Staphylococcal Enterotoxin Gene Cluster: Prediction of Enterotoxin (SEG and SEI) Production and of the Source of Food Poisoning on the Basis of \(v\)Sa\(\beta\) Typing


Research Division Food Microbial Systems, Agroscope, Bern, Switzerland
Laboratory for Food Safety, French Agency for Food, Environmental and Occupational Health & Safety (Anses), Université Paris-Est, Maisons-Alfort, France
Scientific Service Foodborne Pathogens, Sciensano, Brussels, Belgium
Université Paris-Saclay, CEA, INRAE, Département Médicaments et Technologies pour la Santé (DMTS), SPI, Gif-sur-Yvette, France
Université Paris-Saclay, INRAE, MaIAGE, Jouy-en-Josas, France

ABSTRACT
Currently, only 5 (SEA to SEE) out of 27 known staphylococcal enterotoxins can be analyzed using commercially available kits. Six genes (seg, sei, sem, sen, seo, and seu), encoding putative and undetectable enterotoxins, are located on the enterotoxin gene cluster (egc), which is part of the \(Staphylococcus aureus\) genomic island \(v\)Sa\(\beta\). These enterotoxins have been described as likely being involved in staphylococcal food-poisoning outbreaks. The aim of the present study was to determine if whole-genome data can be used for the prediction of staphylococcal egc enterotoxin production, particularly enterotoxin G (SEG) and enterotoxin I (SEI). For this purpose, whole-genome sequences of 75 \(S. aureus\) strains from different origins (food-poisoning outbreaks, human, and animal) were investigated by applying bioinformatics methods (phylogenetic analysis using the core genome and different alignments). SEG and SEI expression was tested \(in vitro\) using a sandwich enzyme-linked immunosorbent assay method. Strains could be allocated to 14 different \(v\)Sa\(\beta\) types, each type being associated with a single clonal complex (CC). In addition, the \(v\)Sa\(\beta\) type and CC were associated with the origin of the strain (human or cattle derived). The amount of SEG and SEI produced also correlated with the \(v\)Sa\(\beta\) type and the CC of a strain. The present results show promising indications that the \(in vitro\) production of SEG and SEI can be predicted based on the \(v\)Sa\(\beta\) type or CC of a strain.

IMPORTANCE
Besides having infectious properties in human and animals, \(S. aureus\) can produce different enterotoxins in food. The enterotoxins can cause vomiting and diarrhea, often involving many people. Most of these outbreaks remain undiscovered, as detection methods for enterotoxins are only available for a few enterotoxins but not for the more recently discovered enterotoxins G (SEG) and I (SEI). For this purpose, whole-genome sequences of 75 \(S. aureus\) strains from different origins (food-poisoning outbreaks, human, and animal) were investigated by applying bioinformatics methods (phylogenetic analysis using the core genome and different alignments). SEG and SEI expression was tested \(in vitro\) using a sandwich enzyme-linked immunosorbent assay method. Strains could be allocated to 14 different \(v\)Sa\(\beta\) types, each type being associated with a single clonal complex (CC). In addition, the \(v\)Sa\(\beta\) type and CC were associated with the origin of the strain (human or cattle derived). The amount of SEG and SEI produced also correlated with the \(v\)Sa\(\beta\) type and the CC of a strain. The present results show promising indications that the \(in vitro\) production of SEG and SEI can be predicted based on the \(v\)Sa\(\beta\) type or CC of a strain.

KEYWORDS
\(Staphylococcus aureus\), egc, enterotoxin

\(Staphylococcus aureus\) can produce a variety of heat-stable enterotoxins, which, when they are secreted in food, can cause staphylococcal food-poisoning outbreaks (SFPO). According to the European Food Safety Authority (EFSA), staphylococcal
enterotoxins (SE) in mixed foods and meat products are among the top 10 pathogen/food vehicle pairs, causing the highest number of hospitalizations in strong-evidence outbreaks. By looking at the number of cases, this trend seems to be on the rise (1). In addition, most of the SFPO are classified as weak-evidence outbreaks, since only the so-called classical enterotoxins (SEA, SEB, SEC, SED, and SEE) can be detected and quantified by commercially available kits (2). Besides these five well-known SE, another 20 have been described recently, and some of them were shown to have an emetic activity (SE) and, hence, could be involved in SFPO (3–9). Enterotoxins for which emetic activity has not yet been proved are considered staphylococcal enterotoxin-like (SEl) proteins. As not all SE can be detected directly in food, different methods have been applied in the past to better characterize the S. aureus strains involved in food-poisoning outbreaks, such as pulsed-field gel electrophoresis typing, PCR for detection of the enterotoxin genes, and other methods (10–14). These methods allowed us to evaluate the toxigenic profile of strains or to establish the link between strains and secreted tox- ins. With the recent advance of whole-genome sequencing (WGS), often each strain involved in an outbreak can be sequenced and characterized genetically, opening new doors to the understanding of the role different SE play in SFPO as well as prediction of antimicrobial resistance and infectivity (15–20).

Twenty years ago, a novel cluster of SE genes, the enterotoxin gene cluster (egc), was described containing the so-called new enterotoxins seg, sei, sem, sen, seo, and seu (21, 22). The egc is located on the genomic island vSaβ and is incorporated in the chromosome as a prophage (16). Literature suggests that about 50% of S. aureus strains harbor an egc (21, 23, 24).

For SEG, SEI, SEM, SEN, and SEO, emetic activity has been demonstrated, and it appears that some SFPO might be caused by these enterotoxins (3, 5). A lot is known on the expression of the classical SE (25–27), yet studies on the expression of the new SE are still very limited (28). Genetic backbones and regulatory systems of SE genes vary among S. aureus strains, causing diverse SE expression patterns. Hence, quantities of toxin production vary between strains (25–27).

Due to the lack of information, new methods and tools need to be developed to better understand and predict the expression and regulation mechanisms of the new enterotoxins, including those of the egc (29). For this reason, the aim of the present study was to determine whether WGS data can be used to predict staphylococcal enterotoxin production of the egc in vitro, particularly of SEG and SEI. These enterotoxins (SEG and SEI) were chosen because they are the only ones (of egc enterotoxins) for which a quantitative method for detection is currently available, allowing a direct link for the corresponding WGS data.

RESULTS

Strain characterization. Multilocus sequence typing (MLST) of the 75 S. aureus strains isolated from different sources, like food, humans, animals, and the environment, showed that the most frequently found clonal complexes (CC) are CC5 (n = 17), CC20 (n = 15), CC30 (n = 13), and CC705 (n = 11), followed by CC45, CC22, CC50, and CC9 (6, 3, 2, and 2 strains, respectively). In contrast, CC10, CC72, and CC121, as well as an unknown CC, were detected only once (Table 1).

The strains from the most frequently found CCs (CC5 and CC30) originated from a vast geographical range and were isolated from either human or food. In contrast, the CC20 and CC705 strains, always originating from France, Italy, and Switzerland, were isolated either from dairy products or bovine mastitis (Table 1).

spa typing of the 75 strains revealed that in most cases the strains belonging to a single CC were allocated to different spa types. Perfect agreement between CC and spa type was found only for CC705 (n = 11), where all strains were allocated to t529. For 15 strains, spa typing resulted in an unknown type, of which the majority belonged to CC30 and CC20 (5 and 6 unknown spa types, respectively).

Besides egc, the 75 strains also harbored other non-egc SE genes (Table 1). Indeed,
TABLE 1 Genotypic characteristics (i.e., clonal complex, enterotoxin genes present on the genome, \( \text{vSa}\beta \text{ type}, \text{ and spa type} \)) and origins of the 75 studied strains\(^a\)

<table>
<thead>
<tr>
<th>Strain</th>
<th>Country</th>
<th>Origin</th>
<th>Source of isolation</th>
<th>CC</th>
<th>Enterotoxin genes</th>
<th>( \text{vSa}\beta \text{ type} )</th>
<th>spa type</th>
</tr>
</thead>
<tbody>
<tr>
<td>07CEB94STA</td>
<td>Belgium</td>
<td>Food (SFPO)</td>
<td>Ready to eat</td>
<td>5</td>
<td>a, g, i, m, n, o</td>
<td>I</td>
<td>t704</td>
</tr>
<tr>
<td>11CEB145STA</td>
<td>Japan</td>
<td>Human</td>
<td>Infection</td>
<td>5</td>
<td>a, c, g, i, m, n, o, x</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>13CEB178STA</td>
<td>Ireland</td>
<td>SFPO</td>
<td>NA</td>
<td>5</td>
<td>d, j, g, i, m, n, o, r, x</td>
<td>I</td>
<td>t463</td>
</tr>
<tr>
<td>13CEB188STA</td>
<td>Ireland</td>
<td>Food (SFPO)</td>
<td>Milk product</td>
<td>5</td>
<td>g, i, m, n, o</td>
<td>I</td>
<td>t5829</td>
</tr>
<tr>
<td>13CEB191STA</td>
<td>Ireland</td>
<td>Food (SFPO)</td>
<td>Milk product</td>
<td>5</td>
<td>d, g, i, j, m, n, o, r, x</td>
<td>I</td>
<td>t837</td>
</tr>
<tr>
<td>13CEB329STA</td>
<td>Belgium</td>
<td>Human (SFPO)</td>
<td>Nose and throat</td>
<td>5</td>
<td>g, i, m, n, o</td>
<td>I</td>
<td>t7506</td>
</tr>
<tr>
<td>15SBC1507STA</td>
<td>Ireland</td>
<td>Food (SFPO)</td>
<td>Meat</td>
<td>5</td>
<td>g, i, m, n, o</td>
<td>I</td>
<td>t586</td>
</tr>
<tr>
<td>15SBC1550STA</td>
<td>Ireland</td>
<td>Food (SFPO)</td>
<td>Ready to eat</td>
<td>5</td>
<td>g, i, m, n, o</td>
<td>I</td>
<td>t450</td>
</tr>
<tr>
<td>17SBC108STA</td>
<td>France</td>
<td>Food (SFPO)</td>
<td>Meat</td>
<td>5</td>
<td>g, i, m, n, o</td>
<td>I</td>
<td>t111</td>
</tr>
<tr>
<td>17SBC907STA</td>
<td>France</td>
<td>Food (SFPO)</td>
<td>Meat</td>
<td>5</td>
<td>g, i, m, n, o</td>
<td>I</td>
<td>t7506</td>
</tr>
<tr>
<td>17SBC1585STA</td>
<td>Bulgaria</td>
<td>Food (SFPO)</td>
<td>Ready to eat</td>
<td>5</td>
<td>a, d, g, i, j, m, n, o, r, x</td>
<td>I</td>
<td>t535</td>
</tr>
<tr>
<td>17SBC1585STA</td>
<td>Bulgaria</td>
<td>Food (SFPO)</td>
<td>Ready to eat</td>
<td>5</td>
<td>a, d, g, i, j, m, n, o, r, x</td>
<td>I</td>
<td>t535</td>
</tr>
<tr>
<td>502A</td>
<td>USA</td>
<td>Human</td>
<td>Infection</td>
<td>5</td>
<td>g, i, m, n, o</td>
<td>I</td>
<td>t010</td>
</tr>
<tr>
<td>Mu50</td>
<td>Japan</td>
<td>Human</td>
<td>Infection</td>
<td>5</td>
<td>a, c, g, i, m, n, o, tst, x</td>
<td>I</td>
<td>t002</td>
</tr>
<tr>
<td>N315</td>
<td>Japan</td>
<td>Human</td>
<td>Human faeces</td>
<td>5</td>
<td>c, g, l, i, m, n, o, p, tst, x</td>
<td>I</td>
<td>t002</td>
</tr>
<tr>
<td>NZAK3</td>
<td>New Zealand</td>
<td>Human</td>
<td>Skin</td>
<td>5</td>
<td>c, g, i, l, m, n, o, p, x</td>
<td>I</td>
<td>t002</td>
</tr>
<tr>
<td>ST288</td>
<td>England</td>
<td>Human</td>
<td>Urine</td>
<td>5</td>
<td>g, i, m, n, o</td>
<td>I</td>
<td>t1003</td>
</tr>
<tr>
<td>18SBL779</td>
<td>Switzerland</td>
<td>Food (SFPO)</td>
<td>Milk product</td>
<td>9</td>
<td>g, i, m, n, o, u, x, y, 27</td>
<td>XIII</td>
<td>t899</td>
</tr>
<tr>
<td>G19F</td>
<td>Italy</td>
<td>Animal</td>
<td>Mastitis (cow)</td>
<td>9</td>
<td>g, i, m, n, u, x, y</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>13CEB177STA</td>
<td>Ireland</td>
<td>NA</td>
<td>FPO</td>
<td>10</td>
<td>c, g, i, m, n, o, u, x</td>
<td>XI</td>
<td>t704</td>
</tr>
<tr>
<td>11CEB277STA</td>
<td>Italy</td>
<td>Food</td>
<td>Milk product</td>
<td>20</td>
<td>g, i, m, n, o, u, x, y</td>
<td>XII</td>
<td>t325</td>
</tr>
<tr>
<td>11CEB279STA</td>
<td>Italy</td>
<td>environment</td>
<td>NA</td>
<td>20</td>
<td>g, i, m, n, o, u, x, y</td>
<td>XII</td>
<td>t948</td>
</tr>
<tr>
<td>15SBC1428STA</td>
<td>France</td>
<td>Food (SFPO)</td>
<td>Milk product</td>
<td>20</td>
<td>g, i, m, n, o, u, x, y</td>
<td>XII</td>
<td>t1544</td>
</tr>
<tr>
<td>17SBC202STA</td>
<td>France</td>
<td>Food (SFPO)</td>
<td>Milk product</td>
<td>20</td>
<td>g, i, m, n, o, u, x, y</td>
<td>XII</td>
<td>t164</td>
</tr>
<tr>
<td>17SBC208STA</td>
<td>France</td>
<td>Food (SFPO)</td>
<td>Milk product</td>
<td>20</td>
<td>g, i, m, n, o, u, x, y</td>
<td>XII</td>
<td>t10134</td>
</tr>
<tr>
<td>18SBL680</td>
<td>Switzerland</td>
<td>Food</td>
<td>Milk product</td>
<td>20</td>
<td>g, i, m, n, o, u, x, y</td>
<td>XII</td>
<td>t164</td>
</tr>
<tr>
<td>18SBL667</td>
<td>Switzerland</td>
<td>Food</td>
<td>Milk product</td>
<td>20</td>
<td>g, i, m, n, o, u, x, y</td>
<td>XII</td>
<td>t325</td>
</tr>
<tr>
<td>G11F</td>
<td>Switzerland</td>
<td>Animal</td>
<td>Mastitis (cow)</td>
<td>20</td>
<td>g, i, m, n, o, u</td>
<td>XII</td>
<td>t2736</td>
</tr>
<tr>
<td>13CEB1795STA</td>
<td>Ireland</td>
<td>NA</td>
<td>FPO</td>
<td>22</td>
<td>c, g, i, m, n, o, u, x</td>
<td>XVI</td>
<td>*</td>
</tr>
<tr>
<td>15SBC1517STA</td>
<td>Ireland</td>
<td>Food (SFPO)</td>
<td>Meat</td>
<td>22</td>
<td>c, g, i, l, m, n, u, x</td>
<td>XVI</td>
<td>t645</td>
</tr>
<tr>
<td>15SBC1527STA</td>
<td>Ireland</td>
<td>Food (SFPO)</td>
<td>Ready to eat</td>
<td>22</td>
<td>g, i, m, n, o, u</td>
<td>XVI</td>
<td>*</td>
</tr>
<tr>
<td>13CEB181STA</td>
<td>Ireland</td>
<td>Food (SFPO)</td>
<td>Ready to eat</td>
<td>30</td>
<td>a, g, i, m, n, o, u</td>
<td>III</td>
<td>t3018</td>
</tr>
<tr>
<td>13CEB312STA</td>
<td>Belgium</td>
<td>Food (SFPO)</td>
<td>Ready to eat</td>
<td>30</td>
<td>a, g, i, m, n, o</td>
<td>III</td>
<td>t022</td>
</tr>
<tr>
<td>13CEB313STA</td>
<td>Belgium</td>
<td>Human (SFPO)</td>
<td>Human faeces</td>
<td>30</td>
<td>g, i, m, n, o</td>
<td>III</td>
<td>*</td>
</tr>
<tr>
<td>13CEB317STA</td>
<td>Belgium</td>
<td>Human (SFPO)</td>
<td>Nose and throat</td>
<td>30</td>
<td>g, i, m, n, o</td>
<td>III</td>
<td>*</td>
</tr>
<tr>
<td>13CEB318STA</td>
<td>Belgium</td>
<td>Human (SFPO)</td>
<td>Nose and throat</td>
<td>30</td>
<td>g, i, m, n, o</td>
<td>III</td>
<td>*</td>
</tr>
<tr>
<td>13CEB327STA</td>
<td>Belgium</td>
<td>Human (SFPO)</td>
<td>Nose and throat</td>
<td>30</td>
<td>g, i, m, n, o</td>
<td>III</td>
<td>*</td>
</tr>
<tr>
<td>13CEB328STA</td>
<td>Belgium</td>
<td>Human (SFPO)</td>
<td>Nose and throat</td>
<td>30</td>
<td>g, i, m, n, o</td>
<td>III</td>
<td>*</td>
</tr>
<tr>
<td>18SBC761</td>
<td>Switzerland</td>
<td>Food (SFPO)</td>
<td>Milk product</td>
<td>30</td>
<td>g, i, m, n, o, u, x, z</td>
<td>III</td>
<td>t021</td>
</tr>
<tr>
<td>18SBC765</td>
<td>Switzerland</td>
<td>Food</td>
<td>Ready to eat</td>
<td>30</td>
<td>g, i, m, n, o, u, x, z</td>
<td>III</td>
<td>t166</td>
</tr>
<tr>
<td>18SBC768</td>
<td>Switzerland</td>
<td>Food</td>
<td>Ready to eat</td>
<td>30</td>
<td>g, i, m, n, o, u</td>
<td>III</td>
<td>t021</td>
</tr>
<tr>
<td>ATCC 25923</td>
<td>USA</td>
<td>Human</td>
<td>Skin</td>
<td>30</td>
<td>g, i, m, n, o, u</td>
<td>III</td>
<td>t021</td>
</tr>
<tr>
<td>K590</td>
<td>Switzerland</td>
<td>Food (SFPO)</td>
<td>Ready to eat</td>
<td>30</td>
<td>g, i, m, n, o</td>
<td>III</td>
<td>t021</td>
</tr>
<tr>
<td>MRSA252</td>
<td>USA</td>
<td>Human</td>
<td>Infection</td>
<td>30</td>
<td>g, i, m, n, o</td>
<td>III</td>
<td>t018</td>
</tr>
<tr>
<td>07CEB90STA</td>
<td>Belgium</td>
<td>Food (SFPO)</td>
<td>Ready to eat</td>
<td>45</td>
<td>c, g, i, m, n, o, u</td>
<td>XXII</td>
<td>t1040</td>
</tr>
<tr>
<td>18SBC677</td>
<td>Switzerland</td>
<td>Food</td>
<td>Ready to eat</td>
<td>45</td>
<td>g, i, m, n, o, u</td>
<td>XXII</td>
<td>t505</td>
</tr>
<tr>
<td>18SBC679</td>
<td>Switzerland</td>
<td>Food (SFPO)</td>
<td>Milk product</td>
<td>45</td>
<td>g, i, m, n, o</td>
<td>XXII</td>
<td>t015</td>
</tr>
<tr>
<td>18SBC677</td>
<td>Switzerland</td>
<td>Food (SFPO)</td>
<td>Milk product</td>
<td>45</td>
<td>g, i, m, n, o</td>
<td>XXII</td>
<td>t015</td>
</tr>
<tr>
<td>USA600</td>
<td>USA</td>
<td>Human</td>
<td>Infection</td>
<td>45</td>
<td>g, i, m, n, o</td>
<td>XXII</td>
<td>t004</td>
</tr>
<tr>
<td>18SBC672</td>
<td>Switzerland</td>
<td>Food</td>
<td>Milk product</td>
<td>50</td>
<td>i, m, n, o, u, x, z</td>
<td>XXI</td>
<td>t246</td>
</tr>
<tr>
<td>GN3</td>
<td>Japan</td>
<td>Human</td>
<td>NA</td>
<td>50</td>
<td>i, m, n, o</td>
<td>XXI</td>
<td>t185</td>
</tr>
<tr>
<td>13CEB323STA</td>
<td>Belgium</td>
<td>Human (SFPO)</td>
<td>Nose and throat</td>
<td>72</td>
<td>c, x, g, i, m, n, o, u, x</td>
<td>XIX</td>
<td></td>
</tr>
<tr>
<td>05CEB52STA</td>
<td>NA</td>
<td>Human</td>
<td>FPO</td>
<td>121</td>
<td>b, g, i, m, n, o, u, y, x</td>
<td>XIX</td>
<td>*</td>
</tr>
</tbody>
</table>

(Continued on next page)
from genome assembly, all 27 SE genes were detected in one of the strains at least once, yet it is noteworthy that the five strains belonging to CC5 often carried additional SE genes, such as selx (in 4 strains), sea (in 3 strains), and a plasmid containing sed, selx, and ser (in 2 strains). Furthermore, CC30 (n = 13) harbored sea in 6 strains and tst (toxic-shock toxin) in 2 strains.

CC705 was comprised of sec, tst, selx, and sel, whereas CC20 often carried selx and sly (in 14 and 11 out of 15 strains, respectively).

Allocation of the strains to their \(v\)Sa\(\beta\) types and diversity of SEG and SEI. In 59 of 75 strains (79%), the \(v\)Sa\(\beta\) type could be allocated to an existing one with overall similarities of \(\geq 90\%\). For the remaining 16 strains, new \(v\)Sa\(\beta\) types were defined by numbering continuously from XVI onward (Fig. 1), resulting in seven new \(v\)Sa\(\beta\) types (XVI to XXII). Three strains were allocated to \(v\)Sa\(\beta\) type XVI, two strains to \(v\)Sa\(\beta\) type XXI. One strain carrying \(v\)Sa\(\beta\) type IV, XI, XII, and XXII were only found in humans (including infections) and food isolates. No animal strains harbored these \(v\)Sa\(\beta\) types.

The seven newly defined \(v\)Sa\(\beta\) types (Fig. 1) all contained, in addition to the egc genes, virulence-associated and hypothetical genes. \(v\)Sa\(\beta\) types XVII and XVIII carry bacteriocins and serine proteases, whereas \(v\)Sa\(\beta\) type XIX was notably (approximately 20,000 bp) longer than the other \(v\)Sa\(\beta\) types and carried numerous genes coding for hypothetical proteins. \(v\)Sa\(\beta\) type XXII was shorter than all other \(v\)Sa\(\beta\) types (approximately 13,000 bp) and did not carry any additional virulence-associated genes besides the egc genes.

Within each \(v\)Sa\(\beta\) type, an amino acid identity of 100% for each SE was observed. However, SE differences were observed among different \(v\)Sa\(\beta\) types (Table 2). Among all strains included in the study, the SEG amino acid similarity varied between 96% and 100%, with a maximum of 9 amino acids of difference, compared to strain Mu50 (reference). For SEI, the similarity varied between 93% and 100%, with a maximum of 9 amino acids of difference, compared to strain Mu50 (reference).

Phylogenetic analysis of the core genome. To evaluate the evolutionary relationship of \(S\). \(a\)ureus strains included in the present study, their phylogeny was evaluated based on their core genomes. The tree shows a perfect concordance between the phylogenetic clades, CCs, and \(v\)Sa\(\beta\) type of the strain (Fig. 2). For \(v\)Sa\(\beta\) type IV, XI, XII, and XIII, a perfect concordance was observed between strains isolated from milk products, and animal mastitis can be observed (no human strains harbored these \(v\)Sa\(\beta\) types). On the other side, strains harboring \(v\)Sa\(\beta\) type I, III, and XXII were only found in humans (including infections) and food isolates. No animal strains harbored these \(v\)Sa\(\beta\) types. SFPO strains were found in every \(v\)Sa\(\beta\) type.

Enterotoxin production. SEG production ranged from below the limit of detection (LOD; 0.001 ng/ml) to 4.26 \(\pm\) 0.78 ng/ml, with a median of 1.17 ng/ml. SEG production below the LOD (0.001 ng/ml) was observed for \(v\)Sa\(\beta\) IV and XXI. One strain carrying \(v\)Sa\(\beta\) III (18SBCL675) showed nondetectable quantities of SEG, whereas the other two

### Table 1 (Continued)

<table>
<thead>
<tr>
<th>Strain</th>
<th>Country</th>
<th>Origin</th>
<th>Source of isolation</th>
<th>CC</th>
<th>Enterotoxin genes</th>
<th>(v)Sa(\beta) type</th>
<th>spa type</th>
</tr>
</thead>
<tbody>
<tr>
<td>18SBCL669</td>
<td>Switzerland</td>
<td>Food</td>
<td>Milk product</td>
<td>479</td>
<td>d, g, i, m, n, o, u, x</td>
<td>XI</td>
<td>t7013</td>
</tr>
<tr>
<td>G68P</td>
<td>Switzerland</td>
<td>Animal</td>
<td>Mastitis (cow)</td>
<td>479</td>
<td>g, i, m, n, o, u</td>
<td>XI</td>
<td>t7013</td>
</tr>
<tr>
<td>13CEB182STA</td>
<td>Ireland</td>
<td>Food (SFPO)</td>
<td>Milk-product</td>
<td>705</td>
<td>c, i, m, n, o, tst, u, x</td>
<td>IV</td>
<td>t529</td>
</tr>
<tr>
<td>13CEB190STA</td>
<td>Ireland</td>
<td>Food (SFPO)</td>
<td>Milk-product</td>
<td>705</td>
<td>c, i, m, n, o, tst, u, x</td>
<td>IV</td>
<td>t529</td>
</tr>
<tr>
<td>15SBCL1438STA</td>
<td>France</td>
<td>Food (SFPO)</td>
<td>Milk-product</td>
<td>705</td>
<td>c, i, m, n, o, tst, u, x</td>
<td>IV</td>
<td>t529</td>
</tr>
<tr>
<td>18SBCL670</td>
<td>Switzerland</td>
<td>Food</td>
<td>Milk-product</td>
<td>705</td>
<td>c, i, m, n, o, u, x, y, y</td>
<td>IV</td>
<td>t529</td>
</tr>
<tr>
<td>M1280</td>
<td>Switzerland</td>
<td>Animal</td>
<td>Mastitis (cow)</td>
<td>705</td>
<td>c, i, m, n, o, u</td>
<td>IV</td>
<td>t529</td>
</tr>
<tr>
<td>M1655</td>
<td>Switzerland</td>
<td>Animal</td>
<td>Mastitis (cow)</td>
<td>705</td>
<td>c, i, m, n, o, u</td>
<td>IV</td>
<td>t529</td>
</tr>
<tr>
<td>M2323</td>
<td>Switzerland</td>
<td>Animal</td>
<td>Mastitis (cow)</td>
<td>705</td>
<td>c, i, m, n, u, o, tst, u</td>
<td>IV</td>
<td>t529</td>
</tr>
<tr>
<td>M2662</td>
<td>Switzerland</td>
<td>Animal</td>
<td>Mastitis (cow)</td>
<td>705</td>
<td>c, i, m, n, u, o</td>
<td>IV</td>
<td>t529</td>
</tr>
<tr>
<td>M2839</td>
<td>Switzerland</td>
<td>Animal</td>
<td>Mastitis (cow)</td>
<td>705</td>
<td>c, i, n, o, tst, u</td>
<td>IV</td>
<td>t529</td>
</tr>
<tr>
<td>M3783</td>
<td>Switzerland</td>
<td>Animal</td>
<td>Mastitis (cow)</td>
<td>705</td>
<td>i, m, n, o, u</td>
<td>IV</td>
<td>t529</td>
</tr>
<tr>
<td>RF122</td>
<td>Ireland</td>
<td>Animal</td>
<td>Mastitis (cow)</td>
<td>705</td>
<td>c, i, m, n, o, u, tst, x, y, z</td>
<td>IV</td>
<td>t529</td>
</tr>
<tr>
<td>175BCL13STA</td>
<td>France</td>
<td>Food (SFPO)</td>
<td>Meat</td>
<td>**</td>
<td>a, g, i, m, n, o, x</td>
<td>XVIII</td>
<td>t13785</td>
</tr>
</tbody>
</table>

*NA, data not available; SFPO, food poisoning outbreak; *, unknown spa type; **, unknown clonal complex (CC).
strains harboring vSaβ III had values between 0.26 ± 0.01 and 0.78 ± 0.13 ng/ml. All the other strains showed values between 0.80 ± 0.11 and 4.26 ± 0.78 g/ml. By visual data inspection (Fig. 3A), two levels of SEG production can be distinguished: 9 strains that generated low (L) and 23 strains that produced high (H) concentrations of SEG. The median concentration for the L producer was 0 ng/ml (minimum [min], 0 ng/ml; maximum [max], 0.26 ± 0.01 ng/ml) and for the H producer was 1.42 ± 0.14 ng/ml (min, 0.783 ± 0.13 ng/ml; max, 4.26 ± 0.78 ng/ml). The difference between medians was highly significantly (P < 0.001).

**TABLE 2** Amino acid similarity of SEG and SEI compared to the reference strainsα (Mu50 and vSaβ type I)

<table>
<thead>
<tr>
<th>vSaβ type</th>
<th>SEG</th>
<th>SEI</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>100α</td>
<td>100α</td>
</tr>
<tr>
<td>III</td>
<td>97</td>
<td>95</td>
</tr>
<tr>
<td>IV</td>
<td>*</td>
<td>95</td>
</tr>
<tr>
<td>XI</td>
<td>97</td>
<td>93</td>
</tr>
<tr>
<td>XII</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>XIII</td>
<td>99</td>
<td>100</td>
</tr>
<tr>
<td>XVI</td>
<td>100</td>
<td>99</td>
</tr>
<tr>
<td>XVII</td>
<td>100</td>
<td>99</td>
</tr>
<tr>
<td>XVIII</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>XIX</td>
<td>97</td>
<td>93</td>
</tr>
<tr>
<td>XX</td>
<td>99</td>
<td>100</td>
</tr>
<tr>
<td>XXI</td>
<td>*</td>
<td>97</td>
</tr>
<tr>
<td>XXII</td>
<td>100</td>
<td>99</td>
</tr>
</tbody>
</table>

*Each vSaβ type sequence is represented based on 100% intergroup similarity. Superscript R, reference; *, gene absent.
FIG 2 Maximum likelihood phylogenetic tree based on the core genome (nucleotidic sequences) showing the evolutionary relationship among 75 isolates of *Staphylococcus aureus* (all strains positive for the enterotoxin gene cluster) recovered from human, animal, environment, and food samples (left). At the right, for each strain its clonal complex (CC), origin of the strain, source of the strain, and involvement in staphylococcal food poisoning outbreak (SFPO) is given. Bootstrap values of >80 are shown. Production of enterotoxin G (SEG) and I (SEI) for the 32 analyzed strains is also given (last two columns). These are shown as L for low enterotoxin production and H for high enterotoxin production. *, statistical outliers; **, unknown CC.
The amount of SEI produced (LOD, 0.037 ng/ml) by the strains ranged from 1.06 ± 0.17 ng/ml to 61.43 ± 10.29 ng/ml (median, 14.31 ng/ml) (Table 1 and Fig. 3B). According to their SEI production, strains could again be visually allocated to two different levels, L producers (producing 1.06 ± 0.17 to 3.85 ± 0.99 ng/ml; median, 2.22 ng/ml) and H producers (10.77 ± 1.22 to 61.43 ± 10.29 ng/ml; median, 21.51 ng/ml). The L strains belonged to the vSaβ types III, IV, and XI, whereas the H strains belonged to vSaβ types I, XII, XIII, XVI, XVII, XVIII, XIX, XX, XXI, and XXII (P < 0.001 between L and H).

To assess a possible relationship between SEG and SEI production, first a robust linear regression (see Fig. S1 and S2 in the supplemental material) was performed,
identifying four outliers (G68P, 18SBCL669, Mu50, and 05CEB52). These outliers were not taken into consideration for a second, ordinary least-square linear regression analysis (Fig. S3). This regression was modeled to \[ \text{SEI} = 15.49 \times \text{SEG} + 0.63 \], with \( R = 0.940 \) \( (P < 0.001) \), where brackets indicate the SE concentrations in nanograms per milliliter.

**DISCUSSION**

In the present study, we demonstrate that SEG and SEI production in vitro can be predicted using genomic data. In fact, there are strong indications that the amount these SE produced depends on the \( vSa\beta \) type. Furthermore, with the analysis and findings described here, it is now possible to infer the origin of an egc-containing \( S.\ aureus \) strain (human derived, cattle derived) that is involved in an SFPO. As the \( vSa\beta \) type is perfectly linked to the CC of a strain, as shown in the present study and in a previous report from Kläui et al. (30), the SE production and the origin of the SFPO also can be predicted based on the CC of the strain obtained by MLST, a typing method that is well established.

Previous studies already demonstrated that different strains can produce different amounts of SE, but in most cases the link to the genome was missing (31, 32).

In this study, the focus was on the egc enterotoxins that, according to previous studies (3, 12, 33, 34), are harbored by about 50% of \( S.\ aureus \) strains. The importance of the egc enterotoxins regarding food safety has been shown by Johler et al. (3), who described the probable egc enterotoxins’ involvement in foodborne outbreaks. However, strong evidence could not be confirmed, as the enterotoxin measurement in the food and from the bacteria could not be performed due to lack of appropriate methods. This could also be the reason why a lot of egc-caused SFPO remain undiscovered. In this study, for two enterotoxins (SEG and SEI) out of the five egc enterotoxins, an enzyme-linked immunosorbent assay (ELISA) method was available, whereas for the other egc enterotoxins this still is not the case. \( seu \) was not considered at all, as there is no literature demonstrating its emetic activity. Due to this lack of information about egc enterotoxins, new methods and tools need to be developed to better understand and predict their expression and regulation mechanisms (29). As a consequence, the aim of the present study was to determine whether WGS data can be used to predict staphylococcal enterotoxin production of the egc in vitro, particularly of SEG and SEI.

**Prediction of SEG and SEI production in vitro.** For the present study, 75 strains were chosen, originating from both human hosts and animal (cattle) as well as from environmental and food sources, with special attention on SFPO strains (35). Out of the 75 strains, 60 were allocated to the 15 previously defined \( vSa\beta \) types (30). The remaining 15 strains could be grouped into 7 newly defined \( vSa\beta \) types (Fig. 1). According to these new insights, using the \( vSa\beta \) types seems to be a very precise tool to characterize the different egc present in \( S.\ aureus \) strains instead of using egc types I to VI, as has been described previously (14, 21, 22, 36, 37).

The present study shows that there are two groups of SE producers, strains that produce low levels of SEG and SEI and strains with increased SE production (for both, SE \( P < 0.001 \)). A special case is the absence of SEG production for \( vSa\beta \) IV and XXI. This is explained by the fact that both had a truncated seg gene, resulting in an incomplete, nondetectable protein.

A very high linear dependency was observed between the production of SEG and SEI (\( R = 0.98, P < 0.0001 \)), while the amount of SEI measured was approximately 16 times higher than that of SEG. The high correlation between SEG and SEI production suggests that both SE are regulated primarily by the same transcription factor as that proposed by Kusch et al. (38). This hypothesis, however, neglects the fact that the SEG production is 16\( \times \) lower than that for SEI, accounting for a fine tuning by additional transcription factors, as observed for other SE (38, 39).

During the first robust linear regression analysis, outliers were observed (G68P, 18SBCL669, and 05CEB52). For these strains, all members of \( vSa\beta \) types XI and XIX, the production of SEI was always lower than SEG production (see Fig. S2 in the
supplemental material). As demonstrated in Table 2, SEI of both vSaβ types showed the lowest similarity (93%) compared to the reference (Mu50). These findings indicate that the monoclonal antibody used for the present study matches incompletely with the SEI epitopes produced by vSaβ type XI- and XIX-producing strains, resulting in a reduced detection of SEI quantities. Besides the technical aspect, it cannot be ruled out; however, regulation of SEI production is special for these vSaβ types. To clarify these ambiguities, additional studies are required.

The results of the present European study were not in agreement with the results published by Omoe et al. (40), who detected SEI in only 40% of the strains and SEG was not detected at all. In our study, SEG was produced by 96% of the strains and SEI for 100% of the strains, being positive for the two enterotoxin genes detected by NAuRa (35). Only for one strain (18SBCL678) was seg predicted, but SEG enterotoxin was not detected. As our results were generated from a large variety of strains, the involvement of the egc enterotoxins in SFPO should be reconsidered.

Infering the origin of an SFPO-involved strain. Looking at the major vSaβ types found in this study (I, III, IV, XII, and XXII), it was observed that in each group there are SFPO-associated strains (isolated from food) and strains that are human (infection) or cattle (mastitis) derived but never both for the same vSaβ type.

In addition to our previous study with 15 allocated vSaβ type observed (30), we found 9 new types. Again, a perfect concordance between vSaβ type and CC was found, confirming this observation as a general principle in *S. aureus*. This principle can now be applied for evaluation of egc-containing strains involved in SFPO. In fact, instead of inferring the vSaβ type involved in the SFPO, the common and simpler method of CC assessment can be performed. This is particularly easy for WGS data, as the reads can be directly uploaded to an Internet app, such as cge.cbs.dtu.dk, for inferring of the sequence type (ST), which is then used together with the pubMLST database program (41) to obtain the corresponding CC. If WGS data are not available, the standard MLST procedure can be performed using standard PCR and Sanger sequencing for the seven housekeeping genes (*arcC, aroE, glpF, gmk, pta, tpiA*, and *yqiL*) (42). Instead of the original primers (35), the newly designed primer by Boss et al. (43) can be applied. They enable unidirectional Sanger sequencing, which considerably saves cost, work, and time.

The suspected reason for this strong link between CC and vSaβ type is that vSaβ acquisition by horizontal transfer in the *S. aureus* genome happened immediately before or simultaneously to clonal diversification of *S. aureus* (30). This hypothesis is also supported by the phylogenetic analysis of the core genomes of the present study (Fig. 2), showing a perfect agreement between the phylogenetic clade, CC, and vSaβ type.

The CC can be used to perform an association of an egc-carrying SFPO strain to a specific origin (human or cattle). As can be seen in Fig. 2, CC705 and CC20 are strains strictly associated with bovine mastitis and dairy products. In fact, CC705-positive strains are classical pathogens of bovine mastitis observed in- and outside of Europe (43–45). In addition, they are also frequently present in delivered milk (43) and cheese (46). CC705 strains are uniquely positive for *spa* type t529 (Table 1) and are typical colonizers of bovine skin as well as infections of bovine wounds (47). Similar findings are also true for CC20-positive strains. These can also cause bovine mastitis and are present in delivered milk, but they are less abundant than CC705 (43).

On the other hand, strains allocated to CC5, CC30, and CC45 were exclusively isolated from human samples (infection, skin, feces, nose, and throat) and from food (Fig. 2), where human handling was very likely (ready-to-eat products). Furthermore, these CCs are widely described in the literature as being found in human infections (48–51). This is a further advantage of CC nomenclature as literature about them is broad (23, 52, 53), enabling us to extend the scope beyond an egc enterotoxin-caused SFPO.

Application of new insights in evaluation of egc-caused SFPO. The involvement of egc enterotoxins in foodborne outbreaks is highlighted by the fact that vSaβ types (and the corresponding CC) from *S. aureus* strains producing high levels of SEG and SEI
are also described to be involved in foodborne outbreaks, especially CC5, CC20, and CC45 (23, 34, 53). Furthermore, certain strains of CC45 (harboring egc) do not harbor any classical enterotoxin (34, 54), yet these strains could have been involved in foodborne outbreaks.

As an example, we deal with strain 18SBC673, which was involved in a foodborne outbreak related to artisanal goat cheese in southern Switzerland (54) and was included in the present study. It is characterized by the presence of just egc enterotoxins, as shown by NAuRa, and produces a substantial amount of SEG (2.04 ± 0.33 ng/ml) and SEI (40.58 ± 9.03 ng/ml). It is positive to vSaβ type XXII and CC45. As the strain had been isolated from goat cheese, it could be hypothesized that goat milk was the probable source. However, according to the present study (Fig. 2), it is clear that the origin of the involved strain is, with a high probability, human. As a consequence, the SFPO caused by this strain was a highly human contamination during cheese manufacturing. This conclusion is supported by the fact that CC45 is never found in goats and goat milk (55, 56).

Conclusions. The presented study demonstrates that the in vitro production of SEG and SEI can be predicted based on the vSaβ type and the CC of a strain. Furthermore, the vSaβ type/CC enables us to predict the source of an egc-positive SFPO strain (animal or human derived). Due to the perfect correlation between CC and vSaβ type, the use of common CC typing is an easy and quick way to characterize a strain involved in an SFPO. Therefore, it is a good alternative to the proposed egc typing (I to IV), a method that results in only four biologically irrelevant types.

This information will enhance the ability to better understand the involvement of the egc enterotoxins in SFPO. The fact that the egc is found in more than 50% of the S. aureus strains and, according to our study, exactly 75% expressed SEG and 100% expressed SEI are further indications that these and other egc enterotoxins are involved in SFPO.

MATERIALS AND METHODS

Bacterial strain and genome collection. The general aim was to use egc-harboring S. aureus strains representing a large diversity in their genomes and origins. To achieve this, 75 strains and genomes from different sources (food, environment, animal, and human) as well as different geographical origins were chosen (Table 1). SFPO genome sequences and strains (42 genomes and strains) were obtained from the collection of the European National Reference Laboratory for Coagulase-Positive Staphylococci (EURL CPS; Maisons-Alfort, France). Nine Swiss bovine mastitis strains were used from the Agroscope strain and genome collection; these strains were sampled previously by Fournier et al. (57) and their genome sequenced by Kläui et al. (30). For genomic and phylogenetic analysis, seven strains of human and animal origins were obtained from NCBI (reference sequence database; https://www.ncbi.nlm.nih.gov) to increase the sample size and variation of the strains. Two strains (Mu50 and N315) were obtained from P. Moreillon (University of Lausanne). Thirteen egc-containing strains were isolated from food in Switzerland (details are described below). An overview of the bacterial strain collection used in study is listed in Table 1.

Characterization of Swiss food strains. Forty-five Swiss S. aureus strains originating from food were obtained from the Federal Food Safety and Veterinary Office (kindly provided by A. Baumgartner). The presence of egc genes in these strains was determined by applying a real-time PCR assay with melting curve analysis (mPCR) for detection of seg, sei, sem, sen, and seo. For detection of seg and sei, primers and PCR conditions were applied as described by Cosandey et al. (58). For detection of sem, sen, and seo, new primers were designed (Table 3). The detection of seu was omitted, as its emetic activity has not been shown so far. After being cultured at 37°C for 24 h on blood agar (bioMérieux Suisse s.a., Geneva, Switzerland), DNA was extracted from single colonies of S. aureus. A colony was picked and resuspended in 100 µl of 10 mM Tris-HCl and 10 mM EDTA (pH 8.5), incubated at 95°C for 10 min, and immediately stored on ice. The lysates were diluted 1:100 in H2O to be used as templates for the different mPCRs (43). For all mPCRs, the total volume was 20 µl, containing 300 nM corresponding forward and reverse primer (Table 3), 1 × Kapa Sybr Fast (Kapa Biosystems Inc., Woburn, MA), and 2.5 µl of 1:100 diluted DNA template. The mPCR run began with an initial denaturation at 95°C for 3 min, followed by 35 cycles of denaturation at 95°C for 3 s, annealing and extension in a single step at 60°C for 50 s, and a final extension step at 60°C for 5 min. Melting of the amplicons was performed between 60°C and 94°C, with increments of 1°C and a 5-s waiting time at each step. The mPCRs were performed using a Rotor-Gene 6000 real-time thermal cycler (Corbett Life Science, Mortlake, Australia).

Primer specificity (Table 3) was tested with S. aureus strains that were previously sequenced, namely, G11F, G19P, M1280, M1655, M2323, M2682, M3783, Mu50, and N315 (Table 1).

Applying the mPCR for detection of the egc genes showed that only 40% of the strains were egc positive. Based on the diversity of their origins, 14 egc-positive strains were selected.
These Swiss strains, isolated from food, were sequenced as follows. Strains were cultured at 37°C for 24 h on blood agar, 3 to 4 colonies were suspended in 4.5 ml tryptic soy broth (TSB; Becton, Dickinson), and incubated 18 h (37°C, with shaking). From this overnight culture (ONC), 1 ml was suspended in 500 ml TSB and incubated under the same conditions. The resulting ONC was centrifuged for 23 min (7°C, 6,000 × g) (Celssep 6/720R; Henderson Biomedical Ltd., Lower Sydenham, UK). The supernatant was discarded and the pellet resuspended in 15 ml 10 mM Tris-HCl, pH 7.8, and transferred to a falcon tube, and incubated 18 h (37°C, 13,500 × g). After centrifugation, the pellet was treated using the NucleoBond Xtra Maxi kit (Machery Nagel, Düren, Germany) according to the manufacturer’s protocol, with the following modifications: instead of resuspending the pellet directly in 24 ml RES (from the kit), the pellet was resuspended in 2 ml RES containing 350 mg glass beads (425 to 600 μm; Merck, Darmstadt, Germany) and shaken on a Bead Ruptor at level 6 for 45 s (Bead Ruptor Elite; Omni International, Kennesaw, GA, USA). After centrifugation for 5 min (4°C, 13,500 × g), 22 ml was added to the supernatant, and DNA was then extracted according to the protocol of the manufacturer of the kit. The pellet was resuspended in 200 μl ddH₂O (double-distilled water) and further purified by applying the High Pure PCR template preparation kit protocol (Roche, Basel, Switzerland). DNA quality was considered sufficient if the optical density at 260 nm (OD₂₆₀/OD₂₈₀) was ≧1.8 and OD₂₆₀/OD₂₃₀ was ≧1.9 (measured with a QuickDrop spectrophotometer; Molecular Devices, San Jose, CA). The extracted DNA (representing the whole genome) was sequenced by an Illumina HiSeq at Eurofins GATC (Constance, Germany), generating more than 1.5 Gb of reads.

**Bioinformatics.** The reads from the strains from EURL CPS were obtained from the European Nucleotide Archive database (https://www.ebi.ac.uk/ena). For these reads and the reads from the Swiss food strains, the method for assembly and annotation was applied according to Merda et al. (35). Before the assembly, reads were normalized using BBnorm (https://jgi.doe.gov/data-and-tools/bbtools/) to have a maximum coverage of 100 ×. Normalized reads were trimmed using Trimmomatic (59). Quality filtering then was performed, removing reads shorter than 50 bp as well as excluding bases having a Phred quality score lower than 30. With these filtered reads, assembly was performed in three steps: (i) a de novo assembly was generated using SPAdes (v.3.9.1) (60) applying the default parameters, (ii) scaffold folding was performed in MeDuSa (61), using the nearest complete public genome of *S. aureus* estimated by Mash (62), and (iii) gaps were closed using GMcloser (63). The quality of each assembled genome was assessed with QUAST (v.4.3) (64). The assemblies were annotated using Prokka (v.1.11) (65) and RAST (66) for the prediction of coding sequences (CDSs).

**MLST, spa type, and *vSaβ* type allocation.** For all 75 genomes used in this study, three typing methods were applied to further characterize the strains genomically: (i) multilocus sequence typing (MLST), (ii) spa typing, and (iii) *vSaβ* typing. The MLST of the seven housekeeping genes (67) and spa typing (68) were done by using the Center for Genomic Epidemiology online platform (http://www.genomicepidemiology.org/). In the pubMLST database program (41), the sequence types (STs) from MLST were used to allocate each strain to a CC. For ST504 in the actual pubMLST database, no corresponding CC is available; as a consequence, this ST was allocated to CC705, as also described in the literature (43). *vSaβ* islands were identified in the genome by applying the method described by Kläui et al. (30). Briefly, if the *vSaβ* island of a strain had a sequence similarity of ≥90% to the reference strain of any existing *vSaβ* type, it was considered of the same type (30). If the sequence similarity was <90%, the *vSaβ* island was defined as a new type. All alignments were performed by using the Needleman-Wunsch algorithm of Clone Manager Professional 9 software (Scientific & Educational Software, Denver, CO).

**Enterotoxin gene profiles.** The enterotoxin gene profiles of the *S. aureus* strains, based on the WGS, were determined using the NAuRA tool (https://github.com/afelten-Anses/NAuRA). The screening of the enterotoxins was performed using the gene sequence and their relative protein sequence of the already described 27 SE and the estimated parameters of BLAST by Merda et al. (35). The screening of the enterotoxins was performed using the gene sequence and their relative protein sequence of the already described 27 SE and the estimated parameters of BLAST by Merda et al. (35).

**Phylogenetic analysis.** The core genome of each of the 75 strains was determined by the Roary pipeline (69). For this, previously obtained GFF3 file from Prokka was used as an input containing all of the strains’ genes as detected by the software. All genes of a strains’ core genome were then concatenated. A multiple-sequence alignment (MSA) (using MAFFT [70]) was performed using the concatenated core genomes of all the strains. The MSA then was imported into the Gblock program (71) for quality checking using the default setup and removing any misaligned regions. A phylogenetic tree was constructed using the maximum-likelihood method in IQtree (72). This program estimated the evolutionary model of sequences, and the best model, according to Akaike criterion, was GTR + I + gamma.

### Table 3: Primers for detection of enterotoxin genes developed and used in this study.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer*</th>
<th>Sequence 5′–3′</th>
<th>Amplicon size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>seb</em></td>
<td>Gsem_S</td>
<td>GATGTCGGAGGTITTGAATCTTA</td>
<td>584</td>
</tr>
<tr>
<td></td>
<td>Gsem_AS</td>
<td>ACTTTACGGTTGCGCTGTT</td>
<td></td>
</tr>
<tr>
<td><em>sen</em></td>
<td>Gsen_S</td>
<td>TTCTTCCAGTTAAGGCTACACA</td>
<td>218</td>
</tr>
<tr>
<td></td>
<td>Gsen_AS</td>
<td>CTGATAAAGGTCGCAATTAG</td>
<td></td>
</tr>
<tr>
<td><em>seo</em></td>
<td>Gseo_S</td>
<td>TAAACGCGATTGCTGCGATGCAG</td>
<td>348</td>
</tr>
<tr>
<td></td>
<td>Gseo_AS</td>
<td>ACATCAGTAGTATTGGTCTGATA</td>
<td></td>
</tr>
</tbody>
</table>

*Sequences were determined using the NAuRA tool (https://github.com/afelten-Anses/NAuRA), applying the Wunsch algorithm of Clone Manager Professional 9 software (Scientific & Educational Software, Denver, CO).*
The branch support was calculated by the bootstrap method, using 1,000 replicates. The graphic representation of the phylogeny was obtained by using iTOL web viewer (https://itol.embl.de/). The development of the ELISA solution (100 μl/well) containing tetramethylbenzidine (TMB; Thermo Fisher Scientific, Waltham, MA) was used for detection at RT for 30 min and washed 5 times again. Substrate solution (100 μl/well) containing tetramethylbenzidine (TMB; Thermo Fisher Scientific, Waltham, MA) then was added for 30 min. Finally, the reaction was stopped by addition of 100 μl of 4N hydrochloric acid (HCl). Absorbances were read at 450 nm on a microplate reader (SAFAS; Monaco). Quantification was performed by using a calibration curve based on the quadratic fit model. Validation data (sensitivity, specificity, and repeatability) of the above-described method are unpublished (Cécile Féraudet-Tarisse, Céline Goulard-Huet, Yacine Nia, Karine Devilliers, Dominique Marcé, Chloé Dambrune, Donatien Lefebvre, Jacques-Antoine Hennekinne, and Stéphanie Simon, unpublished data).

Statistical analysis. For analysis of potential correlation between production of SEG and SEI, a regression analysis was performed. First, the robust method was applied to verify the regression model and to identify outliers. Four outliers were identified and eliminated from the data set before calculating an ordinary least-square regression model.

To prove the two different levels of SEG and SEI production, a Kruskal-Wallis test was performed. For all statistical analyses, measured values under the limit of detection were taken as the value 0.

The graphical presentation of the enterotoxin data was performed using R (version 3.4.4) with the packages ggplot (75), ggsci (76), and ggpubr (77). With these packages, the data of the enterotoxin production of the single strains were plotted in increasing order of production (means ± standard deviations) and a color given according to their relative vSaβ type.

Data availability. Sequencing data for all isolates analyzed in this study have been deposited in the NCBI GenBank database under BioProject accession number PRJNA633807. Accession numbers for individual genomes and assembly statistics are listed in Tables S2 and S3.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 0.4 MB.

ACKNOWLEDGMENTS

This work was financially supported by the Agroscope research program. The work is a collaboration project between Anses and Agroscope. The development of the ELISA method for G and I was funded in part by the French joint ministerial program of R&D against CBRNE threats.
In addition, a particular thanks goes to L. Fritsch for support in using R as well as to N. Vingadassalon and I. Ivanovic for the help in their laboratories.

REFERENCES


