

SCIENTIFIC REPORTS



OPEN

Characterization of a community-acquired-MRSA USA300 isolate from a river sample in Austria and whole genome sequence based comparison to a diverse collection of USA300 isolates

Sarah Lepuschitz^{1,2}, Steliana Huhulescu¹, Patrick Hyden³, Burkhard Springer¹, Thomas Rattei³, Franz Allerberger¹, Robert L. Mach² & Werner Ruppitsch^{1,4}

The increasing emergence of multi-resistant bacteria in healthcare settings, in the community and in the environment represents a major health threat worldwide. In 2016, we started a pilot project to investigate antimicrobial resistance in surface water. Bacteria were enriched, cultivated on selective chromogenic media and species identification was carried out by MALDI-TOF analysis. From a river in southern Austria a methicillin resistant *Staphylococcus aureus* (MRSA) was isolated. Whole genome sequence analysis identified the isolate as ST8, *spa* type t008, SCCmecIV, PVL and ACME positive, which are main features of CA-MRSA USA300. Whole genome based cgMLST of the water isolate and comparison to 18 clinical MRSA USA300 isolates from the Austrian national reference laboratory for coagulase positive staphylococci originating from 2004, 2005 and 2016 and sequences of 146 USA300 isolates arbitrarily retrieved from the Sequence Read Archive revealed a close relatedness to a clinical isolate from Austria. The presence of a CA-MRSA USA300 isolate in an aquatic environment might pose a public health risk by serving as a potential source of infection or a source for emergence of new pathogenic MRSA clones.

Methicillin resistant strains of *Staphylococcus aureus* (MRSA) are the leading cause of nosocomial infections¹. Manifestations vary from minor skin infection to fatal disease. MRSA have an outstanding ability to acquire antibiotic resistance genes leading to resistance to multiple antibiotic classes². Since the late 1960s hospital acquired MRSA (HA-MRSA) became endemic in hospitals worldwide³. In the 1990s new clones affecting also healthy individuals in the community, featuring increased virulence as well as the power to spread easily, arose⁴. This so called community acquired MRSA (CA-MRSA) became prevalent worldwide with predominant clones in certain geographic areas. Clonal Complex (CC) 1 (USA400) and CC8 (USA300) are the major lineages in the United States, CC80 is the predominant lineage in Europe, CC59 and CC30 appear in Asia and the southwest Pacific^{5,6}. CA-MRSA is now epidemic in the United States mainly due to dissemination of the USA300 clone which belongs to multi locus sequence type (MLST) 8/SCCmec IV and harbours the *lukS-lukF* genes, encoding the Panton-Valentine leukocidin (PVL) and the arginine catabolic mobile element (ACME) cluster⁷. While virulence of these strains enhances dissemination, the potential to acquire resistance to multiple antibiotic classes hinders treatment of MRSA infections¹. CA-MRSA clones are not restricted to a geographical region. The ability

¹Austrian Agency for Health and Food Safety, National Reference Laboratory for Coagulase Positive Staphylococci including *Staphylococcus aureus*, Graz, Austria. ²Vienna University of Technology, Research Area of Biochemical Technology, Institute of Chemical, Biological and Environmental Engineering, Vienna, Austria. ³University of Vienna, Department of Microbiology and Ecosystem Science, Vienna, Austria. ⁴University of Natural Resources and Life Sciences, Department of Biotechnology, Vienna, Austria. Correspondence and requests for materials should be addressed to S.L. (email: sarah.lepuschitz@ages.at)

of geographically predominant CA-MRSA clones to spread worldwide and to cause infections on other continents has been demonstrated². In several European countries, where ST80 still represents the prevalent CA-MRSA clone, also infections due to the North American USA300 clone have been reported^{8–11}. Consequently the identification of new potential infection sources is essential for infection control.

Antimicrobial resistance (AMR) is an increasing global problem and the fact that resistant pathogens are not restricted to clinical settings but can be increasingly found in the environment is alarming^{12–14}. In aquatic systems, horizontal gene transfer (HGT) of resistance genes between bacteria leads to the evolution of AMR bacteria, which in turn find back into clinical settings. AMR dissemination of resistant strains occur through direct contact of humans to the AMR bacteria in the aquatic system or from contact to resistant environmental bacteria which pass resistance genes to human or animal pathogens¹⁵. So called high risk clones (HRCs), which are characterized by enhanced virulence and multiple antibiotic resistances, pose a serious public health risk¹⁶. Aanensen *et al.*⁸ identified three key elements to tackle the public health threat caused by HRCs: genetic population structure and identification of HRCs, assessment of risks posed by virulence and resistance determinants and risk management by implementation of prevention and control strategies.

The Austrian Agency for Health and Food Safety, the major Austrian organization responsible for consumer protection and public health, started a pilot project analysing surface water samples from rivers and bathing sites with the aim to identify possible public health risks due to AMR HRCs in 2016¹⁷. Environmental screening for HRCs and high discriminatory typing identified a CA-USA300 isolate in a river water sample in Austria in 2016.

In this report we describe a detailed characterization of this USA300 MRSA isolate and its phylogenetic relatedness to clinical CA-MRSA USA300 isolates.

Results

Screening of the 12 water samples for MRSA revealed one sample from a river in southern Austria (46°38'14"N, 14°17'44"E) positive for MRSA (sample ID: W1). Whole genome sequencing (WGS) identified the river water isolate W1 as a USA300 strain. The isolate showed classical multi locus sequence type (MLST) sequence type (ST) 8, *spa* type t008, harboured SCCmecIV, the Panton-Valentine leucocidin (PVL) genes, the arginine catabolic mobile element (ACME) cluster and USA300 specific *cap5* mutations.

Genetic comparison of ST8 isolates. To assess relationship of the river water isolate W1 to clinical and food associated isolates, 18 clinical isolates from the Austrian national reference laboratory for coagulase positive staphylococci were characterized by WGS and sequences of 146 arbitrarily chosen USA300 isolates were retrieved from the Sequence Read Archive (SRA).

From the comparison of isolate W1 to in total 164 isolates, isolate W1 was indistinguishable from 99 USA300 isolates concerning ST8, *spa* type t008, SCCmecIV, PVL genes, the ACME cluster and *cap5* mutations. Eighteen isolates carried no ACME cluster, six isolates carried no PVL and four isolates carried no ACME cluster and no PVL. Thirty isolates had ST8, carried the ACME cluster and PVL but had different *spa* types: t024 (n = 5), t063 (n = 2), t068 (n = 1), t121 (n = 6), t16000 (n = 1), t2031 (n = 11), t622 (n = 3), t681 (n = 1). One isolate had ST8, *spa* type t622, PVL and no ACME cluster; one isolate had ST8, *spa* type t203, the ACME cluster and no PVL; four isolates had ST8, different *spa* types (t024 (n = 1), t190 (n = 2), t5271 (n = 1)) and carried no ACME cluster and no PVL. One isolate had ST3885, *spa* type t008, the PVL genes and the ACME cluster.

Whole genome based cgMLST phylogenetic analysis including *S. aureus* reference strain COL (accession no. NC_002951.2) (ST250), water isolate (W1), the reference strains FPR3757 (ATCC® BAA-1556™), TCH1516 (USA300-HOU-MR, ATCC® BAA-1717™), all clinical USA300 isolates (n = 18) from the Austrian national reference laboratory for coagulase positive staphylococci (from 2004, 2005 and 2016), and 143 genomes from isolates retrieved from SRA was performed and a minimum spanning tree was calculated (Fig. 1). Distance calculation between all 164 samples revealed a maximum allelic distance between samples of 351 and an average allelic distance of 93.5 across the MST. Clinical Austrian isolates differed among each other in minimum of two, a maximum of 327, and an average distance of 100 alleles. Based on the defined complex threshold (CT) of 24 allelic differences¹⁸ 25 different complexes were obtained. Six out of 18 clinical Austrian isolates (H1/04-H3/04; H4/05-H6/05) showed close relatedness with a maximum allelic difference of 21 and were all located in complex 2 (Fig. 1) respectively in complex 1 (Fig. 2). Isolate H3/04, a human isolate from Austria in 2004, showed close relatedness to strains C2406 with an allelic difference of 14, USFL055 isolated in the United States in 2009 with an allelic difference of 22, N28973PS isolated in the United States in 2011 from turkey meat with an allelic difference of 24, and reference strain FPR3757 with an allelic difference of 23. Surface water isolate W1 differed by 54 alleles from C2406 and was closest related to the Austrian clinical isolate H5/16 differing by two alleles in cgMLST and by four alleles in the pan genome including the following targets from the reference genome *S. aureus* (SA)COL (accession no. NC_002951.2): SACOL1678 (luciferase, YP_186518.1, G996T), SACOL1708 (type III leader peptidase, YP_186547.1, C218T) in the core genome and the targets (n = 2) SACOL0289 (hypothetical protein, YP_185183.1, G56T, C57T, G84A, A105C) and SACOL0651 (hypothetical protein, YP_185536.1, C524A) in the accessory genome.

A total of 69 genes were identified referring virulence attributes in USA300 samples W1, H1/04-H3/04, H1/05-H6/05, H1/06, H1/16-H8/16, C2406 and FPR3757. Water isolate W1 harboured 61 virulence gene loci. Comparison of all these 21 isolates revealed 26 identical virulence genes and 60 identical virulence alleles between W1 and H5/16, 59 identical virulence alleles between W1, C2046 and reference strain FPR3757, and 64 identical virulence alleles between H3/04 and C2406 (Supplementary Table S1). All 21 isolates (W1, H1/04-H3/04, H1/05-H6/05, H1/06, H1/16-H8/16, C2406, FPR3757) carried the *agr* allele 1, all except H6/16 and H7/16 carried Panton-Valentine leucocidin (*lukS*-PV/*lukF*-PV), seventeen isolates carried the USA300 characteristic arginine catabolic mobile element (ACME) (W1, H1/04-H3/04, H3/05-H6/05, H1/06, H1/16, H3/16-H6/16, H8/16, C2406, FPR3757), and sixteen isolates carried *seq* (H1/04-H3/04, H2/05-H6/05, H1/06, H2/16-H5/16, H8/16, C2406, FPR3757).

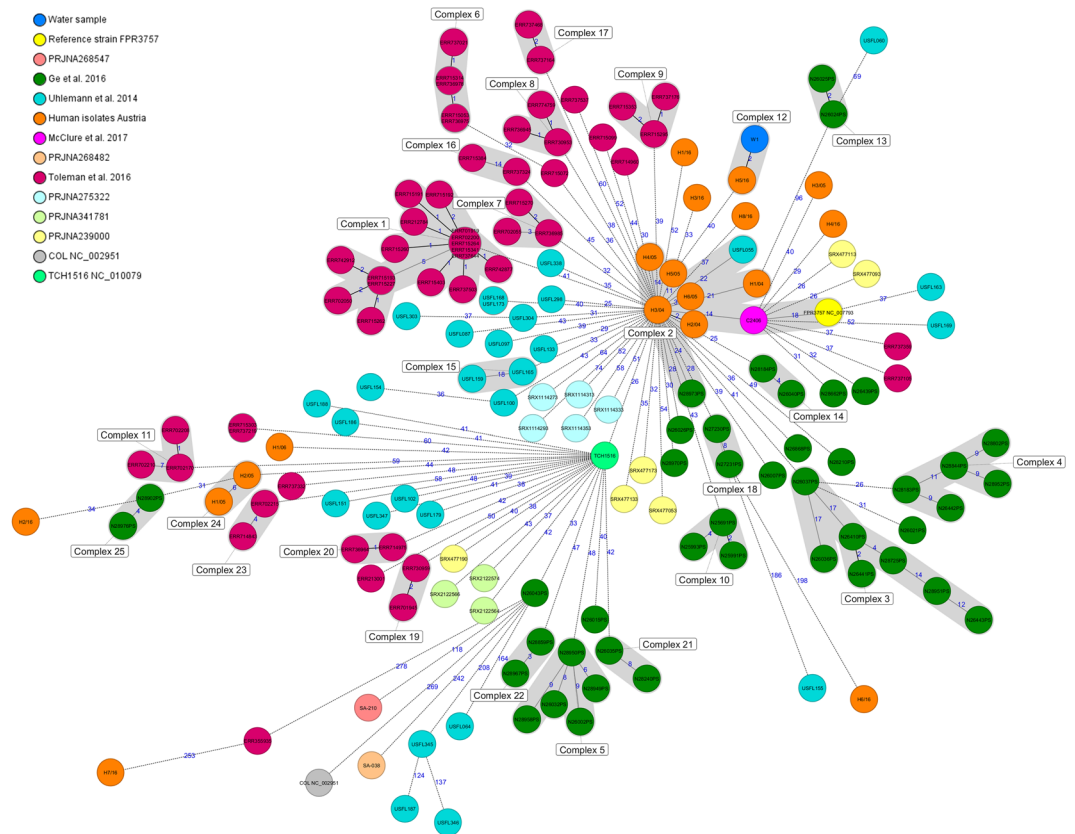


Figure 1. Minimum spanning tree for 165 MRSA isolates based on the cgMLST of *S. aureus*. Colours correspond to the origin of the samples. Each circle represents isolates with an allelic profile based on the sequence of 1,861 core genome targets. Blue numbers refer to the allelic differences between two isolates. Isolates with closely related genotypes were identified with a maximum of 24 allelic differences and are shaded in grey.

Phenotypic susceptibility testing revealed that the water isolate W1 was resistant to beta-lactams (benzylpenicillin, amoxicillin-clavulanic acid, ceftiofur), fluoroquinolones (ciprofloxacin, moxifloxacin) and erythromycin among the seventeen antibiotics tested (Table 1).

In total 25 genes conferring resistance to various antibiotics were identified in USA300 samples W1, H1/04-H3/04, H1/05-H6/05, H1/06, H1/16-H8/16 C2406 and FPR3757. The water isolate W1 harboured in total 14 resistance genes (Table 2) and showed complete concordance to H5/16, sharing nine identical resistance targets with FPR3757 and *fosB* with C2406. SNPs conferring resistance to ciprofloxacin were identified in 15 out of 21 isolates (except H1/04, H1/05, H2/05, H1/06, H2/16, H6/16), due to a L84S change in *gyrA*.

The presence of antimicrobial resistance against penicillin, methicillin, erythromycin and ciprofloxacin in isolate W1 was confirmed by *in silico* analysis using Mykrobe predictor.

Comparative analysis of isolate W1 to the closest relative isolate H5/16 via OrthoFinder revealed following differences in presence and absence of protein coding genes: additional presence of two domain-containing proteins (DUF443) and one hypothetical protein in isolate W1 and the additional presence of ten hypothetical proteins, four transposases, two family proteins (TIGR01741), one domain-containing protein (DUF5079), one immunoglobulin G-binding protein A and one transcriptional regulator in isolate H5/16.

Comparison of isolate W1 to isolate C2406 revealed nine additional hypothetical proteins, three domain-containing proteins (DUF443 (n = 2), DUF536), three transcriptional regulators, three transferases, two transporter, one penicillinase repressor BlaI, beta-lactam sensor/signal transducer BlaR1, penicillin-hydrolyzing class A beta-lactamase BlaZ, quinone oxidoreductase, one transposase and one ABC-F type ribosomal protection protein MsrA in isolate W1 and additional eight hypothetical proteins, two family proteins (TIGR01741), two transposases, one domain-containing protein (DUF5079) and one transcriptional regulator in isolate C2406 (Supplementary Table S2).

Additionally 15 genes in isolate W1 were identified for which no orthologue could be assigned in the other genomes in this comparison: two ATP-binding proteins, two family proteins (TIGR01741), two hypothetical proteins, one acyltransferase, one LysR family transcriptional regulator, one AI-2E family transporter, one homoserine dehydrogenase, one recombination protein RecJ and one DUF5079 domain-containing protein. Three genes revealed no significant result and could not be assigned to any known group of protein coding genes (Supplementary Table S3).

Antibiotic	MIC breakpoint (mg/L)		USA300 W1 (MIC)
	S ≤	R >	
Benzylpenicillin	0.125	0.125	R (>32)
Amoxicillin-clavulanic acid	n.a	n.a	R (48)
Cefoxitin	4	4	R (>128)
Ciprofloxacin	1	1	R (>32)
Moxifloxacin	0.25	0.25	R (1.5)
Amikacin	8	16	S (1.5)
Gentamicin	1	1	S (0.125)
Teicoplanin	2	2	S (0.75)
Vancomycin	2	2	S (1)
Erythromycin	1	2	R (32)
Clindamycin	0.25	0.5	S (0.125)
Minocycline	0.5	1	S (0.064)
Linezolid	4	4	S (1.5)
Fosfomycin	32	32	S (0.5)
Fusidic acid	1	1	S (0.125)
Rifampicin	0.06	0.5	S (0.008)
Trimethoprim	2	4	S (0.38)

Table 1. Phenotypical resistance data for sample W1. MIC (minimum inhibitory concentration) breakpoints accord to EUCAST. S = sensitive, R = resistant, I = intermediate, n.a. = not applicable.

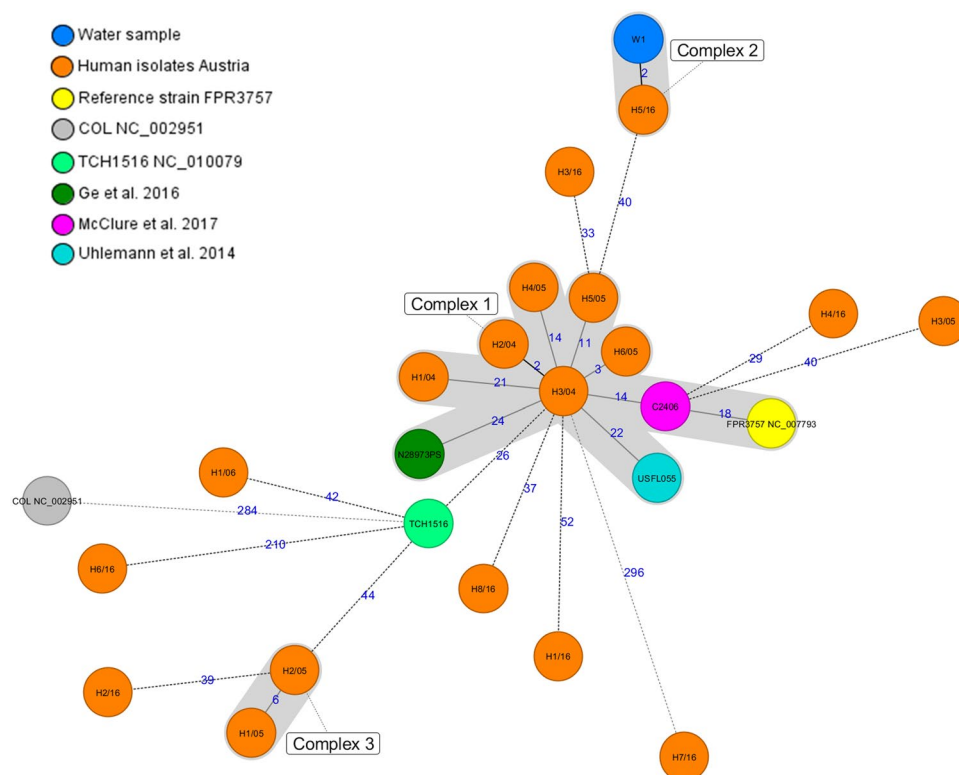


Figure 2. Minimum spanning tree for 25 MRSA isolates based on the cgMLST of *S. aureus*. Colours correspond to the origin of the samples. Each circle represents isolates with an allelic profile based on the sequence of 1,861 core genome targets. Blue numbers refer to the allelic differences between two isolates. Isolates with closely related genotypes were identified with a maximum of 24 allelic differences and are shaded in grey. Austrian isolates: W1, H1/04-H3/04, H1/05-H6/05, H1/06, H1/16-H8/16; reference strains: FPR3757, COL (NC_002951.2); closest related international isolates = C2406²⁶, N28973PS³⁹, TCH1516³⁸, USLF055³⁶.

The circular map (CGView Server V 1.0 (2007)) depicts sequence similarity between the reference strain FPR3757, isolates H5/16 and W1 and sequence similarity between COL (NC_002951.2), FPR3757 and isolate W1 in Fig. 3.

Target	W1	H1/04	H2/04	H3/04	H1/05	H2/05	H3/05	H4/05	H5/05	H6/05	H1/06	H1/16	H2/16	H3/16	H4/16	H5/16	H6/16	H7/16	H8/16	C2406	FPR3757	Resistance against	Accession Nb.
aadD	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	1	ND	ND	ND	aminoglycosides	BA000017.4
aphA3	ND	1	1	1	1	1	1	1	1	ND	1	ND	ND	ND	1	ND	ND	ND	1	ND	ND	aminoglycosides	CP009681.1
blaI	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	ND	1	7	ND	ND	ND	beta-lactam	SRR016154
blaR	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	ND	ND	15	6	ND	ND	beta-lactam	SRR016154
blaZ	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	ND	1	10	6	ND	ND	beta-lactam	SRR016154
ccrA2	6	NAT	6	6	6	6	6	6	6	NAT	6	6	6	6	6	6	ND	17	6	1	6	methicillin	NC010079
ccrB2	6	NAT	6	6	6	6	6	6	6	ND	6	6	15	6	6	6	ND	1	6	1	6	methicillin	NC010079
ccrC	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	NAT	ND	ND	3	ND	ND	methicillin	AP008934.1
dfrA	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	1	ND	1	ND	1	trimethoprim	AE017171.1
ermA	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	1	ND	ND	ND	ND	MLSbK	BA000017.4
fosB	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	fosfomycin	CP000046.1
lmrP	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	ND	1	various	CP000046.1
mecA	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4	3	3	1	3	methicillin	BX571856.1
mecR-intact	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	1	ND	1	ND	ND	methicillin	BA000017.4
mecR-truncated	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	ND	FQC	1	ND	1	methicillin	CP000046.1
merA	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	1	ND	ND	ND	ND	SCCmec	NC013352.1
merB	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	1	ND	ND	ND	ND	SCCmec	AB179623.1
mphC	2	2	2	2	2	2	2	2	2	2	2	2	FQC	FQC	2	2	ND	ND	2	ND	ND	MLSbK	NC017351
mprF	1	1	1	1	1	1	1	1	1	1	1	NAT	1	1	1	1	1	1	1	ND	1	methicillin-oxacillin	CP000046.1
msrA	2	2	2	2	2	2	2	2	2	2	2	2	FQC	2	2	ND	ND	NAT	ND	ND	ND	MLSbK	NC017351
qacA	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	NAT	ND	ND	ND	ND	various	AF053771.1
sdrM	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	NAT	ND	1	various	CP000046.1
tetK	ND	ND	ND	ND	ND	ND	4	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	tetracycline	sampleLGO017
ugpQ	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	ND	1	SCCmec	CP000046.1
xylR	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	2	ND	ND	ND	SCCmec	BA000018.3

Table 2. Identified resistance genes and allele types in USA300 samples W1, H1/04–H3/04, H1/05–H6/05, H1/06, H1/16–H8/16, C2406 and FPR3757. Numbers accord to the respective allele type for each target; FQC = failed target QC procedure; NAT = new allele type; ND = not detected.

Discussion

The discharge, persistence and dissemination of AMR in nature are considered a major public health threat worldwide¹⁹. Therefore a strategy to better control the release of antibiotics into the environment and subsequently to prevent contact of bacteria from human and animal sources with environmental organisms²⁰ is of utmost importance.

Staphylococcus aureus was particularly successful in developing resistance to antibiotics. MRSA is listed as of high priority on WHO's antibiotic-resistant pathogens list (<http://www.who.int/medicines/publications/global-priority-list-antibiotic-resistant-bacteria/en/>). Since the first detection of MRSA with association to healthcare (HA-MRSA) settings in 1961²¹, new epidemic MRSA clones have emerged affecting the community (CA-MRSA) or being associated with animals (LA-MRSA)^{22–24}. Surveillance and characterisation of clinical, animal and environmental isolates is a public health imperative.

The MRSA isolate described was obtained from a water sample from a small river in southern Austria²⁴. Characterisation of this isolate by WGS analysis revealed that the isolate has all characteristic features of CA-MRSA USA300⁷. CgMLST and wgMLST based gene-by-gene comparison¹⁸ to eighteen clinical isolates (USA300) available from the National Reference Laboratory for Staphylococci including *Staphylococcus aureus* revealed a very close relatedness to one clinical USA300 isolate (H5/16) collected 2016 in Vienna, differing only by two alleles in their core- and by four alleles in their whole-genome targets. The water and patient isolates shared 60 identical virulence and 14 antibiotic resistance alleles. We were unable to find an epidemiological link between these two isolates. Four isolates collected in 2016 from a hospital located in one of the settlements next to this river, showed no close relatedness to the USA300 water isolate.

An interesting additional outcome of this WGS based typing study was that clinical USA300 isolates from Austria are a diverse population indicating that USA300 clones were introduced several times in Austria and are not all descendants from the first Austrian USA300 isolate from 2004. According to the WGS patterns all eight clinical Austrian USA300 isolates from 2016 are unrelated to each other. Thus in comparison to previous typing methods that did not allow a differentiation (same MLST, same *spa* type), WGS based typing is an important improvement representing an added information value for hospitals allowing a clear discrimination of outbreak and transmission events from unrelated isolates. In due consideration of staphylococcal mutation rates of one SNP per 15 weeks²⁵ we can assume that the water isolate W1 and several clinical isolates from Austria are descendants of the complex containing the hypervirulent USA300-C2406 strain from a patient with necrotizing pneumoniae isolated in Canada in 2004²⁶ and reference strain FPR3757¹. In contrast to the aforementioned Austrian USA300 isolates from 2016, the first USA300 isolates detected in Austria in 2004 to 2005¹⁰ belong to four different complex types. One complex contained isolates collected from two patients, stationary in the same hospital in 2005, whereof the initial one was derived from a postal worker processing post from the United States,

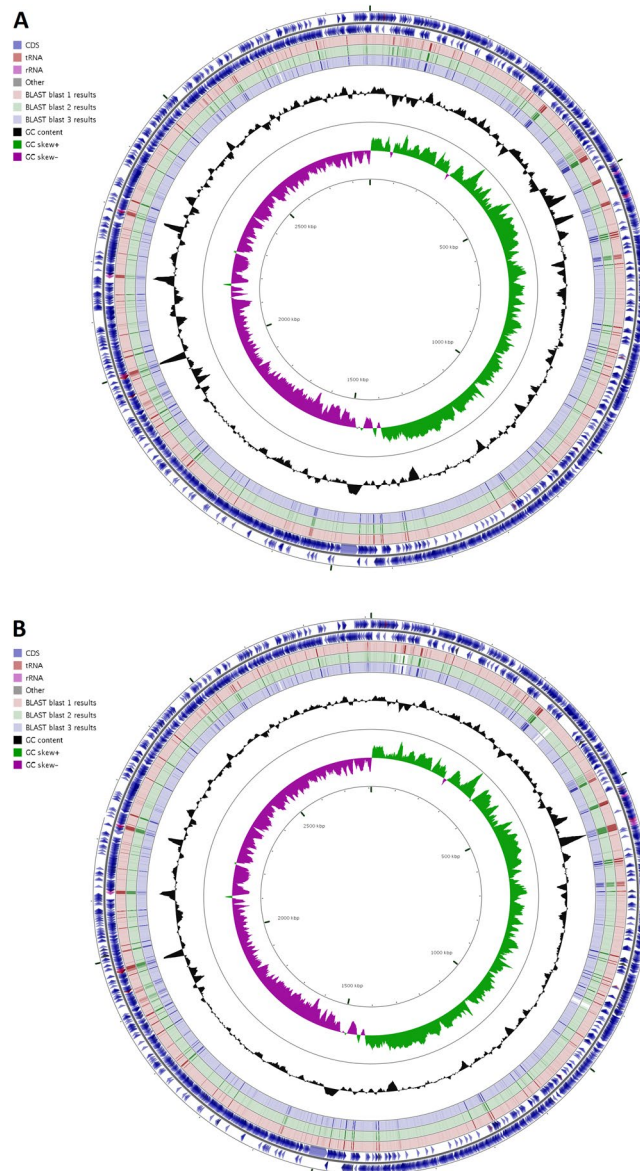


Figure 3. Comparative circular maps of *Staphylococcus aureus* genomes USA300_FPR3757 (NC_007793) and COL (NC_003951.2) to USA300 isolates. (A) The circles display the following information (from outside to inside): CDSs of USA300_FPR3757 on the + strand, CDSs of USA300_FPR3757 on the – strand, blastn result 1: USA300_FPR3757, blastn result 2: H5/16, blastn result 3: W1, GC content, GC skew + (green) and – (violet) value; (B) The circles display the following information (from outside to inside): CDSs of COL (NC002951.2) on the + strand, CDSs of COL (NC002951.2) on the – strand, blastn result 1: COL (NC002951.2), blastn result 2: USA300_FPR3757, blastn result 3: W1, GC content, GC skew + (green) and – (violet) value.

which had a clear transmission link. The second complex contained six isolates indicating a transmission or a common source of infection. Within this complex an epidemiological link to the United States could only be confirmed for one isolate (isolate ID H4/05) as well as the close relatedness between a couple (isolate ID H2/04 & H3/04) and an Austrian isolate from 2005 (isolate H6/05). CgMLST analysis shows that Austrian isolates of the second complex are closely related to several US isolates including C2046 and reference strain FPR3757 isolated from a HIV-positive patient in the United States in 2000¹ demonstrating a transfer. In contrast to the United States, where USA300 has become the most successful MRSA clone in the community as well as in hospitals, USA300 is still rare in most European countries and Austria. However, the fact that USA300 infections transmitted via contaminated environmental sources have been demonstrated²⁷ and the detection in a river water sample with close relatedness of the water isolate to a clinical isolate might indicate a possible risk for public health.

In conclusion, this is to our best knowledge the first genome based characterisation of an environmental USA300 isolate confirming the emerging global health threat caused by environmental AMR. While numerous environmental AMR studies have reported on Gram negative bacteria our example shows that also Gram positive

bacteria might play their role. Considering the historical evolution of MRSA clones the emergence of new pathogenic clones is merely a question of time.

For appropriate actions further investigations are essential to determine the occurrence, risk, environmental impact of AMR in the environment as well as use of optimal wastewater treatment technologies to alleviate this public health threat²⁸.

Methods

Microorganisms and Species Identification. In 2016 we collected 12 water samples from four rivers (Danube (n = 3, 48°14'53"N, 14°25' 37"E; 48°14'45"N, 14°26'1"E; 48°14'32"N, 14°26'41"E), Glan (n = 3, 46°42'46"N, 14°6'23"E; 46°38'1"N, 14°18'56"E; 46°38'14"N, 14°17'44"E), Inn (n = 4, 47°16'35"N, 11°12'50"E; 47°15'55"N, 11°16'58"E; 47°16'14"N, 11°26'16"E; 47°16'6"N, 11°27'42"E), Traun (n = 1, 48°13'2"N, 14°15'26"E)) flowing through the Austrian provinces Carinthia, Salzburg, Tyrol, Lower Austria, Upper Austria, and Vienna. Samples were collected in sterile 500 ml sodium thiosulfate containing bottles (Corning® Gosselin™, NY, USA), 20 to 30 cm below the river surface and 50 to 100 cm apart from the river bank. For MRSA screening 100 ml aliquots of samples were filtrated and bacterial isolates were enriched by incubating the filter in tryptone soya broth (containing 3.5 mg/ml cefoxitin and 75 mg/ml aztreonam) at 37 °C overnight according to the standard protocol for isolating MRSA from dust samples (<http://www.eurl-ar.eu/233-protocols.html>, Isolation of MRSA from dust samples). To minimize the risk of contamination water sample filtration and cultivation were separated spatially, locally and temporally from cultivation of clinical samples. Laboratories were weekly screened for contaminations using air settlement plates and surface swabbing according to a standard operation protocol (SOP) following international guidelines²⁹. For detection of methicillin resistant staphylococci, the overnight cultures were cultivated on BBL™ CHROMagar™ MRSA II according to the manufacturer's instructions (Becton Dickinson, Vienna, Austria). Mauve colonies from chromogenic media were sub-cultured and identified by matrix assisted laser desorption/ionisation time-of-flight (MALDI-TOF) mass spectrometry.

The clinical MRSA USA300 isolates from the Austrian National Reference Laboratory for Coagulase Positive Staphylococci originating from 2016 (H1/16 to H4/16) came from a hospital located in a settlement beside the river (1,470 beds) in Carinthia and four isolates (H5/16 to H8/16) were from Vienna and from the province Lower Austria. Additionally isolates from the Austrian provinces Vienna, Lower Austria, Salzburg and Styria from 2004 (H1/04-H3/04), 2005 (H1/05-H6/05) and 2006 (H1/06) were analysed.

Antibiotic susceptibility testing. *In vitro* susceptibility was tested for 17 antibiotics (benzylpenicillin, amoxicillin-clavulanic acid, cefoxitin, ciprofloxacin, moxifloxacin, amikacin, gentamicin, teicoplanin, vancomycin, erythromycin, clindamycin, minocycline, linezolid, fosfomycin, fusidic acid, rifampicin, trimethoprim) by Etests (bioMérieux, Marcy-l'Étoile, France) and interpreted according to EUCAST (European Committee on Antimicrobial Susceptibility Testing, EUCAST Clinical Breakpoint Tables v.7.1, valid from 2017-03-10).

Whole genome sequencing and data analysis. For whole genome sequencing (WGS) high quality genomic DNA (gDNA) was isolated using the MagAttract HMW DNA Kit (Qiagen, Hilden, Germany) and quantified with a Qubit® 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) using the dsDNA BR Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA). Libraries of bacterial genomes were prepared using Nextera XT DNA Library Preparation Kit (Illumina, San Diego, CA, USA) according to the manufacturer's protocol. A desired coverage of at least 50-fold was calculated with Sequencing Coverage Calculator (www.illumina.com/CoverageCalculator) and bacterial isolates were paired-end sequenced with a read length of 2 × 300 basepairs on a MiSeq instrument (Illumina, San Diego, CA, USA). SPAdes version 3.9.0³⁰ was used for read assembly and SeqSphere+ version 4.1.9 (Ridom, Münster, Germany) for strain characterisation using a recently published cgMLST scheme¹⁸ and calculation of minimum spanning trees (MST). Phylogenetic relatedness between water and related clinical isolates was further visualized using CGView Server V 1.0(2007)³¹.

Spa type³², MLST (multilocus sequence type)³³, cgMLST (core genome MLST)¹⁸, resistance genes and virulence genes³⁴ were extracted from WGS data using SeqSphere+ version 4.1.9. Mutations in *cap5* were identified according to the publication of Boyle-Vavra *et al.* (2015) and mutations in *gyrA* causing ciprofloxacin resistance were identified according to Sreedharan *et al.*^{35,36}. Genotype-based antimicrobial-resistance was predicted for isolate W1 by using Mykrobe predictor for *Staphylococcus aureus*³⁷.

Genomes from Sequence Read Archive. One-hundred and forty-six isolates downloaded from the Sequence Read Archive (SRA) belonged to following Bioprojects: PRJEB2870 (n = 28)³⁸, PRJEB3174 (n = 55)¹¹, PRJNA224116 (n = 3)^{1,39,40}, PRJNA239000 (n = 6), PRJNA268482 (n = 1), PRJNA268547 (n = 1), PRJNA275322 (n = 5), PRJNA311554 (n = 43)⁴¹, PRJNA341781 (n = 3), PRJNA345240 (n = 1)²⁶.

Analysis of gene presence and absence. Detailed analysis of additional genomic information was performed for 10 genomes: Reference strains C2406 (CP019590.1) and FPR3757 (CP000255.1) and isolates H1/04, H2/04, H3/04, H4/04, H5/04, H6/04, H05/16 and W1. The contigs of each assembly was filtered for a minimum length of 1,000 nucleotides. Genes were predicted using prodigal v. 2.6.3⁴² with default parameters, orthologous groups were calculated with OrthoFinder v. 1.1.4⁴³. Orthologous groups with differences in presence/absence were selected. Annotation of the orthologous groups and genes without group assignment was performed using NCBI-BLASTp v. 2.6.0+ with e-value cutoff of 0.01⁴⁴ and the RefSeq nr-protein database release 84⁴⁵. The most frequent annotation in an orthologous groups was selected as description for the group.

Nucleotide accession number(s). This Whole Genome Shotgun project including ST8 isolates (H1/16-H8/16) from the Austrian national reference laboratory for coagulase positive staphylococci in 2016 and

the water isolate (W1) have been deposited at DDBJ/ENA/GenBank under the accession NKCP0000000 to NKCX0000000. The version described in this paper is version NKCP01000000 to NKCX01000000.

Data availability. All data generated or analysed during this study are included in this published article.

References

- Diep, B. A. *et al.* Complete genome sequence of USA300, an epidemic clone of community-acquired methicillin-resistant *Staphylococcus aureus*. *Lancet*. **367**, 731–739 (2006).
- DeLeo, F. R., Otto, M., Kreiswirth, B. N. & Chambers, H. F. Community-associated methicillin-resistant *Staphylococcus aureus*. *Lancet*. **375**, 1557–1568 (2010).
- Fridkin, S. K. *et al.* Methicillin-resistant *Staphylococcus aureus* disease in three communities. *N Engl J Med*. **352**, 1436–1444 (2005).
- DeLeo, F. R. & Chambers, H. F. Reemergence of antibiotic-resistant *Staphylococcus aureus* in the genomics era. *J Clin Invest*. **119**, 2464–2474 (2009).
- Aires-de-Sousa, M. Methicillin-resistant *Staphylococcus aureus* among animals: current overview. *Clin Microbiol Infect*. **23**, 373–380 (2017).
- Lindsay, J. Genomic variation and evolution of *Staphylococcus aureus*. *Int J Med Microbiol*. **300**, 98–103 (2010).
- Tenover, F. C. & Goering, R. V. Methicillin-resistant *Staphylococcus aureus* strain USA300: origin and epidemiology. *J Antimicrob Chemother*. **64**, 441–446 (2009).
- Aanensen, D. M. *et al.* Whole-genome sequencing for routine pathogen surveillance in public health: a population snapshot of invasive *Staphylococcus aureus* in Europe. *MBio*. **7**, <https://doi.org/10.1128/mBio.00444-16> (2016).
- Köck, R. *et al.* Methicillin-resistant *Staphylococcus aureus* (MRSA): burden of disease and control challenges in Europe. *Euro Surveill*. **15**, <https://doi.org/10.2807/ese.15.41.19688-en> (2010).
- Ruppitsch, W. *et al.* Occurrence of the USA300 community-acquired *Staphylococcus aureus* clone in Austria. *Euro Surveill*. **12**, <https://doi.org/10.2807/esw.12.43.03294-en> (2007).
- Toleman, M. S. *et al.* Systematic surveillance detects multiple silent introductions and household transmission of methicillin-resistant *Staphylococcus aureus* USA300 in the east of England. *J Infect Dis*. **214**, 447–453 (2016).
- Mushtaq, A. UN commits to tackling antimicrobial resistance. *Lancet Infect Dis*. **16**, 1229–1230 (2016).
- Viens, A. M. & Littmann, J. Is antimicrobial resistance a slowly emerging disaster? *Public Health Ethics*. **8**, 255–265 (2015).
- WHO: Emerging and other communicable diseases: antimicrobial resistance. World Health Organization, Geneva, Switzerland (1998).
- Taylor, N. G., Verner-Jeffreys, D. W. & Baker-Austin, C. Aquatic systems: maintaining, mixing and mobilising antimicrobial resistance? *Trends Ecol Evol*. **26**, 278–284 (2011).
- Cantón, R. & Ruiz-Garbajosa, P. Co-resistance: an opportunity for the bacteria and resistance genes. *Curr Opin Pharmacol*. **11**, 477–485 (2011).
- WHO: Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics. World Health Organization, Geneva, Switzerland (2017).
- Leopold, S. R., Goering, R. V., Witten, A., Harmsen, D. & Mellmann, A. Bacterial whole-genome sequencing revisited: portable, scalable, and standardized analysis for typing and detection of virulence and antibiotic resistance genes. *J Clin Microbiol*. **52**, 2365–2370 (2014).
- Narciso-da-Rocha, C. & Manaia, C. M. The influence of the autochthonous wastewater microbiota and gene host on the fate of invasive antibiotic resistance genes. *Sci Total Environ*. **57**, 932–940 (2017).
- Baquero, F., Martínez, J. L. & Cantón, R. Antibiotics and antibiotic resistance in water environments. *Curr Opin Biotechnol*. **19**, 260–265 (2008).
- Jevons, M. P. “Celbenin”-resistant staphylococci. *Br Med J*. **1**, 124–125 (1961).
- Bal, A. M. *et al.* Genomic insights into the emergence and spread of international clones of healthcare-, community- and livestock-associated methicillin-resistant *Staphylococcus aureus*: Blurring of the traditional definitions. *J Glob Antimicrob Resist*. **6**, 95–101 (2016).
- Smith, T. C., Moritz, E. D., Leedom Larson, K. R. & Ferguson, D. D. The environment as a factor in methicillin-resistant *Staphylococcus aureus* transmission. *Rev Environ Health*. **25**, 121–134 (2010).
- Lepuschitz, S., Mach, R., Springer, B., Allerberger, F., Ruppitsch, W. Draft Genome Sequence of a Community-Acquired Methicillin-Resistant *Staphylococcus aureus* USA300 Isolate from a River Sample. *Genome Announc*. **5**, <https://doi.org/10.1128/genomeA.01166-17> (2017).
- Worby, C. J., Chang, H. H., Hanage, W. P. & Lipsitch, M. The distribution of pairwise genetic distances: a tool for investigating disease transmission. *Genetics*. **198**, 1395–1404 (2014).
- McClure, J. A. & Zhang, K. Complete genome sequence of a community-associated methicillin-resistant *Staphylococcus aureus* hypervirulent strain, USA300-C2406, isolated from a patient with a lethal case of necrotizing pneumonia. *Genome Announc*. **5**, <https://doi.org/10.1128/genomeA.00461-17> (2017).
- Kazakova, S. V. *et al.* A clone of methicillin-resistant *Staphylococcus aureus* among professional football players. *N Engl J Med*. **352**, 468–475 (2005).
- Waseem, H., Williams, M. R., Stedtfeld, R. D. & Hashsham, S. A. Antimicrobial Resistance in the Environment. *Water Environ Res*. **2017** **89**, 921–941 (2017).
- Eleftheriadou, M. & Tsimillis K. C. (Eds). *Eurachem guide: Accreditation for Microbiological Laboratories*, Second edition, ISBN: 978-91-87017-92-6. Available from www.eurachem.org (2013).
- Nurk, S. *et al.* Assembling genomes and mini-metagenomes from chimeric MDA products. *J Comput Biol*. **20**, 714–737 (2013).
- Grant, J. R. & Stothard, P. The CGView Server: a comparative genomics tool for circular genomes. *Nucleic Acids Res*. **1**, 181–184 (2008).
- Harmsen, D. *et al.* Typing of methicillin-resistant *Staphylococcus aureus* in a university hospital setting by using novel software for *spa* repeat determination and database management. *J Clin Microbiol*. **41**, 5442–5448 (2003).
- Enright, M. C., Day, N. P., Davies, C. E., Peacock, S. J. & Spratt, B. G. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. *J Clin Microbiol*. **38**, 1008–1015 (2000).
- Strauß, L. *et al.* Detecting *Staphylococcus aureus* virulence and resistance genes: a comparison of whole-genome sequencing and DNA microarray technology. *J Clin Microbiol*. **54**, 1008–1016 (2016).
- Boyle-Vavra, S. *et al.* USA300 and USA500 clonal lineages of *Staphylococcus aureus* do not produce a capsular polysaccharide due to conserved mutations in the *cap5* locus. *MBio*. **6**, <https://doi.org/10.1128/mBio.02585-14> (2015).
- Sreedharan, S., Oram, M., Jensen, B., Peterson, L. R. & Fisher, L. M. DNA gyrase *gyrA* mutations in ciprofloxacin-resistant strains of *Staphylococcus aureus*: close similarity with quinolone resistance mutations in *Escherichia coli*. *J Bacteriol*. **172**, 7260–7262 (1990).
- Bradley, P. *et al.* Rapid antibiotic-resistance predictions from genome sequence data for *Staphylococcus aureus* and *Mycobacterium tuberculosis*. *Nat Commun*. **6**, <https://doi.org/10.1038/ncomms10063> (2015).

38. Uhlemann, A. C. *et al.* Molecular tracing of the emergence, diversification, and transmission of *S. aureus* sequence type 8 in a New York community. *Proc Natl Acad Sci USA* **111**, 6738–6743 (2014).
39. Gill, S. R. *et al.* Insights on evolution of virulence and resistance from the complete genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *Staphylococcus epidermidis* strain. *J Bacteriol.* **187**, 2426–2438 (2005).
40. Highlander, S. K. *et al.* Subtle genetic changes enhance virulence of methicillin resistant and sensitive *Staphylococcus aureus*. *BMC Microbiol.* **7**, 99 (2007).
41. Ge, B. *et al.* MRSA and multidrug-resistant *Staphylococcus aureus* in US retail meats, 2010–2011. *Food Microbiol.* **62**, 289–297 (2017).
42. Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics.* **11**, 119 (2010).
43. Emms, D. M. & Kelly, S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **16**, 157 (2015).
44. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J Mol Biol.* **215**, 403–410 (1990).
45. Pruitt, K., Brown, G., Tatusova, T. & Maglott, D. The Reference Sequence (RefSeq) Database Chapter 18. The NCBI Handbook (eds McEntyre, J., Ostell, J.) <https://www.ncbi.nlm.nih.gov/books/NBK21091/> (National Center for Biotechnology Information (US), 2002).

Acknowledgements

Part of the sequencing-work was funded by a grant awarded under the Horizon 2020 Framework Programme (H2020-MSCA-ITN-2015; grant agreement No. 675530) for the project “ANTibioticS and mobile resistance elements in WastEwater Reuse applications: risks and innovative solutions (ANSWER)”.

Author Contributions

Designed the project: S.H., B.S., R.L.M., F.A., W.R.; Performed the experiments: S.L.; Analysed the data: S.L., S.H., P.H., B.S., T.R., W.R.; Wrote the manuscript: S.L., P.H., F.A., W.R. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-27781-8>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as 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 images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018