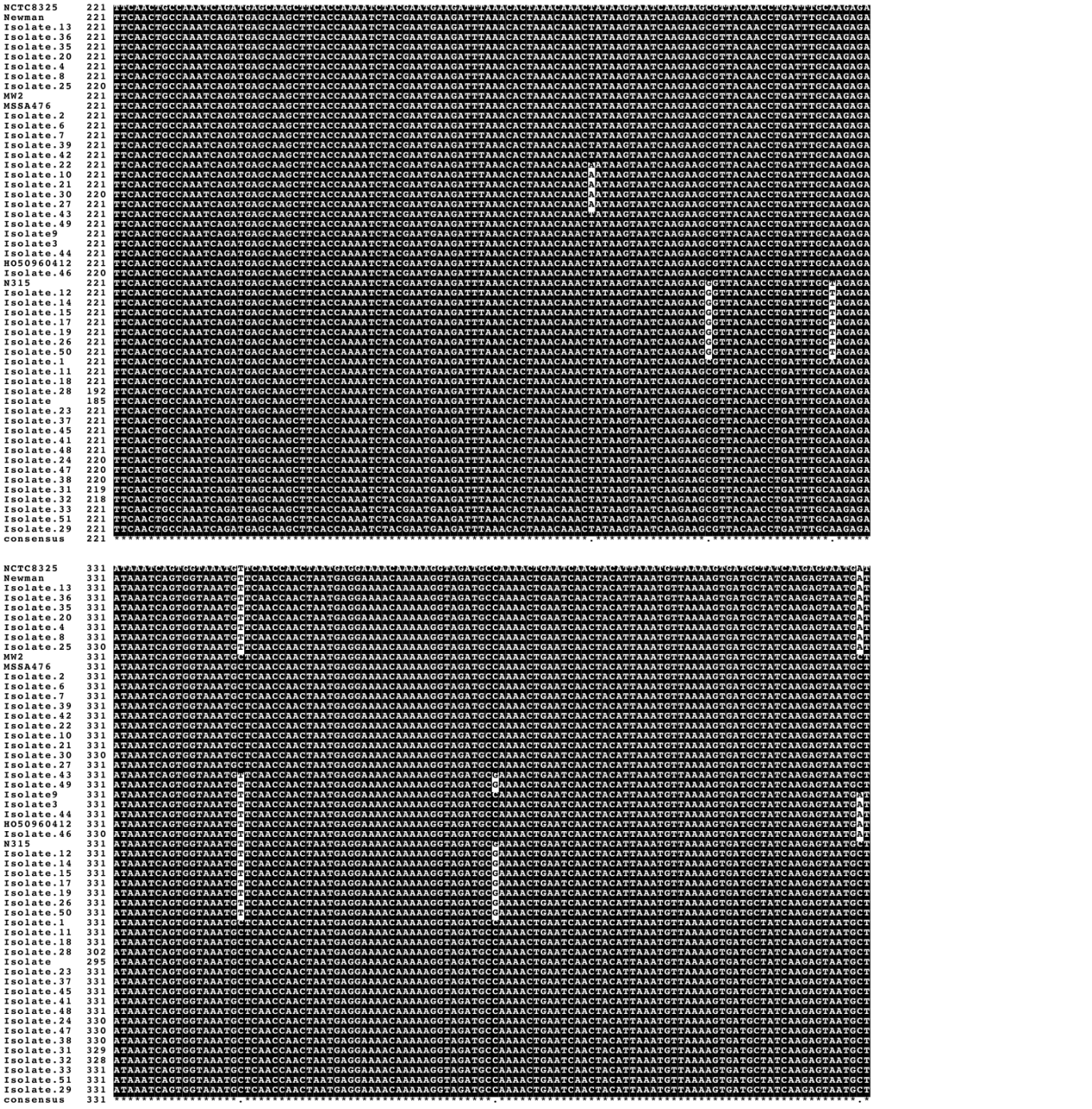
- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

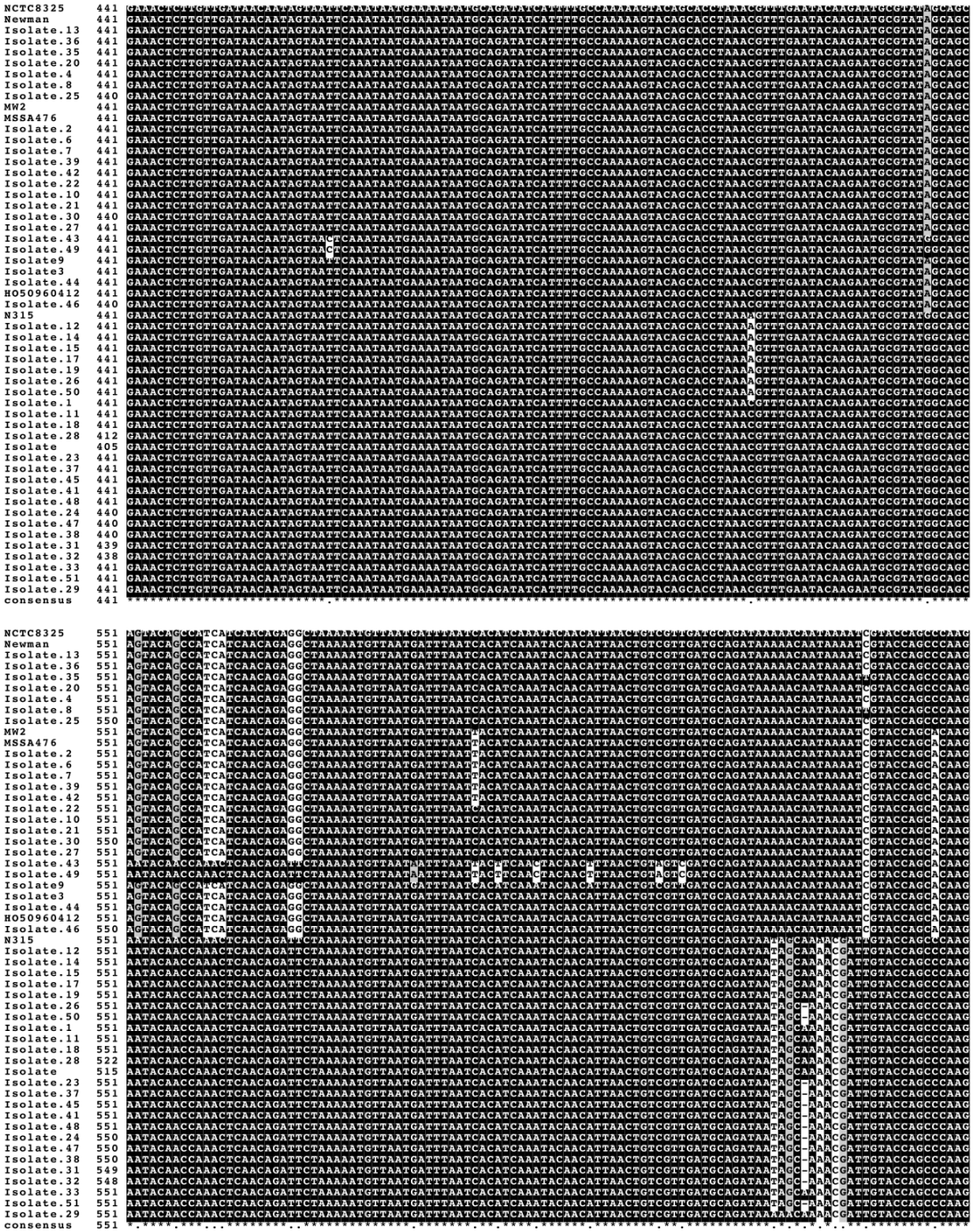
At BMC, research is always in progress.

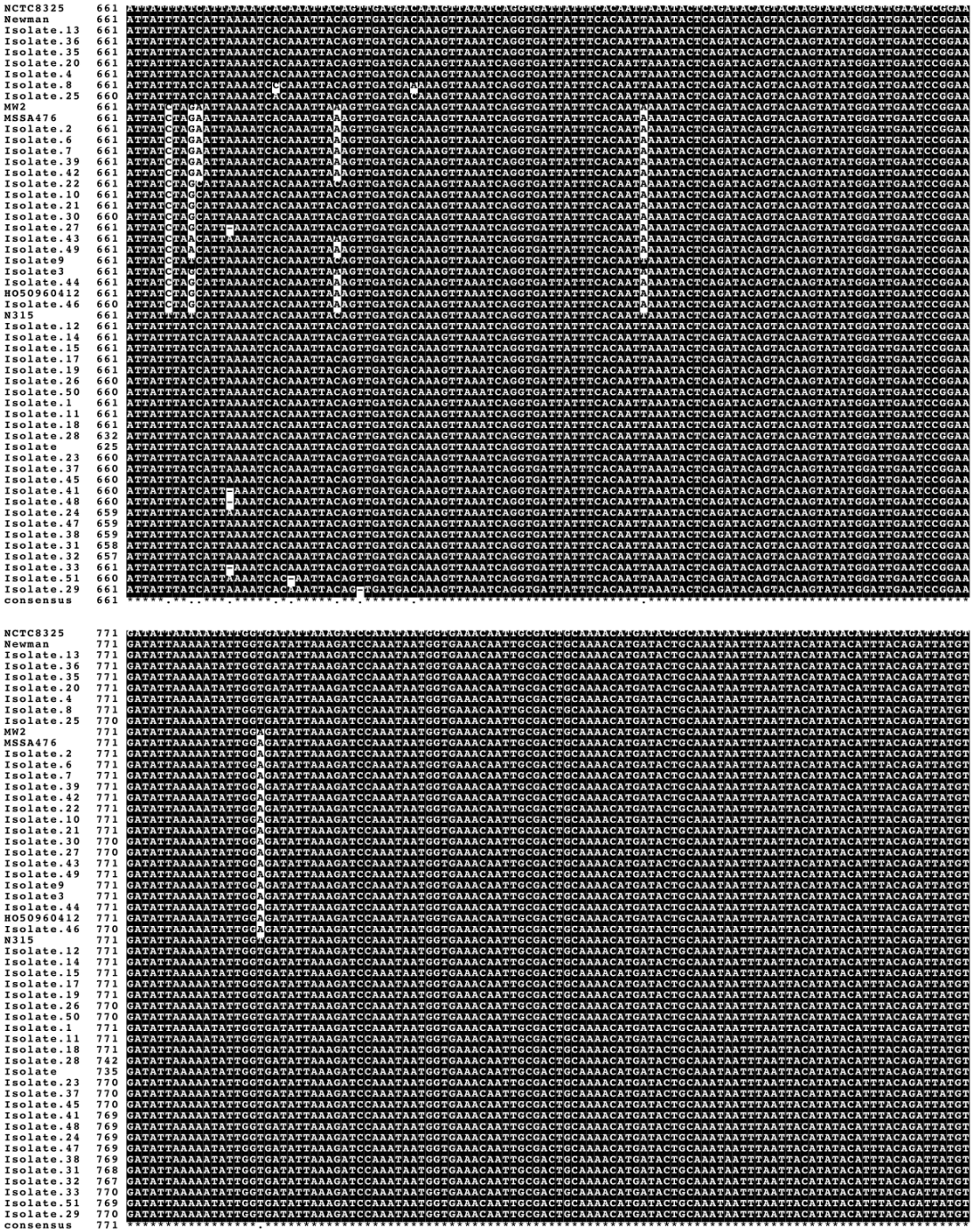
Learn more biomedcentral.com/submissions

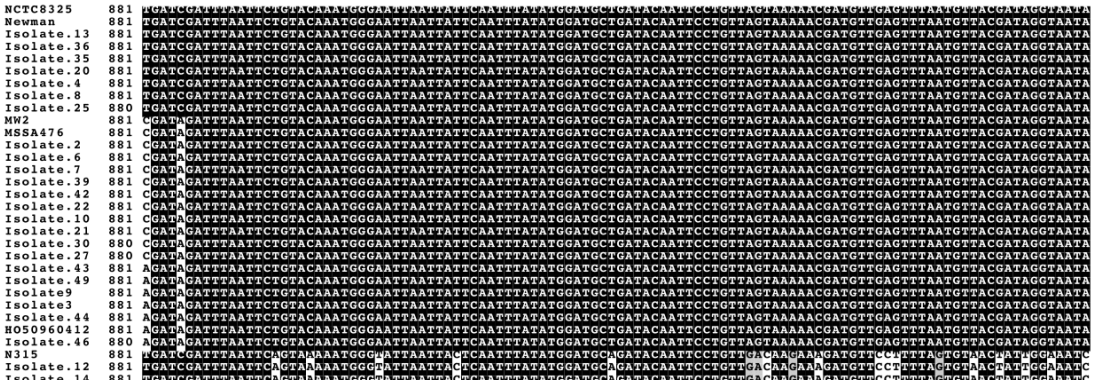
Additional File 1.



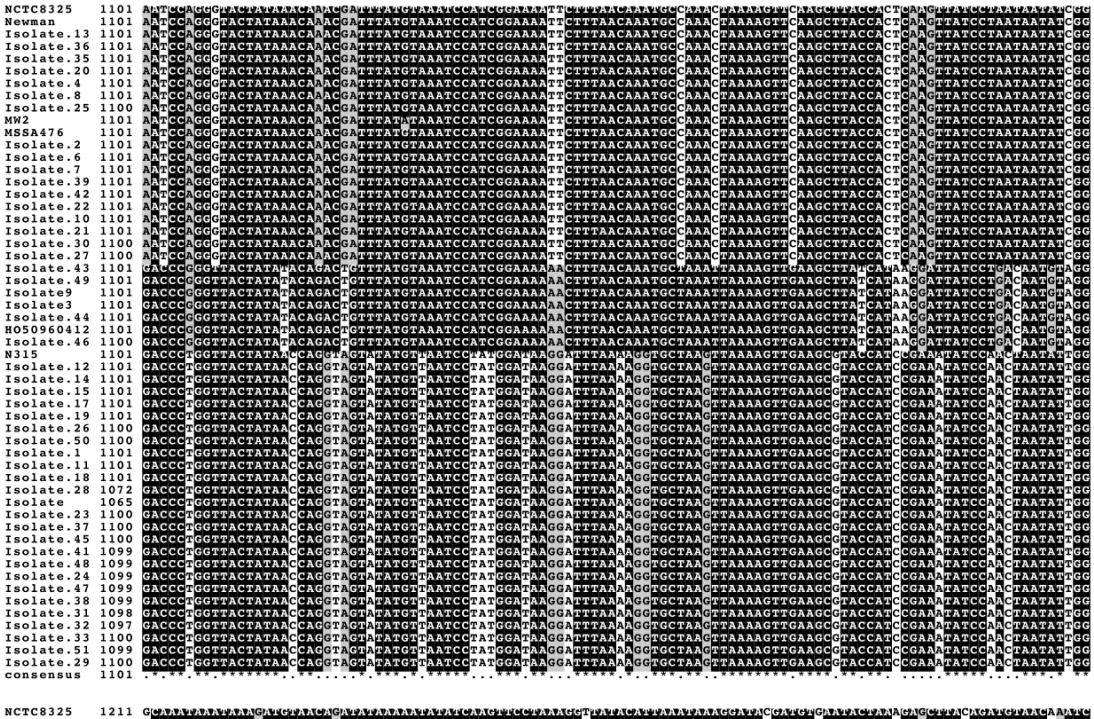


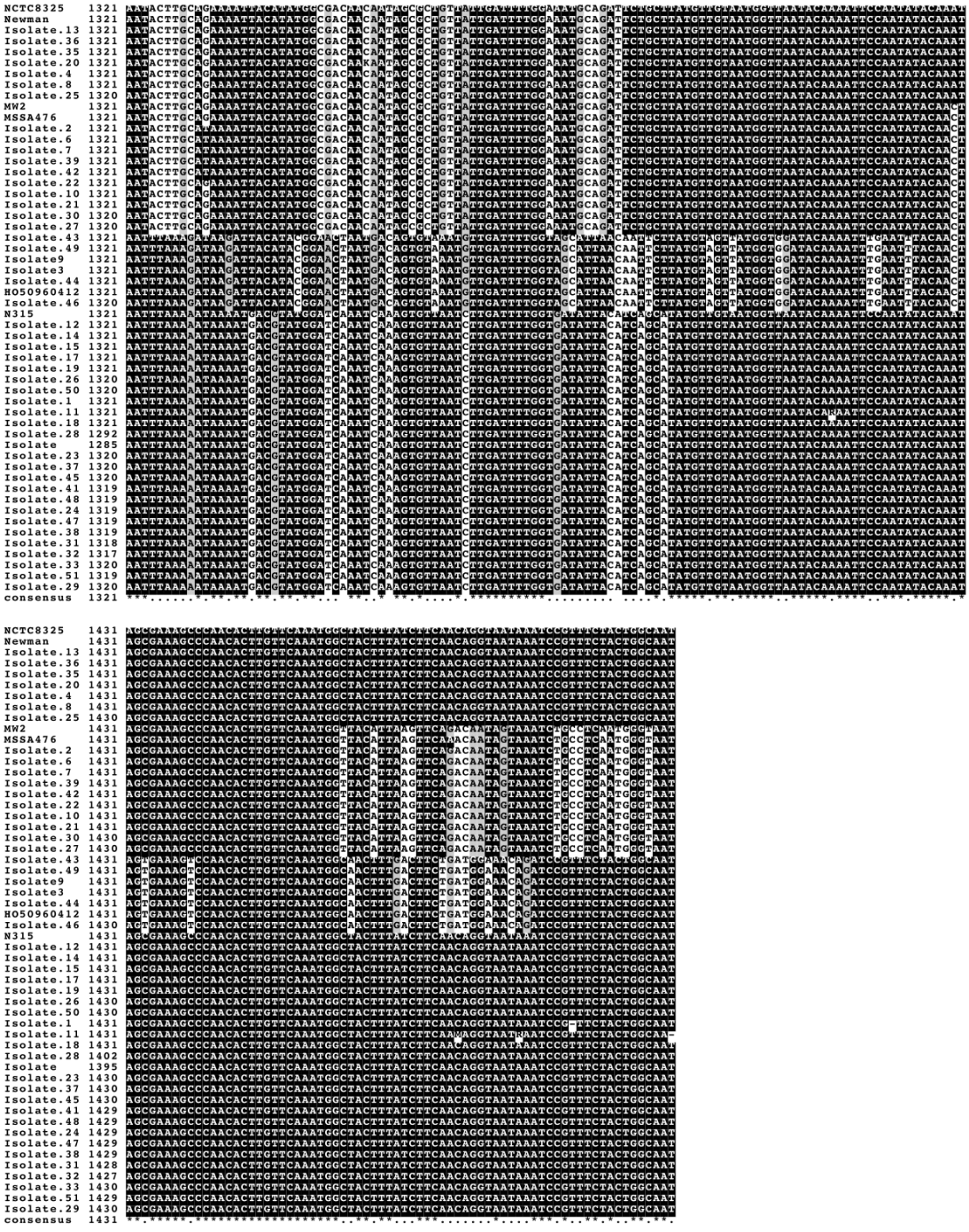




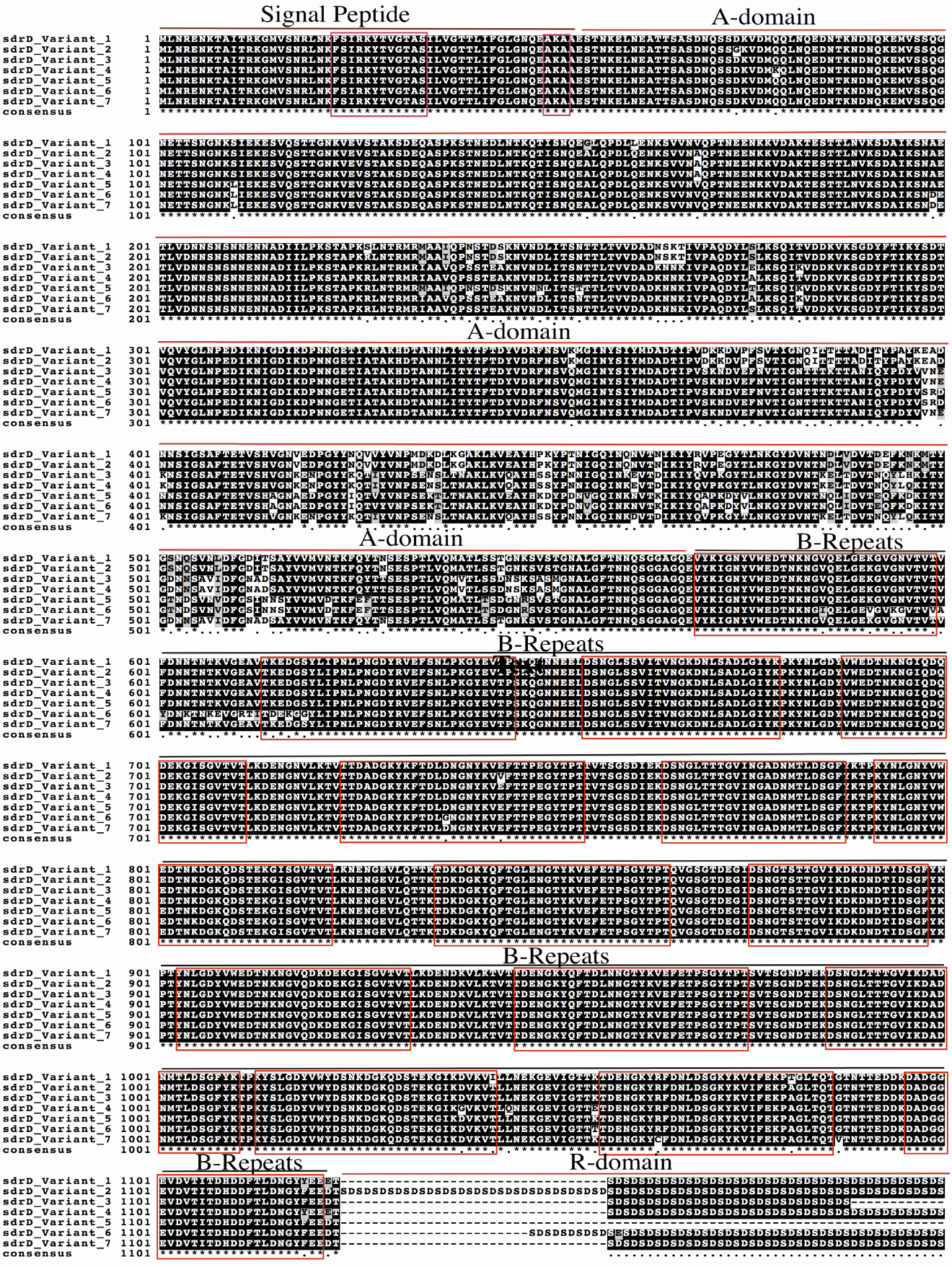

NCTC8325 881
Newman 881
Isolate.13 881
Isolate.36 881
Isolate.35 881
Isolate.20 881
Isolate.4 881
Isolate.8 881
Isolate.25 880
MW2 881
MSSA476 881
Isolate.2 881
Isolate.6 881
Isolate.7 881
Isolate.39 881
Isolate.42 881
Isolate.22 881
Isolate.10 881
Isolate.21 881
Isolate.30 880
Isolate.27 880
Isolate.43 881
Isolate.49 881
Isolate9 881
Isolate3 881
Isolate.44 881
HO50960412 881
Isolate.46 880
N315 881
Isolate.12 881
Isolate.14 881
Isolate.15 881
Isolate.17 881
Isolate.19 881
Isolate.26 880
Isolate.50 880
Isolate.1 881
Isolate.11 881
Isolate.18 881
Isolate.28 852
Isolate 845
Isolate.23 880
Isolate.37 880
Isolate.45 880
Isolate.41 879
Isolate.48 879
Isolate.24 879
Isolate.47 879
Isolate.38 879
Isolate.31 878
Isolate.32 877
Isolate.33 880
Isolate.51 879
Isolate.29 880
consensus 881

NCTC8325 991
Newman 991
Isolate.13 991
Isolate.36 991
Isolate.35 991
Isolate.20 991
Isolate.4 991
Isolate.8 991
Isolate.25 990
MW2 991
MSSA476 991
Isolate.2 991
Isolate.6 991
Isolate.7 991
Isolate.39 991
Isolate.42 991
Isolate.22 991
Isolate.10 991
Isolate.21 991
Isolate.30 990
Isolate.27 990
Isolate.43 991
Isolate.49 991
Isolate9 991
Isolate3 991
Isolate.44 991
HO50960412 991
Isolate.46 990
N315 991
Isolate.12 991
Isolate.14 991
Isolate.15 991
Isolate.17 991
Isolate.19 991
Isolate.26 990
Isolate.50 990
Isolate.1 991
Isolate.11 991
Isolate.18 991
Isolate.28 962
Isolate 955
Isolate.23 990
Isolate.37 990
Isolate.45 990
Isolate.41 989
Isolate.48 989
Isolate.24 989
Isolate.47 989
Isolate.38 989
Isolate.31 988
Isolate.32 987
Isolate.33 990
Isolate.51 989
Isolate.29 990
consensus 991




NCTC8325 1321
Newman 1321
Isolate.13 1321
Isolate.36 1321
Isolate.35 1321
Isolate.20 1321
Isolate.4 1321
Isolate.8 1321
Isolate.25 1320
MW2 1321
MSSA476 1321
Isolate.2 1321
Isolate.6 1321
Isolate.7 1321
Isolate.39 1321
Isolate.42 1321
Isolate.22 1321
Isolate.10 1321
Isolate.21 1321
Isolate.30 1320
Isolate.27 1320
Isolate.43 1321
Isolate.49 1321
Isolate9 1321
Isolate3 1321
Isolate.44 1321
HO50960412 1321
Isolate.46 1320
N315 1321
Isolate.12 1321
Isolate.14 1321
Isolate.15 1321
Isolate.17 1321
Isolate.19 1321
Isolate.26 1320
Isolate.50 1320
Isolate.1 1321
Isolate.11 1321
Isolate.18 1321
Isolate.28 1292
Isolate 1285
Isolate.23 1320
Isolate.37 1320
Isolate.45 1320
Isolate.41 1319
Isolate.48 1319
Isolate.24 1319
Isolate.47 1319
Isolate.38 1319
Isolate.31 1318
Isolate.32 1317
Isolate.33 1320
Isolate.51 1319
Isolate.29 1320
consensus 1321
NCTC8325 1431
Newman 1431
Isolate.13 1431
Isolate.36 1431
Isolate.35 1431
Isolate.20 1431
Isolate.4 1431
Isolate.8 1431
Isolate.25 1430
MW2 1431
MSSA476 1431
Isolate.2 1431
Isolate.6 1431
Isolate.7 1431
Isolate.39 1431
Isolate.42 1431
Isolate.22 1431
Isolate.10 1431
Isolate.21 1431
Isolate.30 1430
Isolate.27 1430
Isolate.43 1431
Isolate.49 1431
Isolate9 1431
Isolate3 1431
Isolate.44 1431
HO50960412 1431
Isolate.46 1430
N315 1431
Isolate.12 1431
Isolate.14 1431
Isolate.15 1431
Isolate.17 1431
Isolate.19 1431
Isolate.26 1430
Isolate.50 1430
Isolate.1 1431
Isolate.11 1431
Isolate.18 1431
Isolate.28 1402
Isolate 1395
Isolate.23 1430
Isolate.37 1430
Isolate.45 1430
Isolate.41 1429
Isolate.48 1429
Isolate.24 1429
Isolate.47 1429
Isolate.38 1429
Isolate.31 1428
Isolate.32 1427
Isolate.33 1430
Isolate.51 1429
Isolate.29 1430
consensus 1431

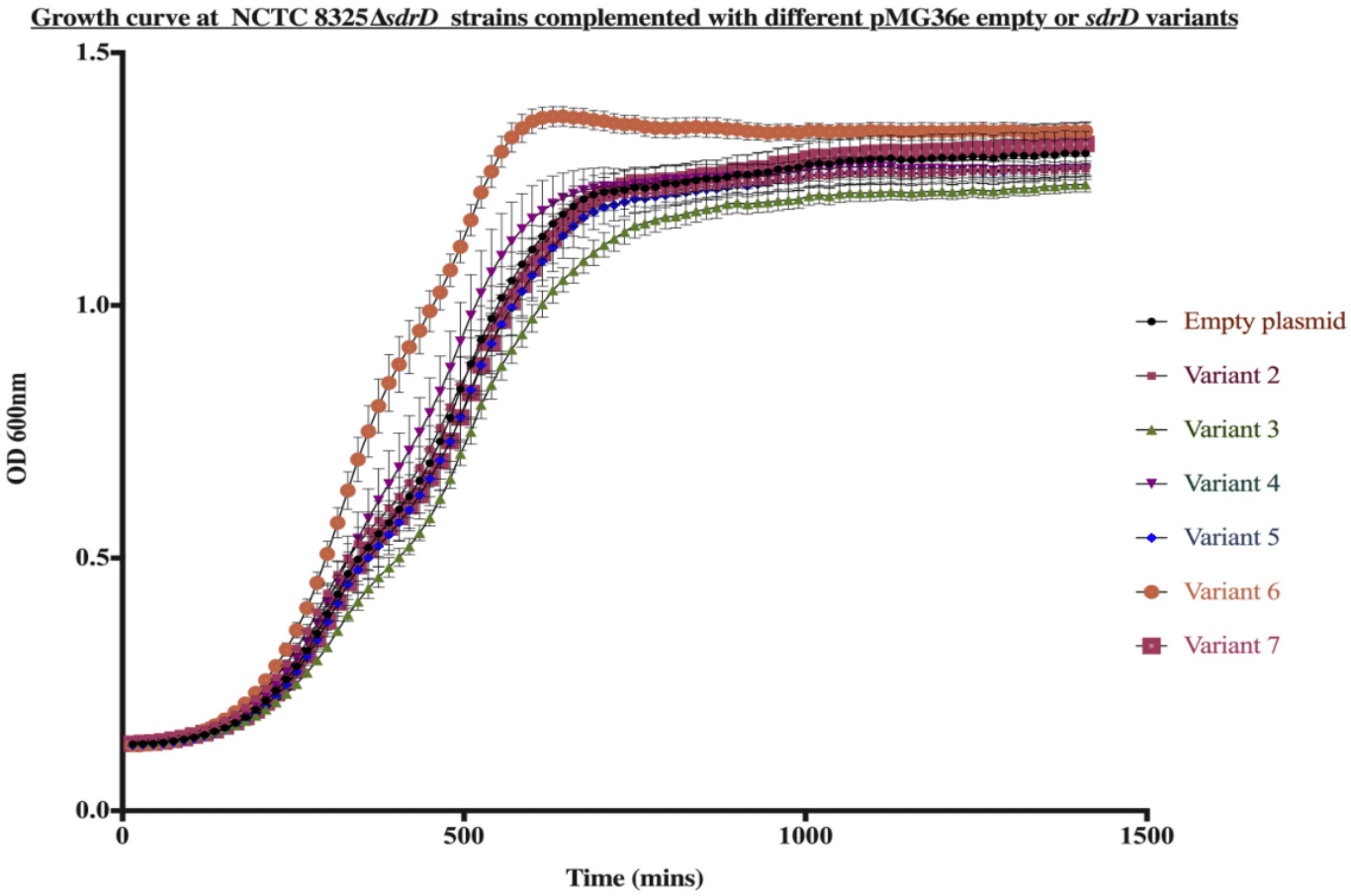
Additional File 2



R-domain
R-domain
WMC (sorting signal)
WMC

Additional file 3.

A.



B

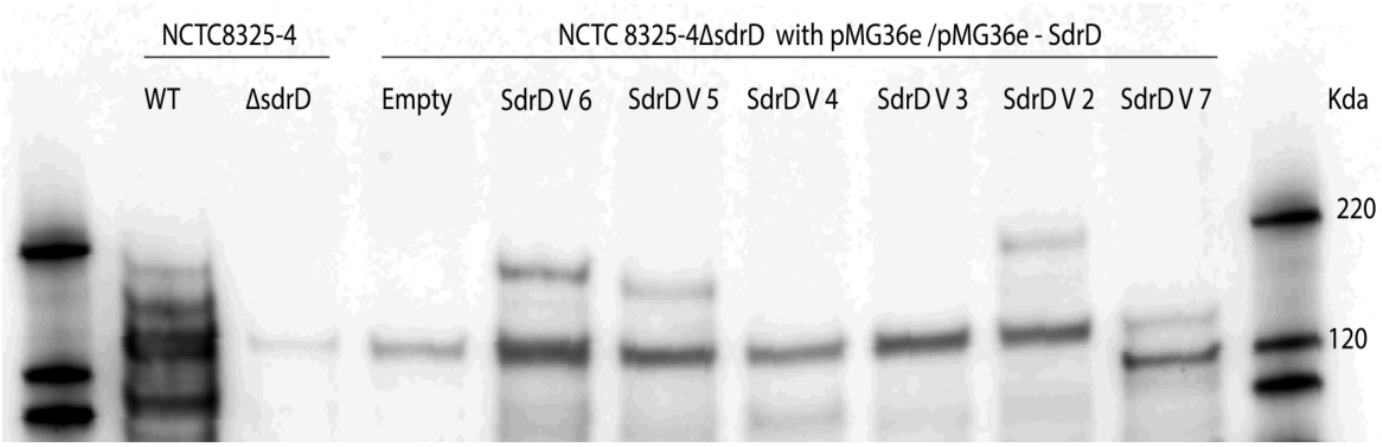


Table S1: Percentage Amino Acid Identity based on amino acid residues for the sdrD A domain for the representative isolates

	SdrD Variant 1	SdrD Variant 2	SdrD Variant 3	SdrD Variant 4	SdrD Variant 5	SdrD Variant 6	SdrD Variant 7
SdrD Variant 1	100	99.3	85.4	85.4	87.1	86.4	86.9
SdrD Variant 2	99.3	100	86,0	86.0	87.3	86.5	87.1
SdrD Variant 3	85.4	86,0	100	99.2	88.7	89.9	97.3
SdrD Variant 4	85.4	86.0	99.2	100	88.3	89.7	97.5
SdrD Variant 5	87.1	87.3	88.7	88.3	100	98.2	89.3
SdrD Variant 6	86.4	86.5	89.9	89.7	98.2	100	90.8
SdrD Variant 7	86.9	87.1	97.3	97.5	89.3	90.8	100,0